

J. Linguistics 58 (2022), 307–343. © The Author(s), 2021.

Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

doi:10.1017/S0022226721000311

The semantics of conversion nouns and *-ing* nominalizations: A quantitative and theoretical perspective¹

ROCHELLE LIEBER

University of New Hampshire

INGO PLAG

Heinrich-Heine-Universität Düsseldorf

(Received 08 December 2020; revised 13 July 2021)

This paper addresses a fundamental problem of derivational morphology: which meanings are possible for the words of a given morphological category, which forms can be chosen to express a given meaning, and what is the role of the base in these mappings of form and meaning? In a broad empirical study we examine the extent to which two types of nominalizations in English – conversion nouns and *-ing* nominalizations – can express either eventive or referential readings, can be quantified as either count or mass, and can be based on verbs of particular aspectual classes (state, activity, accomplishment, achievement, semelfactive). Past literature (for example, Grimshaw 1990 Brinton 1995, 1998 Borer 2013) has suggested an association between conversion nominalization, count quantification, and referential reading on the one hand, and between *-ing* nominalization, mass quantification and eventive reading on the other. Using a subset of the data reported in Andreou & Lieber (2020), we give statistical evidence that the relationship between morphological form, type of quantification, and aspectual class of base verb is neither categorical, as the literature suggests, nor completely free, but rather is probabilistic. We provide both a univariate analysis and a multivariate analysis (using conditional inference trees) that show that the relationship among the variables of morphological form, eventivity, quantification and aspectual class of base is complex. Tendencies sometimes go in the direction suggested by past literature (e.g. *-ing* forms tend to be eventive), but sometimes contradict past predictions (conversion also tends to be eventive). We also document that an important role is played by the specific verb underlying the nominalization rather than the aspectual class of verb. Finally, we consider what the pattern of

[1] The authors would like to acknowledge the anonymous reviewers for very helpful comments and suggestions. We are also grateful for the comments provided by the audiences at Heinrich-Heine-Universität Düsseldorf, December 9, 2020, and at the 43. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft, 23–26 February, 2021, Freiburg. This work has partly been funded by the Deutsche Forschungsgemeinschaft, Project ‘The semantics of derivational morphology’, awarded to Ingo Plag (grant PL 151/11-1), which we gratefully acknowledge.

polysemy that we uncover suggests with respect to theoretical modeling, looking at syntactic models (Distributed Morphology), lexical semantic models (the Lexical Semantic Framework), Analogical Models, and Distributional Semantics.

KEYWORDS: *-ing* nominalization, conversion, eventivity, quantification, aspectual class

1. INTRODUCTION

Andreou & Lieber (2020) have established that conversion and *-ing* nominalizations in English are capable of expressing a wide range of readings, both in terms of quantification (count versus mass) and in terms of eventivity (eventive versus referential readings), contrary to earlier claims in the literature (e.g. Grimshaw 1990 Brinton 1995, 1998 Borer 2013). Based on a large corpus of data drawn from the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC), Andreou & Lieber's study looks at both singular and plural conversion and *-ing* forms for 106 English verbs, covering the full range of aspectual classes (state, activity, accomplishment, achievement, and semelfactive). Andreou & Lieber (2020) argue that the morphological form of the nominalization does not in fact determine whether the nominalization will be construed as either count or mass, either eventive or referential. They argue also (pp. 334–335) that the aspectual class of the nominalization's base has little or no influence on either the form of the nominalization or its construal. Rather, they argue that context plays the most important role in determining the ultimate reading of either type of nominalization.

Here we build on the work of Andreou & Lieber and examine the readings of conversion and *-ing* nominalizations in further depth, looking specifically at how common it is for each morphological form to be construed as either count or mass, or as eventive or referential, and also whether the aspectual class of the nominalization's verbal base has any influence at all on the choice of reading. Although Andreou & Lieber have shown that the whole range of possible readings can be found for both types of nominalization, they do not answer what we might call the 'extent of polysemy' question, that is, whether a particular morphological form might have a PROPENSITY towards one or another reading. Here we take up that 'extent of polysemy' question and look at whether conversion nouns at least tend to be both count and referential in reading and *-ing* nouns both mass and eventive in reading, as had been suggested in prior literature. Further we look at the extent to which the individual verbal base of the nominalization influences its possible readings.

In what follows, we present a quantitative analysis of a carefully constructed subset of the Andreou & Lieber dataset, covering the same 106 base verbs. In this study we look specifically at singular conversion and *-ing* nominalizations in English, examining the extent to which morphological form is linked to count or mass quantification, to eventivity or to aspectual class of the base verb. The paper is organized as follows. Section 2 briefly reviews past literature in which issues of eventivity, quantification, and aspectuality in English nominalizations have been discussed, which leads us to pose our current research questions. Section 3 details the methodology with which we constructed our database and coded and analyzed

our examples. In Section 4 we present our statistical results. We show first that there is no categorical relation between conversion and count quantification, nor between *-ing* nouns and mass quantification. Although there is a strong tendency in this direction, there are nevertheless a significant number of tokens that have the unexpected quantification. We will also show in this section that there is a clear tendency for *-ing* nominalizations to express eventive readings, but referential readings are possible as well. More interesting is that conversion nouns actually favor eventive readings, albeit weakly. Fewer than half the conversion tokens we studied carried a referential reading. With respect to aspectual class, we confirm Andreou & Lieber's finding that there is almost no correlation between the aspectual class of the base verb and the readings available to the nominalizations derived from those verbs. In addition, we will show that the individual verbal base plays an important role in determining whether a conversion or *-ing* nominalization is quantified as count or mass, or construed as eventive or referential.

The most interesting question that our results raise is what this pattern means with respect to linguistic theory, the subject to which we turn in Section 5. There we argue that the statistical tendencies that we find in fact problematize the theoretical analysis of English nominalizations. Among currently available theoretical models, a syntactically based model like Distributed Morphology (DM) (Alexiadou 2001, Harley 2009, Alexiadou, Iordachioaia & Soare 2010, among others) predicts categorical behavior for nominalizations (that is, only a limited number of readings should be acceptable).² Lexical semantic models like Lieber's Lexical Semantic Framework (LSF; Lieber 2004, 2016) predict far freer variation in the readings of nominalizations than we actually find. Our results do not support predictions of either categorical behavior or relatively free variation in reading. The question, then, is how linguistic theory is to deal with behavior that is neither categorical nor free, but somewhere in between. The possibility that we propose is that frameworks like Analogical Modeling (e.g. Daelemans & van Bosch 2005, Skousen & Stanford 2007) and Distributed Semantics Models (DSM) (Boleda 2020, Marelli & Baroni 2015) might potentially be deployed in accounting for probabilistic behavior such as that displayed by these nominalizations.

2. LITERATURE REVIEW AND RESEARCH QUESTIONS

Over a period of several decades and across a number of theoretical frameworks, linguists have suggested that different morphological forms of nominalizations may be specialized for quantification and eventivity. There have also been passing claims in the literature that the aspectual class (also known as aktionsart, Vendler

[2] The same would hold for other syntactically based frameworks like Nanosyntax (Baunaz & Lander 2018) and Borer's (2013) Exoskeletal model.

class, lexical aspect, or inner aspect) of the verbal base has a role in circumscribing the readings available to the nominalization.³

In our analysis we will be primarily concerned with the relationships between morphological form (conversion versus *-ing*), eventivity (eventive versus referential interpretation), quantification (mass versus count), and aspectual class (state, activity, achievement, accomplishment, and semelfactive). We will not be concerned with issues of boundedness or perfectivity in interpretation, topics that have also been discussed in previous literature.

Looking first at quantification, the consensus over many decades has been that *-ing* nominalizations tend towards mass quantification; this association is mentioned very early on by Biese (1941), and can also be found in Grimshaw (1990), Langacker (1991), Brinton (1995, 1998), Alexiadou et al. (2010), Borer (2013), and Park & Park (2017). Biese also says explicitly that conversion nouns tend to be count quantified.⁴

Associations have also been drawn between aspectual class of base verb and quantification of nominalization. For example, Mourelatos (1978) suggests that nominalizations derived from state and activity verbs are mass quantified, while those derived from accomplishments and achievements are count quantified. Brinton (1995) agrees with Mourelatos that state verbs are mass quantified. Bauer, Lieber & Plag (2013) suggest an association between aspectual classes that are punctual or instantaneous (that is, achievements and semelfactives) and count quantification.

Moving on to the association between aspectual classes and morphological form of nominalization, both Brinton (1995) and Grimshaw (2011) suggest that state verbs prefer conversion nominalizations. Brinton (1995) suggests as well that *-ing* nominalizations prefer activity verbs. Alexiadou (2001) and Borer (2013) suggest that accomplishments and achievements (that is, telic verbs) disprefer *-ing* nominalizations, although Iordachioaia & Werner (2019) disagree on this point.

Finally, with regard to the association between eventivity and morphological form, Grimshaw (1990) argues that *-ing* nominalizations are always interpreted as complex events,⁵ whereas conversion nouns cannot be interpreted as complex events. For Grimshaw (1990: 49), conversion nouns can be interpreted as results or as simple events, but to the extent that she discusses simple event nouns at all they are treated in the same way as result nouns. Subsequent literature subsumes nominalizations with simple event interpretation into the category of referential nouns (see Roy & Soare

[3] In both the literature cited and in what follows here, we look at what is called by Filip (2012: 721) the 'lexical aspect' of verbs. We assume, with Andreou & Lieber (2020: 339), that verbs frequently have a default lexical aspect which can of course be modified, depending on syntactic context, such as when a default activity verb like *push* (*I pushed for hours*) is construed as an accomplishment in a sentence like *We pushed the cart to the wall* (see also Verkuyl 1972, Croft 2012, among others). Since we deal here with verbs only as bases of nominalizations, and not as they appear in wider syntactic contexts, it makes sense to restrict our categorization to this default lexical aspect.

[4] Not all of these authors comment specifically on the properties of conversion nouns.

[5] By complex events, Grimshaw (1990, 45) means nominalizations that have obligatory argument structure.

2011: 9). Similar claims can be found in such works by Roeper (1993), Alexiadou (2010, 2011), Borer (2005, 2013) and Fabregas & Marin (2012), among others.⁶

Putting all of this together, we should expect to find a general clustering of characteristics with each morphological form: conversion nouns should be count and referential, and *-ing* nouns mass and eventive. What sort of associations we should expect to find between morphological forms and individual aspectual classes of base verbs does not emerge clearly from the literature, nor does any of the past literature address the role of individual verbal bases in the construal of different kinds of nominalization.

This brings us to the most recent study to look at associations between morphological form, eventivity, quantification, and aspectual class. Andreou & Lieber (2020) is the most empirically grounded of the studies that examine the relationship between the morphological form of a nominalization and its ability to express count or mass quantification and eventive or referential readings. Based on an extensive corpus of conversion and *-ing* nominalizations in context, their study documents that there is in fact no necessary relationship between conversion nominalization, count quantification, and referential reading, nor is there a necessary relationship between *-ing* nominalization, mass quantification, and eventive reading. Different morphological forms of nominalization do not show strict preferences for bases from any particular aspectual class of verbs. Indeed, they show that individual conversion or *-ing* nominals can express any combination of quantification, eventivity, or aspectual class.⁷ What Andreou & Lieber do not do, however, is to document the frequency with which various combinations of readings arise, or whether there is any statistical correlation between the type of verb on which the nominalization is based, the appearance of mass or count quantification, or the reading as eventive or referential. This is what we set out to do in the present study.

In what follows we pose the following questions. First, we explore the extent to which conversion nominalizations tend towards readings that are count and referential and *-ing* nominalizations towards readings that are mass and eventive, as has been claimed in past literature. That is, although we know from Andreou & Lieber (2020) that all combinations of these semantic attributes are possible with either type of nominalization, we seek to explore whether some combinations of semantic attributes with morphological category are more likely than others. Second, we seek to discover what, if anything, is the relationship between the aspectual class of base verbs and the quantification and eventivity of the resulting nominalization. Is it the case, as Mour-elatos (1978) suggests, that nominalizations of activity and state verbs prefer mass quantification? Is it the case, as Brinton (1995) suggests, that *-ing* nominalizations

[6] It should be noted, however, that several works including Cetnarowska (1993), Newmeyer (2009) and Lieber (2016) have shown that this simple association of *-ing* nominalizations with eventive readings and conversion with referential readings does not hold.

[7] Andreou & Lieber (2020) also consider whether conversion nouns are restricted to aspectually bounded readings and *-ing* nouns to aspectually unbounded readings, a distinction that we will not touch on here.

prefer activity verbs as bases? Is it the case, as Borer (2013) says, that *-ing* nominalizations tend to be formed only on atelic bases?

The answers to these questions matter with respect to morphological theory. For example, if it turns out to be the case that the association between conversion, referentiality, and count quantification or between *-ing* nominalization, eventive reading, and mass quantification is at least close to categorical, this would speak in favor of deterministic models such as syntactic analyses in frameworks like Distributed Morphology (Alexiadou 2001) or the Exoskeletal model (Borer 2013). If, on the other hand, the associations between morphological form and these various semantic characteristics are highly variable, this would favor lexical models such as Lieber's (2004, 2016) Lexical Semantic Framework. If the pattern of data we uncover is neither clearly categorical nor largely free, this suggests that other approaches to modeling the relation between form and meaning need to be considered.

3. METHODOLOGY: CONSTRUCTION AND CODING OF THE DATABASE, STATISTICAL ANALYSIS

3.1 *Constructing the database*

Exploring the extent to which relationships exist between morphological form, quantification, and eventive versus referential readings demands a database amenable to statistical analysis. To achieve this, we selected and analyzed a subset of the database used in the Andreou & Lieber study. The Andreou & Lieber dataset comprised more than 57,000 singular and plural tokens of conversion and *-ing* nominalizations from 106 base verbs.⁸ This massive data set was reduced to a manageable size in several steps.

The first step was to focus on the singular forms, for two reasons. Our aim was to compare forms with the same meaning potential. Plural forms, however, are always count quantified while singular forms have the potential to be either count or mass quantified. Having only singulars means comparing like with like. Second, we wanted to avoid complications that may arise with abstract nouns through the complex interaction of plurality, cumulativity and distributivity on the one hand, and the count/mass distinction and aspectuality on the other. In a nutshell, the pluralization of an abstract noun can lead to differences in their interpretation vis-à-vis the singular form that go beyond the mere one-many distinction (see Sutton &

[8] To construct this larger dataset, Andreou & Lieber (2020: 342) first chose 98 verbs from the 1000 most frequent verbs listed in the COCA frequency list, looking for verbs from all frequency ranges that exhibited both *-ing* and conversion nominalizations. To this list they added eight low frequency verbs. Conversion pairs were restricted to those in which verb was attested earlier than the noun in the OED. Up to 300 tokens were extracted from COCA for each of the singular and plural conversion and *-ing* nominalizations for each verbal base, and tokens were added from the BNC if fewer than 300 tokens were available in COCA. After standard data-cleaning procedures (removing tokens that contained typos, misspellings, and highly lexicalized forms), the resulting dataset contained approximately 57,700 tokens.

Filip 2019 for discussion). By restricting ourselves to singular forms we avoided this potential confound.

We then chose every eighth token of the singular conversion nominalizations and similarly every eighth token of the singular *-ing* nominalizations. If after taking every eighth token for conversion and *-ing* nominalizations derived on a particular verbal base there were fewer than ten examples for the conversion or *-ing* forms, we supplemented so that there were at least ten examples of each. If the original sample had fewer than ten examples for a morphological form for a particular verb to begin with, we kept however many there were in the original sample.⁹ This procedure yielded a sample of 4282 tokens, which we then coded as we describe in Section 3.2.

3.2 Coding of the database

Tokens were hand-coded by both authors concerning eventivity, quantification and aspectual class. The coding was based on the following criteria:

3.2.1 Count versus mass

In coding each token as count or mass, we looked for positive evidence of quantification in the surrounding context. Positive evidence could be any of the following:

(1) *Positive evidence for count quantification*

- A determiner *a/an* precedes the noun.
- There is an ordinal like *first, second*, etc. or the cardinal *one*.
- There is a modifier like *final, last*, or *single*.
- There is a quantifier like *many, each, every*.

(2) *Positive evidence for mass quantification*

- The nominalization is singular and there is no determiner preceding it.
- There is a quantifier like *much, some*.

(3) provides illustrative examples:¹⁰

(3) (a) Conversion count (evidence: indefinite article)

Accent on Living 1992: And if enough people tell you you're no good, you end up believing it and become what they say you are. A **change** has to come.

(b) Conversion mass (evidence: singular form, no determiner)

[9] In the resulting data base 17 base verbs (of 106) have 10 or less attestations of conversion nouns, and 52 verbs (of 106) have 10 or less attestations of *-ing*. Overall, the number of tokens per verbal base ranges between one and 37 tokens for conversion nouns, and between none and 37 tokens for *-ing* nouns (medians: 28 for conversion nouns, 11 for *-ing* nouns).

[10] The nominalization is set in bold, with the relevant part of the context underlined.

Independent School 2006: Overall, then, the single greatest obstacle to implementing **curricular change** and, over time, establishing a culture that values continuous reflection and improvement in a school, is the general predisposition of educators to resist change itself.

- (c) *-ing* count (evidence: indefinite article)
BNC 1992: ... tiny phosphorescent sparks around its hands, small ripples in the stone beneath its feet, a gentle breeze around its head, a sudden dampness and **dripping** of water from the stones of the walls around it.
- (d) *-ing* mass (evidence: singular form, no determiner)
BNC 1978: Listen to the rain. Compare slow **dripping**, fast gushing, trickling, etc.

One issue that arose in our coding was that there were quite a few examples in which no positive evidence for count or mass quantification was available in context. For instance, consider the example in (4):

- (4) *No positive evidence for quantification*
BNC 1993: Inside the wall there was a new blackness, free of mist, and the most persistent sound was the **dripping** and trickling of myriad streams into the storage below.

Because the determiner in this example is *the* and *the* occurs with both count and mass nouns (*the squirrel* and *the milk* are equally possible), there is no overt evidence of quantification in such cases. Rather than deciding on an impressionistic basis whether we thought the use of the nominalization was count or mass quantified in a given token, we chose to disregard such examples from all statistical analyses of this variable. The remaining data amounted to a total of 2379 tokens of conversion and *-ing* nominalizations with unambiguous evidence for mass or count readings.

3.2.2 *Eventive versus referential*

Coding for eventive versus referential is complicated, as there are all sorts of cues to either interpretation, and many of them are subtle. Some are syntactic: for example, modification by adjectives like *frequent* or *repeated* suggest eventivity. Others are semantic, based on paraphrases that seem appropriate in a given context. The following is a list of cues that we used in coding particular tokens as eventive or referential, and a selection of examples that illustrate some of those cues.

- (5) *Positive evidence for E readings*
- (a) SYNTACTIC: kinds of PPs, possessives or other modifiers, synthetic compounds
- The context shows a tensed verb that suggests an eventive reading for the nominalization.

- There is a temporal adjectival or adverbial modifier like *frequent*, *continuous*, *repeated*, etc. that suggests an eventive reading.
- There is a possessive subject or a *by* phrase with an agentive interpretation.
- There are other PPs that indicate participants in an event (e.g. *of* PPs).
- There are manner modifiers like *intentional*, *personal*, or *nervous* that suggest eventivity.
- There is a temporal preposition like *after* or *before* that suggests eventivity.
- The nominalization occurs as the second element in a synthetic compound and the first element is construed as the internal argument of the base verb.

(b) SEMANTIC: paraphrases that suggest eventivity¹¹

- The feeling of (*worry*)
- The state of (*worry*)
- The activity of (*climb*)
- An instance of (*drive, kiss, win*)

(6) *Positive evidence for R readings*

(a) SYNTACTIC

- coreference with another noun in the context that is clearly concrete

(b) SEMANTIC: paraphrases that suggest referential reading (N.B. the verbs listed with each reading often have different referential readings in different contexts)

- A kind of (*blend*)
- A time span (*break*)
- An entity, product, copy product, or information object (*report, smell*)
- A place or location (*pass, car-wash*)
- An amount, measure, or extent (*reach, pay, return, spread*)
- The thing, stuff or person V-ed (*catch, hire*)
- The name for a kind of a feeling (*concern*)¹²
- A symptom or condition (*cough*)
- The name for a kind of hobby, pastime, occupation, or job (*rock-climbing, medical assisting, bungee jumping*)
- An instrument (*drip, punch*)

[11] By 'paraphrases that suggest eventivity' we mean paraphrases that we felt could be substituted and convey the contextually relevant reading.

[12] Cetnarowska (1993: 26) characterizes the construal we mean here as 'feeling as a general phenomenon'.

- The name of a sound (*gurgle, knock, slam, tap*)
- A quality (*flash, float, cast, heat, good repair*)¹³
- A manner (*hold*)¹⁴

(7) gives some examples to illustrate the sort of syntactic cues and paraphrases we used to code for eventivity and (8) some examples of syntactic cues or paraphrases that suggested referential readings:

- (7) (a) Eventive reading: verb in context suggests eventivity
Massachusetts Review 1995: ... a muffled banging began, a thumping, not rapid but steady, like the **drip** of water on a slab, a noise as though someone above was stamping one foot, heavily booted, on a bare board floor.
- (b) Eventive reading: temporal adverbial modifier suggests eventivity
Massachusetts Review 1995: ... a muffled banging began, a thumping, not rapid but steady, like the **drip** of water on a slab, a noise as though someone above was stamping one foot, heavily booted, on a bare board floor.
- (c) Eventive reading: prepositional phrases suggesting participants in an activity (of PPs, by PPs)
Christian Science Monitor 2004: It was discovered by accident in January 1936, the same year as the first **find** of a Clovis point, when a Works Progress Administration crew was digging a drainage ditch and unearthed most of a foreleg from a juvenile mammoth.
- (d) Eventive reading: manner adjective or adverb suggests eventivity
Shawshank Redemption 1994: Norton keeps his eyes on Andy, looking for a wrong glance or nervous **blink**.
- (e) Eventive reading: temporal preposition suggests eventivity
Field and Stream 2005: When you see a doe break out of cover during a **drive**, get ready, as the buck is likely just a few steps behind.
- (f) Eventive reading: nominalization is part of a synthetic compound that suggests eventivity
Prevention 1997: And getting more large muscle groups like these into the act means greater calorie **burn**.
- (g) Eventive reading: paraphrase suggests eventivity (an instance of)
Bicycling 1999: This was only the first **climb**, but Armstrong had proven himself.
- (h) Eventive reading: paraphrase suggests eventivity (the feeling of)
Rolling Stone 1990: There was **worry** on the boy's part about his academic work. 'There was Logan ordering him upstairs to his room, only to appear in a state of undress'.

[13] For example, *float* as a property of something: *The tube's float was superior to that of any other raft.*

[14] For example: *Her hold was technically superior to that of other rock-climbers.*

- (8) (a) Referential reading: coreference with a clearly concrete noun in syntactic context
Mag. Inc. 1995: “Today,” says Holman, “employees in the cleaning department, for example, know not to work on the easiest-to-clean **casting** or on the one that happens to be on the top of the pile.”
- (b) Referential reading: paraphrase suggests a kind of activity
New York Times 2012: Around Moab, the bureau, part of the Department of the Interior, decrees special-recreation permits for activities including river rafting, **rock climbing**, mountain biking, four-wheeling and hot-air ballooning.
Associated Press 2012: Santorum said Obama and his allies want to frighten people about alleged dangers of petroleum-extraction techniques, including **hydraulic fracturing**, or “fracking,” which might lower energy prices.
- (c) Referential reading: paraphrase suggests that the nominalization refers to a participant in an event (e.g. instrument)
Consumer Research Magazine 1997: Electric resistance systems, such as electric baseboard **heating**, also require thermostats capable of directly controlling 120 volt or 240 volt line-voltage circuits.

Given that the coding of eventive vs. referential readings is not always straightforward, we implemented a very conservative strategy. Each observation was rated independently by the two authors using the criteria explained and discussed above. Four values were used for this variable: eventive, referential, ambiguous and unclear. The statistical analysis of the rater agreement showed very high agreement between the two raters. The authors agreed in 78.1 percent of the cases using the four values (Cohen’s Kappa for 2 Raters: 0.584, $z = 48.5$, $p = 0$). The truly contradictory ratings, i.e. those where one rater said ‘eventive’, the other ‘referential’) were only 7.8 percent of the cases. We decided to restrict the analysis presented in this paper to those tokens where both authors rated a given token as eventive or both authors rated it as referential. Items where at least one rating was ‘ambiguous’ or ‘unclear’ were disregarded. This reduced the data set for the analysis of eventivity to 3286 items.

3.2.3 Aspectual class of base verb

In coding the verbal bases of conversion and *-ing* nominalizations for aspectual class, we looked solely at default lexical aspect, that is, at those elements of meaning which are contributed by the verb alone, without considering surrounding syntactic context.¹⁵ As has been done before in the literature (Smith 1991, Brinton 1995,

[15] For example, we considered a verb like *push* to be an activity, even though in the context of a sentence like *We pushed the cart into the shed*, it could be construed as having an endpoint, and

among others), we considered each aspectual class to be characterized by binary features, as shown in (9):¹⁶

(9)	ASPECTUAL CLASS	EXAMPLE	DYNAMIC	DURATIVE	IMPLIED ENDPOINT
	State	<i>know, love</i>	–	+	–
	Activity	<i>push, float</i>	+	+	–
	Accomplishment	<i>cook, cover</i>	+	+	+
	Achievement	<i>arrive, find</i>	+	–	+
	Semelfactive	<i>blink, knock</i>	+	–	–

Each verbal base in the sample was coded for aspectual class. A verb was coded as [+dynamic] if it implied an input of energy, [–dynamic] if not (Comrie 1976). Verbs of change of state, directed change of position, or verbs which otherwise implied an outcome or product were coded as [+implied endpoint]. Other verbs were coded as [–implied endpoint]. Verbs which implied a passage of time were coded as [+durative] and verbs that implied a punctual event were coded as [–durative].

For the majority of our verbal bases it was possible to identify a relatively stable core meaning that can be classified as one of the five aspectual classes detailed in (9). Some verbs, however, display nuances of reading or multiple senses even out of context that required us to place them in more than one aspectual class. For example, several verbs in our sample are ambiguous between state readings and activity readings (that is, the verb *taste* as what something tastes like as opposed to the act of tasting), and were therefore coded [+/-dynamic], indicating that they could go either way with respect to this feature; those verbs appear in the chart below as act/state for ‘activity or state’. There was one verb, *pass*, for which an endpoint could be implied or not, depending on the sense of the verb in context. Specifically, *pass* can be an activity verb in the context of football, but an achievement in the context of exam-taking. This verb was coded as [+/- implied endpoint] and appears as act/ach for ‘activity or achievement’ in the chart below. Again, with respect to the feature [durative], in cases where punctuality was not necessarily clear, we coded the verbal base as [+/- durative]; for example, does the verb *exit* refer exclusively to the moment of leaving or does it include the motion towards the leaving-point? Such verbs appear in Table 1 as acc/ach for ‘accomplishment or achievement’ or as sem/act for ‘semelfactive or activity’ depending on whether the verbs involved implied an endpoint, as would be the case for accomplishments or achievements, or not, as would be the case for semelfactives or activities.¹⁷

therefore being construed as an accomplishment. See Andreou & Lieber (2020) for a justification of this move.

[16] The feature [+/-telic] is often used instead of [+/-implied endpoint]. We prefer the latter, as it makes clear that what is relevant is that an endpoint may be lexically implied, although whether that endpoint is reached or not might only be supplied in context. The term ‘telic’ is often ambiguous between those two possibilities.

[17] Croft (2012) makes a good case for the need to multiply categories of lexical aspect beyond the five that we make use of here. However, as much of the literature that we are responding to makes use of those five classes, we stick with them here.

Accomplishment/achievement	bet, cast, cut, divide, draw, exit, grant, hire, offer, order, pay, return, shift
Accomplishment	bend, blend, burn, change, cover, display, fall, fix, form, heat, melt, mix, repair, report, rise, spread, strip, surround, take, transfer, wash
Achievement	break, burst, catch, chop, find, finish, fracture, kill reach, reveal, slam, split, win
Activity/achievement	pass
Activity/state	smell, taste
Activity	account, assist, chat, climb, dance, design, drink, drive, embrace, float, gurgle, hold, hurt, kiss, leak, play, push, rest, ride, run, sail, spin, stay, step, stir, stretch, stumble, swim, wait
Semelfactive/achievement	pop, strike
Semelfactive/activity	beat, shake, shove
Semelfactive	blink, cough, drip, flash, hit, jump, kick, knock, poke, punch, spring, tap
State	concern, desire, doubt, fear, hate, hope, lack, love, stink, worry

Table 1

List of verbs coded by aspectual class.

We limited our analysis of the role of aspectual class to those verbs whose aspectual class was unambiguous, which resulted in a data set for aspectual analysis of 3405 tokens.

3.3 *Statistical analysis*

The statistical analyses were carried out using R (R Core Team 2020). After an overview of the distributions of relevant variables in the different data sets, we provide a univariate analysis of all pair-wise associations between pertinent variables of interest (morphological form, quantification, eventivity, aspectual class, aspectual feature, base verb). This will be followed by multivariate analyses that take all variables into account at the same time. The details of the analyses are introduced and explained as we go along.

4. RESULTS

4.1 *Overview*

Overall, our data sets show the distributions of forms and meanings as shown in Table 2.

Claims about the relation of form and meaning in nominalizations have come in two flavors, semasiological and onomasiological. The statistical analysis will also take these two perspectives. In the semasiological perspective we will ask which

Semantic dimension	Semantic category	Number of tokens
Quantification	mass	1079
	count	1282
Eventivity	referential	982
	eventive	2304
Aspectual classes	accomplishment	928
	achievement	548
	activity	1152
	semelfactive	421
	state	356

Table 2
Distribution of forms and meanings.

semantic properties conversion and *-ing* nominalizations can express. In the onomasiological perspective we investigate which formal means (conversion or *-ing*-suffixation) speakers and writers choose in order to express a particular meaning. We will first look at each semantic dimension independently in a univariate analysis. To assess the variability brought into the game by the individual base verb and its nominalization(s), we also analyze the variation within and across individual base verbs and their nominalizations. Finally, we provide multivariate analyses in which we take all variables into account at the same time (i.e. all semantic dimensions and the role of the individual base word).

4.2 Univariate analyses

4.2.1 *Mass vs. count*

Here we look at the extent to which morphological form and type of quantification are interdependent, with the two questions:

- The semasiological view: Do conversion nouns tend to be count in quantification while *-ing* nominalizations tend to be mass?
- The onomasiological view: Are count meanings expressed by using conversion, and mass meanings by using *-ing* nominalizations?

Our analysis shows that there is indeed statistically significant tendency in these directions (chi-squared = 435.7, $df = 1$, $p < 2.2e-16$, Cramer's $V = 0.43$), as illustrated in the mosaic plots in [Figure 1](#). In mosaic plots the size of the area represents the number of observations in that combination of variable values. [Table 3](#) shows the same thing giving percentages.

CONVERSION NOUNS AND *-ING* NOMINALIZATIONS

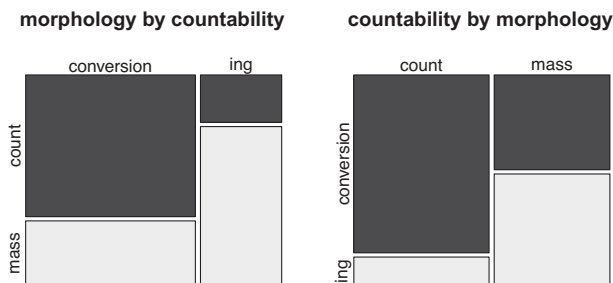


Figure 1
Morphology and quantification, N = 2379.

Semasiological			Onomasiological	
	Conversion	<i>-ing</i>	Count	Mass
Count	69	23	Conversion	86
Mass	31	77	<i>-ing</i>	14
Total	100	100	Total	100

Table 3
Distribution of countability and morphology, in percentages (N = 2379).

It is important to note that the association of form and meaning is by no means categorical. The left panel of Figure 1 shows us the semasiological view: here almost one third of all conversion nouns have a mass interpretation, and almost a quarter of the *-ing* nominalizations are count nouns. From an onomasiological perspective (right panel of Figure 1) we can say that the vast majority of count meanings (86 percent) are expressed by conversion, while mass meanings are almost equally expressed by the two morphological categories (46 percent conversion, 54 percent *-ing*).

4.2.2 Eventive vs. referential

In this section, we examine the extent to which there is a link between eventivity and morphological form, that is, whether eventive readings are strictly associated with *-ing* nominals and referential readings with conversion nominals.

Figure 2 and Table 4 give the distributions of the two variables. From the semasiological point of view (left panels), we find that there is indeed a strong tendency (89 percent) for *-ing* nominalizations to express eventive readings. However, rather unexpectedly given the claims in the literature, conversion nouns also have a slight majority of eventive readings; only 44 percent of the conversion nouns are referential. From the onomasiological perspective (right panel), to express an eventive reading the two morphological means are almost equally likely (47 percent conversion as against 53 percent *-ing*), whereas referential readings

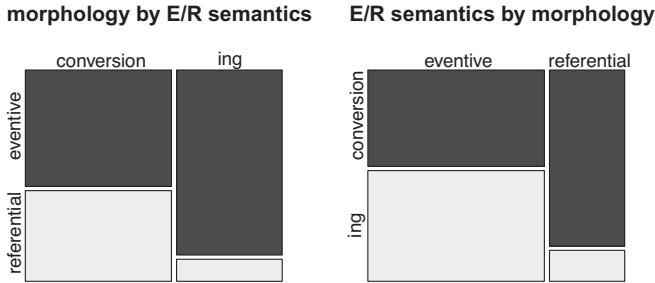


Figure 2 Morphology and eventivity, N = 3286.

	Conversion	-ing	Eventive	Referential
Eventive	56	89	47	85
Referential	44	11	53	15
Total	100	100	100	100

Table 4 Distribution of eventivity and morphology, in percentages (N = 3286).

have a strong tendency to select conversion nouns (85 percent). The differences in distribution are statistically significant (chi-squared = 417.53, *df* = 1, *p* < 2.2e-16, Cramer’s *V* = 0.36).

4.2.3 Aspectual classes

Next, we seek to discover whether there is any relationship between the aspectual class of base verbs and the quantification and eventivity of the resulting nominalization. Is it the case, as Mourelatos (1978) suggests, that activity and state verbs prefer mass quantification? Is it the case, as Brinton (1995) suggests, that *-ing* nominalizations prefer activity verbs as bases? Is it the case, as Borer (2013) says, that *-ing* nominalizations tend to be formed only on atelic bases?

We start with the analysis of the relation of aspectual class and morphological form.

The overall differences in the distribution of morphology and aspectual class are statistically significant (chi-squared = 114.46, *df* = 4, *p* < .001, Cramer’s *V* = 0.18). Both morphological forms can take verbs of all categories with similar proportions across the five aspectual classes. The only notable difference occurs with state verbs. Conversion has a sizable number of state verbs as bases, *-ing* very few. This difference becomes even clearer in the right panel of Figure 3. State verbs overwhelmingly take conversion as the preferred nominalization strategy.

In Figure 4 we can see that both types of quantification go together with all kinds of base verbs, and that differences mainly concern achievement verbs,

CONVERSION NOUNS AND *-ING* NOMINALIZATIONS

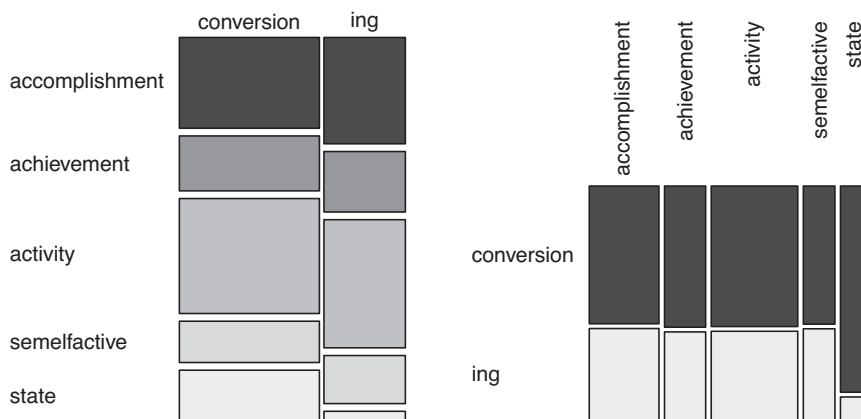


Figure 3

Morphology by aspectual class (N = 3405, intersecting set of all morphology-coded nominals and all unambiguously aspectual class-coded nominals).

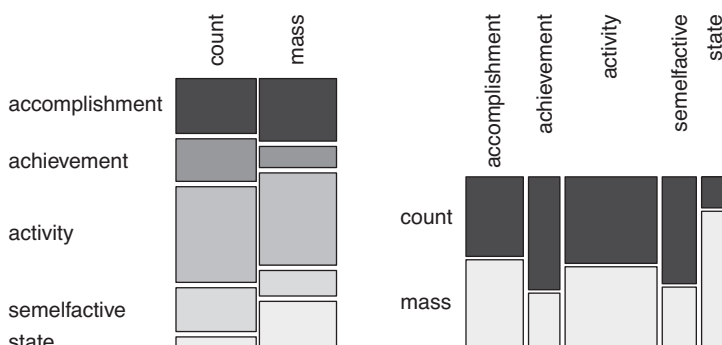


Figure 4

Quantification by aspectual class (N = 1934, intersecting set of all non-default-countability-coded nominals and all unambiguously aspectual class-coded nominals).

semelfactives and state verbs. While achievements and semelfactives are more strongly associated with count interpretations, state verbs are more strongly associated with mass interpretations of their nominalizations. These differences are statistically significant (chi-squared = 139.4, $df = 4$, $p < 2.2e-16$, Cramer's $V = 0.27$).

Finally, Figure 5 shows that both eventive and referential interpretations of nominals can be found with any kind of base verb aspectual class. The proportion of accomplishment verbs among referential nominals is highest. Conversely, the proportion of referential interpretations is highest among the nominals based on accomplishments. The differences in the distribution are statistically significant (chi-squared = 157.2, $df = 8$, $p < 2.2e-16$, Cramer's $V = 0.25$).

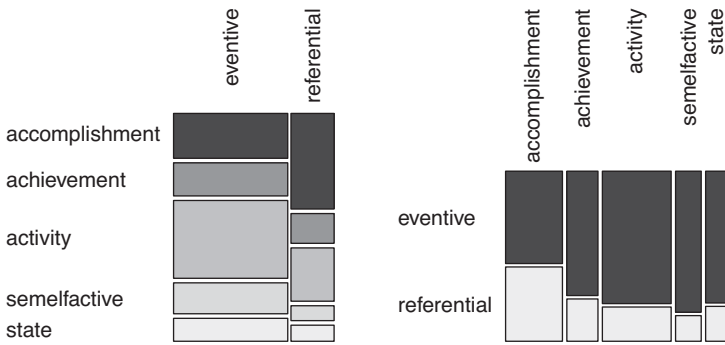


Figure 5

Eventivity by aspectual class (N = 2568, intersecting set of all unanimously eventivity-coded nominals and all unambiguously aspectual class-coded nominals).

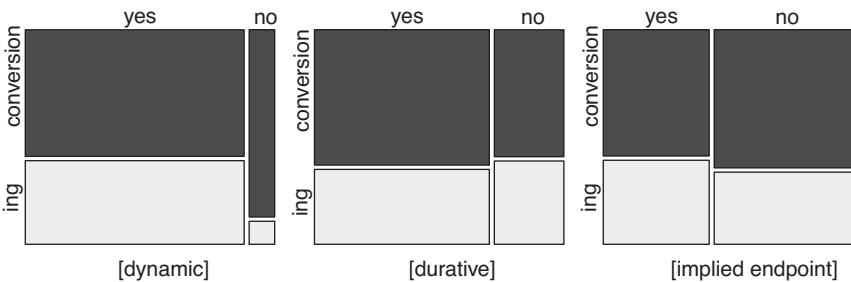


Figure 6

Aspectual features by morphology ('yes' indicates a positive value for the feature, 'no' a negative value for the feature, N = 3405).

4.2.4 Aspectual features

The data sets for the analysis of the aspectual features are the same as in the analysis of the aspectual classes in the previous section.

4.2.4.1 Morphology

Figure 6 shows the distributions of these features by morphological category.

The difference in distribution of aspectual feature by morphology is significant for each feature ([dynamic]: chi-squared = 112.71, $df = 1$, $p < .001$, [durative]: chi-squared = 4.5916, $df = 1$, $p = .03$, [endpoint]: chi-squared = 11.258, $df = 1$, $p < .001$), but the effect sizes are tiny for [durative] and [implied endpoint] (Cramer's V : 0.04 and 0.06, respectively), as compared to the at least small effect size for the [dynamic] feature (Cramer's $V = 0.18$). This means that only the feature [dynamic] is really associated with a morphological category: nominalizations of stative verbs tend strongly towards conversion, while the other two aspectual features do not show a pronounced trend towards a particular form of nominalization.

4.2.4.2 Quantification

The relationship between aspectual features and quantification is significant for all features ([dynamic]: chi-squared = 100.02, $df = 1$, $p < .001$, [durative]: chi-squared = 62.673, $df = 1$, $p < .001$, [endpoint]: chi-squared = 6.2654, $df = 1$, $p = .01231$, Cramer's V : 0.23, 0.18, 0.06 for [dynamic], [durative] and [implied endpoint], respectively). The feature [dynamic] has the strongest influence. Figure 7 illustrates this.

4.2.4.3 Eventivity

The relationship between aspectual features and the semantic interpretation of the derivative is significant for all features ([dynamic]: chi-squared = 5.7, $df = 1$, $p = .02$, [durative]: chi-squared = 21.1, $df = 1$, $p < .001$, [endpoint]: chi-squared = 105.6, $df = 1$, $p < .001$, Cramer's V : 0.05, 0.09, 0.20 for [dynamic], [durative] and [implied endpoint], respectively). Figure 8 illustrates this.

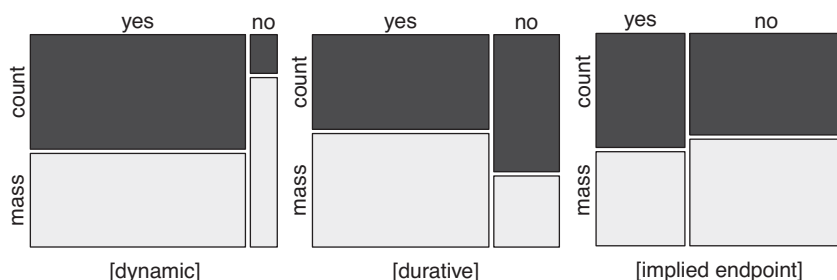


Figure 7

Aspectual features by quantification ('yes' indicates a positive value for the feature, 'no' a negative value for the feature, $N = 1934$).

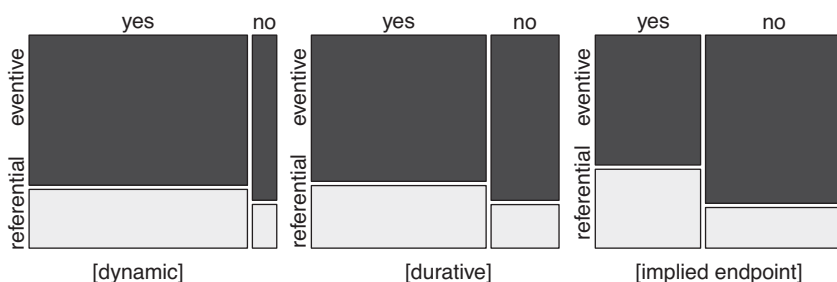


Figure 8

Aspectual features by eventivity ('yes' indicates a positive value for the feature, 'no' a negative value for the feature, $N = 2568$).

4.2.5 *Variation within and across base verbs and their nominalizations*

The variability in the choice of nominalizations and the expression of eventivity and quantification can also be seen when we look at individual base verbs. Almost all verbs are attested with both kinds of nominalization, but individual verbs often exhibit rather different profiles with respect to quantification and eventivity.

First, we note that individual verbs can vary greatly with respect to quantification. We find some verbs whose conversion forms are always count quantified (for example, *bet*, *blend*, *gurgle*). For some verbs we find that their *-ing* nominalizations are always mass quantified (*betting*, *hitting*, *kissing*). And for still other verbs we find that their conversion and *-ing* nouns are more evenly split with respect to quantification (*float*, *fracture*, *cutting*, *drawing*). We also find some verbs that tend toward mass quantification for conversion forms (*hate*, *change*) or count quantification for *-ing* forms (*covering*, *offering*), exactly the opposite of what the literature predicts.

With regard to the distribution of eventive vs. referential readings by verb, we find a similar picture. There are verbs like *design*, whose two nominalizations, *design* and *designing*, show a neat complementary distribution of form and meaning, but many other bases are much more flexible in both directions, for example *cut* and *dance*. Table 5 gives the token counts for the nominalizations for these three base verbs to illustrate this point.

Let us look at the distribution of eventivity across tokens to illustrate the problem on the large scale of the whole data set. For each verb we can derive a proportion of tokens with an eventive reading for each morphological form. In an ideal categorical world, all verbs should behave like *design*, with a proportion of 0 eventive readings for the conversion nominal and a proportion of 1.0 eventive readings for the *-ing* nominalization.

Figures 9 and 10 give an overview of the within-type variability of eventive vs. referential interpretations. Figure 9 shows the distribution of the proportions for conversion nouns, Figure 10 for *-ing* nouns.

We can see that 18 verbal bases have a proportion of 0 eventive readings in their conversion nouns. The nouns *account* and *blend* belong to this group. At the other end of the plot there is, perhaps surprisingly, an even larger group of 27 conversion nouns that are attested exclusively with eventive readings. The nouns *blink*, *burst*, *embrace* are examples from this set of forms. Hence, taken together, 45 out of

	CUT		DANCE		DESIGN	
	<i>cut</i>	<i>cutting</i>	<i>dance</i>	<i>dancing</i>	<i>design</i>	<i>designing</i>
Eventive	3	5	5	11	0	9
Referential	10	7	4	2	23	0

Table 5
Distribution of E/R readings by verb and nominalization.

CONVERSION NOUNS AND *-ING* NOMINALIZATIONS

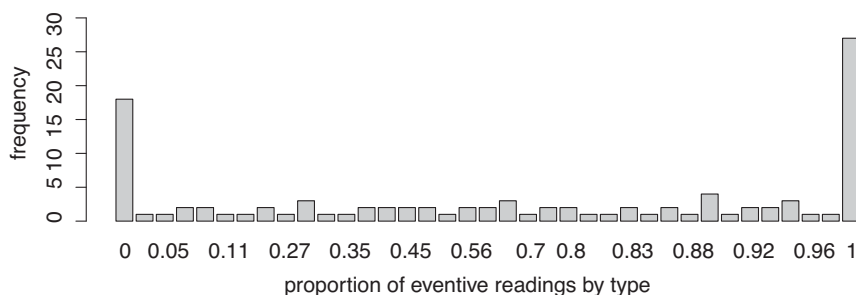


Figure 9

Frequency of proportions of eventive readings of conversion nouns by verbal base (type).

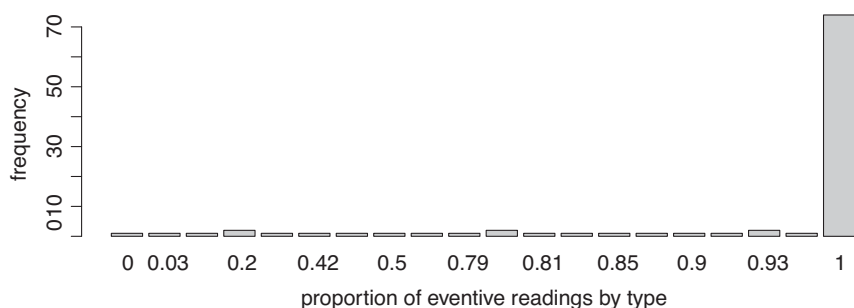


Figure 10

Frequency of proportions of eventive readings of *-ing* nominals by verbal base (type).

103 conversions (i.e. types) show no variation in their interpretation, and sixty percent of these invariably interpreted conversion nouns have an eventive reading. The majority of conversions lend themselves to both interpretations (to varying degrees). Overall, 62 of the 103 conversion nouns have a proportion of eventive readings larger than 0.5, which means that they have more tokens with eventive readings than with referential readings. The association between conversion and referentiality is thus chiefly a matter of individual lexemes.

The distribution of readings for *-ing* nominals depicted in Figure 10 tell a different story. The vast majority of *-ing* nouns show no variation across their tokens. They are always interpreted as eventive. The nouns *beating*, *blending* and *blinking* are examples from this very large set of nouns. Only one noun, *running*, occurs exclusively in combination with verbs or modifiers that suggest a referential reading ('type of sport'), as in *marathon running*, *road running* or *river running*. (10) shows the attestation of *river running* for illustration:

- (10) *Chicago-Sun Times 1995*: They taught us the fine art of river **running**

Although the majority *-ing* nouns show a categorical behavior with regard to eventivity, many *-ing* nouns do not, and allow for both readings (see again [Table 5](#) above).

In order to assess the magnitude of the base verb effect quantitatively we devised three mixed effect regression models with MORPHOLOGY, QUANTIFICATION and EVENTIVITY as dependent variables, respectively, and the random effect of VERB as the only predictor. We used the R packages *lme4* (Bates et al. 2015) and *MuMin* (Barton 2009) for this analysis. We computed the R^2 values (using the `r.squaredGLMM()` function of the *MuMin* package) to see how much of the variance is explained by VERB. In the model predicting morphological form the theoretical R^2 was 0.22, in the model predicting quantification the theoretical R^2 was 0.39, and in the model predicting eventivity the theoretical R^2 was 0.61. This shows that the base verb plays a significant, but moderate, role in the choice of the morphological form. The rather high R^2 values for the two semantic variables demonstrate, however, that the individual base verb plays a very important role in the interpretation of either nominalization form.

4.3 *Multivariate analyses*

Given that a number of semantic variables may be associated with morphological form, a multivariate analysis is also called for to investigate the effect of a given variable in the presence of other variables. Given that there are claims in the literature concerning words with particular constellations of semantic properties and morphological form, a method is needed that can find effects for complex interactions of variables. Mixed effects regression cannot handle complex three-way or four-way interactions in an easily interpretable way. A different method is therefore called for. Conditional inference trees are well suited and standardly used for this purpose and are employed in this paper.

For the multivariate analyses we used a subset of the data in which we used only observations for which eventivity and all aspectual features were coded unambiguously and for which quantification was coded as non-default. In other words, we are working with the intersecting set of the three data sets used in [Sections 4.2.1–4.2.3](#), and described in [Table 2](#). This intersecting set has 1421 observations (with 150 types based on 84 verbs).

We used conditional inference trees as implemented in the *partykit* package of R (Hothorn & Zeileis 2009). We devised three trees. In the first, we predicted morphological form on the basis of quantification, eventivity and the three aspectual features. In the second we predicted the quantification reading on the basis of the other variables, and in the third analysis we predicted the eventive vs. referential reading on the basis of the other variables.

The trees given below are to be read as follows. Each node contains the name of the variable according to which the data show a significant split. This means that the subset with a particular constellation of variable values behaves differently from

CONVERSION NOUNS AND *-ING* NOMINALIZATIONS

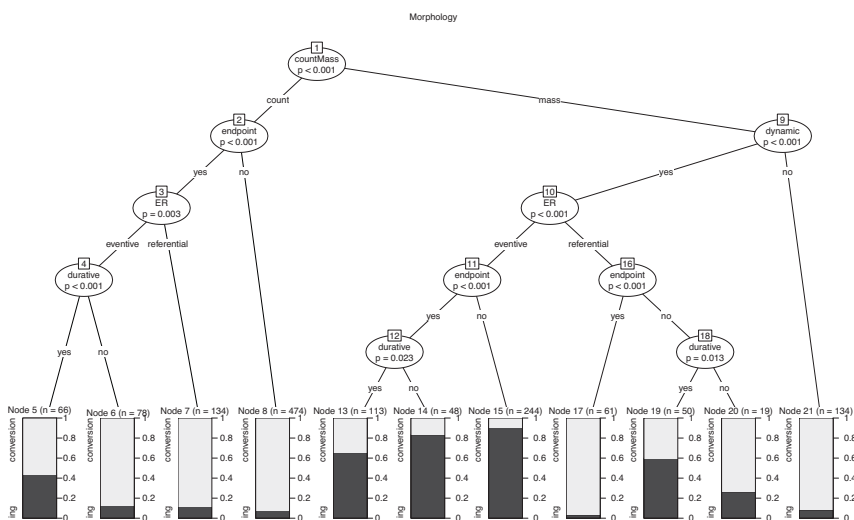


Figure 11
Conditional inference tree for the choice of morphological form.

another subset with a different constellation of values. Note that not all of the splits might be theoretically interesting because they might only show smaller differences between tendencies that go in the same direction. We will concentrate on those splits that show significant differences in their majority choice. The nodes are numbered for easier reference. The terminal nodes give the distribution of the dependent variable for the respective constellation of features in terms of a bar chart for each subset and the total number of observations in this set.

For the prediction of morphology, as shown in [Figure 11](#), we see that all variables have some influence for at least some subsets of the data. The model predicts the correct morphological category for 86 percent of all tokens. At the top level, quantification is the most important predictor. In those cases where the derivative has a count interpretation, conversion is the clear majority choice (see nodes 5–8). Within the count nominals, [implied endpoint], eventivity and [durative] also play a role for certain subsets. For the nominals with mass interpretation, the picture is more complex. If they are based on state verbs ([dynamic]: no), the clear majority choice is conversion (node 21), but if they are based on dynamic verbs ([dynamic]: yes), things become complicated: eventive nominals based on dynamic verbs favor *-ing* (nodes 13–15), and the degree of this preference is modified by the presence or absence of the features [implied endpoint] and [durative]. If referential, the majority choice depends on the values of the features [implied endpoint] and [durative] (nodes 17, 19, 20). The take-home message of all this seems to be that simple generalizations involving only one or two variables are not able to capture the complexity of what is going on in the data.

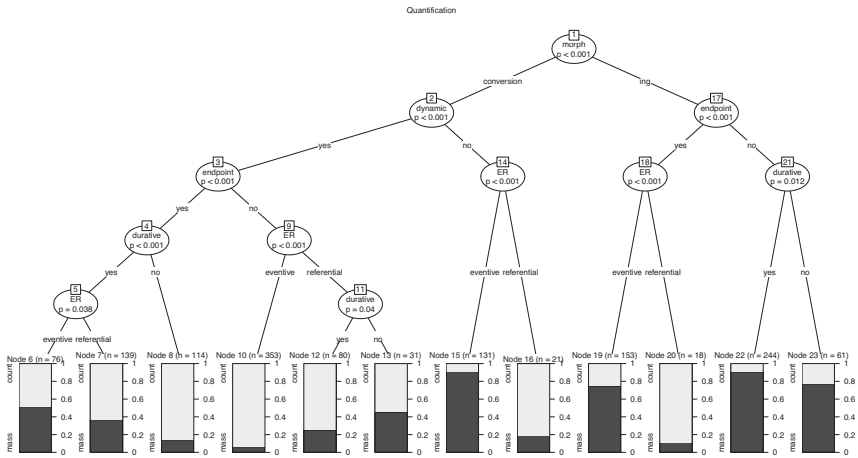


Figure 12
Conditional inference tree for the choice of count vs. mass reading.

A similarly complex picture presents Figure 12, which shows the tree for predicting quantification readings. The model predicts the correct quantification reading for 82 percent of all tokens. Again, all variables have a say, but for different subsets of variable constellations. The largest subset with the same majority choice is on the left (nodes 6–8, 10, 12, 13). Conversion nouns based on dynamic verbs show a majority of count readings. The size of the majority depends on the aspectual features of the base verb and on the eventive vs. referential reading of the nominal. Another subset with a clear majority choice is on the right. Nominals in *-ing* with a base verb that has no implied endpoint strongly tend to occur with a mass interpretation. The role of eventivity is also interesting. For nominals in *-ing* with base verbs implying an endpoint, referential interpretations go with count interpretations (node 20), and eventive interpretations go together with mass interpretations (node 19, compare also nodes 14–16). This direction in the association of quantification and eventivity is, however, not a given. Eventive conversion nouns based on dynamic verbs without implied endpoint almost categorically receive a count interpretation (node 10).

The tree predicting eventive vs. referential readings (Figure 13) also makes use of all predictor variables and is also quite successful in predicting the correct reading. The model predicts eventivity correctly for 80 percent of all tokens. On the right-hand side we see the largest uniform subset, the derivatives ending in *-ing*. Independent of their quantification reading or aspectual base verb features, they show a clear preference for eventive readings (nodes 17, 18, 20, 21). Such a preference is, however, also observable for certain subsets of conversion nouns. Count-interpreted conversion nominals based on dynamic verbs with no implied endpoint (node 10), and mass-interpreted conversion nominals based on state verbs (and no implied endpoint) are both strongly associated with eventive readings.

CONVERSION NOUNS AND *-ING* NOMINALIZATIONS

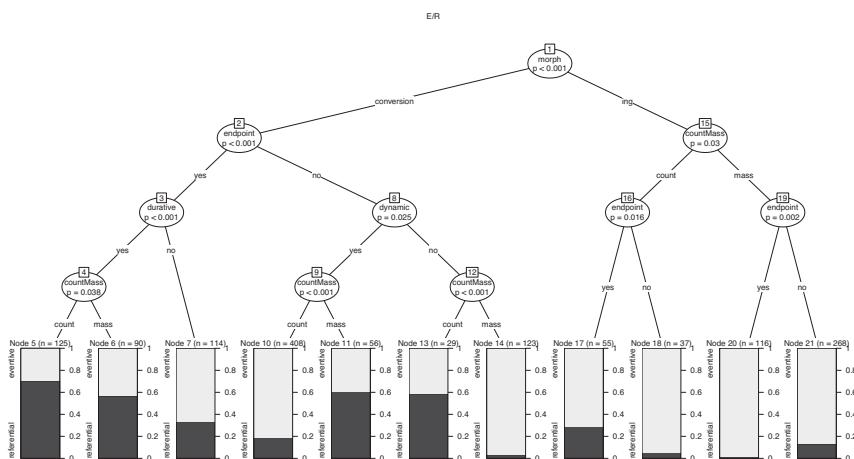


Figure 13
Conditional inference tree for the choice of referential vs. eventive reading.

The trees also provide another intriguing insight. Depending on which constellation of variables, i.e. subsets of the data, we are looking at, the effect of eventivity on quantification sometimes goes in one direction, sometimes in the other direction. The same holds for the effect of quantification on eventivity.

In sum, the classification tree analysis demonstrates that there is a complex interaction of the different semantic and morphological predictors in all models. Sweeping generalizations of certain effects of particular variables are wrong.

4.4 Summary of findings

Before we go on to consider the theoretical ramifications of these results, we briefly summarize what our analysis has revealed and compare our findings to the varied claims in the literature set out in Section 2. We look at both the semasiological perspective where we seek to find how the two morphological forms, conversion and *-ing*, tend to express various semantic nuances (quantification, aspectuality, eventivity), and at the onomasiological perspective where we seek to find out how speakers tend to encode these semantic nuances in morphological form (that is, what form they choose to express what meaning).

Looking first at the relation of morphological form and quantification, we find that from the semasiological perspective, conversion tends to express count quantification (66 percent), which is largely but not entirely in line with the claims of authors like Biese (1941). Nominalizations in *-ing* somewhat more strongly tend to express mass quantification, as claimed by Grimshaw (1990), Langacker (1991), Brinton (1995, 1998), Alexiadou et al. (2010), Borer (2013), and Park & Park (2017). However, this is not a straightforwardly categorical association as 25 percent of tokens go in the other direction. From the onomasiological perspective,

count meanings tend to choose conversion forms (86 percent). Interestingly, however, mass meanings choose either form in almost equal measure.

We now look at the association of morphological form with eventivity. From the semasiological perspective, conversion expresses eventive readings 56 percent of the time, contrary to the claims of such theorists as Roeper (1993), Alexiadou (2011), Fabregás & Marín (2012), and Borer (2013). However, in line with the claims in that literature, *-ing* forms tend to express eventive meanings (85 percent). From the onomasiological perspective eventive meanings may choose either morphological form, while referential meanings strongly tend to choose conversion (85 percent).

Looking at the relationship between morphological form and aspectual class, we find from the semasiological point of view that conversion nouns have no preference for any particular aspectual class. Nominalizations in *-ing* also have no special preference for aspectual class with the exception of state verbs, which they tend not to take as bases. From the onomasiological point of view, state verbs tend to be nominalized via conversion, with other aspectual classes showing no preference for morphological form. This finding then falsifies claims like that of Brinton (1995) that *-ing* nominalizations have a preference for activity verbs, or that of Alexiadou (2001) that *-ing* nominalizations disprefer accomplishment and achievement base verbs.

We look finally at the relationship between quantification and aspectual class. Semasiologically, count quantification occurs with base verbs of all aspectual classes, as does mass quantification. Onomasiologically, state bases have a preference for mass quantification, as suggested by Mourelatos (1978). However, Mourelatos (1978) is incorrect that activity verbs are associated with mass quantification. On the other hand, Bauer et al. (2013) are correct that punctual verbs (that is, achievements and semelfactives) have at least a tendency towards count quantification.

Turning to the conditional inference models, our analyses reveal small subpatterns in the interactions of these various factors. Predicting morphological form from quantification, eventivity and aspectual features, we find that mass quantification and the feature [–dynamic] together strongly predict conversion nominalization, as does the constellation of mass quantification, referential reading and the feature [+implied endpoint]. Count quantification (as indicated in the univariate analysis) strongly predicts conversion, and the combination of mass quantification and eventivity strongly predicts the *-ing* nominalization.

Predicting quantification from morphological form, eventivity and aspectual features, we find that *-ing* nominalizations from [–implied endpoint] verbs (activities, semelfactives) tend strongly towards mass quantification, as do *-ing* nominalizations that are [+implied endpoint] and eventive. On the other hand *-ing* nominalizations that are [+implied endpoint] and referential tend towards count quantification. Conversion forms from [–dynamic] bases (states) tend towards mass quantification if they are eventive, and tend towards count quantification if they are referential. Conversion forms from [+dynamic] verbs tend towards count quantification.

Finally, predicting eventivity from morphological form, quantification, and aspectual features, we find that *-ing* nominalizations with mass quantification tend towards eventive readings. Mass quantified conversion forms from [–dynamic] verbs tend towards eventive readings, whereas [+implied endpoint, +durative] verbs tend to be referential.

The upshot here is that the relationship between morphological form, quantification, eventivity, and the aspectual class of base verb is neither categorical in nature nor free.

5. THEORETICAL RAMIFICATIONS

The question remains of what sort of theoretical model can best accommodate a pattern of behavior that is neither categorical nor free, but rather seems to be dependent on several variables, including the individual verb involved. Before we look at specific theoretical frameworks, we offer a general observation on the sort of patterning we find.

There are a number of ways of approaching data that appear to be probabilistic. One approach might be to dismiss probabilistic patterns as random idiosyncrasy. No theoretical model expects linguistic rules to be exceptionless. Cases that appear to be exceptions to an otherwise robust pattern might simply be noise in the system that can safely be ignored. Another approach might be to attribute probabilistic behavior to the effects of historical change. Exceptions to the predominant pattern that our model predicts might be historical holdovers or emerging patterns of some sort that could also be safely ignored in a synchronic analysis.

We are reluctant to take either of these approaches, however. For one thing, it seems clear that we are not dealing here with the occasional exception to an otherwise robust pattern in which we find conversion nominalizations associated with count quantification and referential reading and *-ing* nominalizations associated with mass nominalization and eventive reading.

We have shown the interaction of morphological form, eventivity, quantification, and individual verbs to be far more complex and not easily dismissed as exceptional. We therefore choose a third approach here, which is to see to what extent our theoretical models can make sense of this probabilistic behavior.

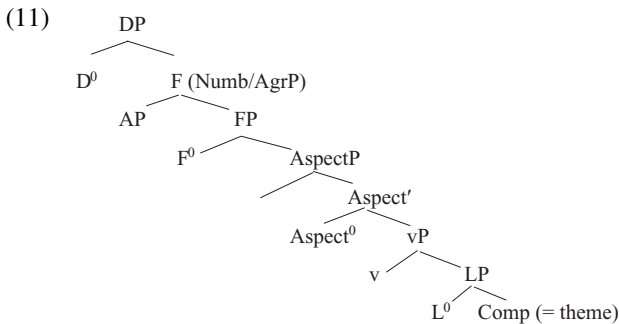
We consider here four possible models, beginning with a syntactic approach, the framework of Distributed Morphology (DM), then looking at Lieber's (2004, 2016) Lexical Semantic Framework (LSF), and finally considering two possible probabilistic approaches, Analogical Modeling (AM) (e.g. Skousen & Stanford 2007) and Distributional Semantics (DSM) (e.g. Marelli & Baroni 2015, Boleda 2020). To be clear from the outset, we do not claim that DM and LSF cannot handle the behavior we have uncovered here. Rather, we claim in what follows that they do not specifically predict the pattern that we find and therefore cannot give us much insight into it. Further, we acknowledge in advance that our presentation of AM and DSM will be far more brief and sketchy than that of DM and LSF. DM and LSF have both offered explicit analyses of nominalizations, whereas AM and

DSM have not. Our treatment of those frameworks must therefore remain rather speculative.¹⁸

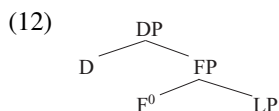
5.1 Syntactic approaches

Perhaps the most widely assumed approaches to the treatment of nominalization in current linguistic theory are syntactic approaches, among them Distributed Morphology (Alexiadou 2001, Harley 2009), Borer’s (2013) Exoskeletal model, and Nanosyntax (Baunaz & Lander 2018). We choose to use Distributed Morphology as representative here, as it has a long history of treating nominalization, but we believe that our assessment would extend to the other syntactic frameworks as well.

Adherents of DM treat nominalizers as functional categories, items that have purely grammatical meaning with fixed representations. One of the earliest treatments of deverbal nominalization within DM is that of Alexiadou (2001), which illustrates the basic components of a syntactic approach. For Alexiadou (2001: 19), the difference between an eventive and a referential reading for a nominalization can be attributed to a difference in syntactic structure. Specifically, the difference is assumed to follow from the location at which the nominalizing affix is attached (that is, higher or lower in the syntactic tree) and from the array of functional projections that occur above or below the nominalizing affix. The structures proposed by Alexiadou (2001: 19) are shown in (11) and (12):



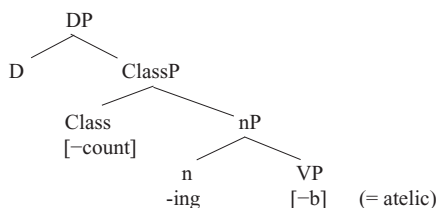
[18] One theory that does recognize the potential for multiple construals among verbs and their nominalizations is Cognitive Grammar (Langacker 1991, Croft 2012). We do not consider Cognitive Grammar here for the following reason: because Cognitive Grammar is not a formal theory in the way that DM and LSF are, it is not clear to us how it could be adapted to predict probabilistic behavior.



The structure in (11) is the one proposed for nominalizations with eventive interpretations, the one in (12) the structure for referential interpretations. The head labeled F^0 (F stands for some functional projection, perhaps NumberP or AgrP) would be the locus at which the nominalizing affix would be inserted. In (11), the nominalizing affix is attached above ‘little v’. The presence of a verbal projection below the affix plus the presence of an AspP (aspect phrase) give rise to the eventive interpretation of the nominalization. The tree in (12), lacking a vP projection has the affix under F^0 closer to the LP (for Lexical Phrase), which is the locus of the root. The referential interpretation follows from the absence of verbal and aspectual projections in the tree. The pattern in (11) is a ‘high attachment’ pattern, the one in (12) a ‘low attachment’ one. A variation on this kind of analysis can be found in Harley (2009).

We turn now to the modeling of quantification. Here, the work of Alexiadou et al. (2010) is relevant. Alexiadou et al. consider *-ing* nominalizations, arguing that although they are generally mass quantified, they may be count quantified if the base of the *-ing* form is a telic verb (which we take to mean [+ implied endpoint] in terms of features).¹⁹ They assume (following Grimshaw 1990) that all *-ing* nominalizations are eventive in reading, so they do not provide structures for forms like *heating* or *drawing* that are clearly referential. Presumably, those would have structures like that in (12) which lack a verbal projection, with one structure for count-quantified referential *-ing* forms and a different one for mass-quantified *-ing* forms. As for eventive *-ing* nominalizations, Alexiadou et al. (2010) argue for a representation like that in (13) for *-ing* forms on atelic bases that have a mass reading, using the [–b] (for unbounded) feature of Jackendoff (1991) to indicate the atelicity of the verbal base.

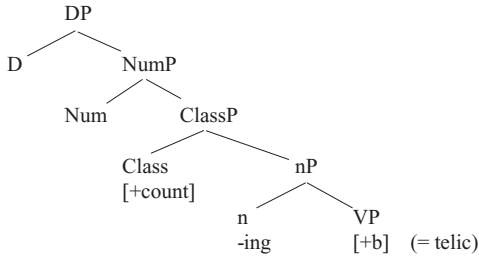
(13) Mass quantified *-ing* nominalizations



[19] As the conditional inference tree in Figure 12 shows, this is true only for cases where the *-ing* form expresses a referential reading, not for eventive cases.

In this structure, the nominalizing affix is located under ‘little n’, which in turn takes the atelic verb base as its complement. ClassP, generated above nP, is the locus of gender and case features in languages that draw those distinctions, but serves as the locus of the [+/-count] feature in English. If the feature is [-count], as they argue is the case on atelic verbal bases, the only projection above ClassP is DP. For *-ing* nominalizations on telic verbal bases, they suggest a slightly different structure, as shown in (14):

(14) *Count quantified -ing nominalizations*



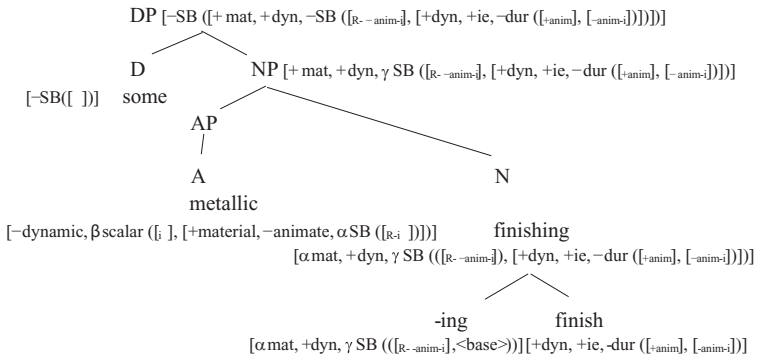
We find the *-ing* nominalization again under little n, which here takes as its complement a VP with a [+b] or telic base (bounded in Jackendoff’s 1991 terminology). The [+count] Class node then licenses the projection of the NumP (Number Phrase), which would be the locus of a plural morpheme if the nominalization is pluralized.

The critical observation for our purposes is that the DM approach to polysemy in nominalizations requires a distinct syntactic structure for each possible reading or combination of readings of a nominalization. For this reason, the DM analysis is most appealing if we find a rather restricted range of readings that can be modeled by a correspondingly small set of structures. In other words, DM works best when we find nominalizations behaving in a categorical manner, displaying a narrow range of readings associated with a clearly restricted range of classes of base verbs for each morphological form. The broader the range of readings we find with a given nominalization and the broader the range of possible constellations of aspectual class, eventivity and quantification, the larger the number of syntactic representations that would be necessary. Significantly, the DM approach would neither lead us to expect some representations to be more probable than others, nor would it predict that the individual verbal base should be an important factor in determining the range of readings we find with individual nominalizations.

5.2 *Lexical Semantic Framework*

Another possibility for modeling the behavior of these nominalizations is that offered by Lieber’s (2004, 2016, etc.) Lexical Semantic Framework (LSF). LSF

(17) *some metallic finishing*



The semantic representation is composed from the bottom up. The bottom-most branching layer shows *finish* composing with the nominalizing affix. The *-ing* nominalization is then modified by the adjective *metallic*. *Metallic*, in turn, is composed of a base *metal*, which is lexically concrete ([+material]) plus the adjective-forming affix *-ic*.²² In the context of an adjective with a [+material] base, the NP *metallic finishing* is naturally construed as concrete as well, so the feature [α material] in the nominalization is resolved as [+material]. Once the noun *finishing* is fixed as concrete, it must then be construed as a referential noun – eventives are abstract – and therefore its R argument comes into play.²³ Finally, in the context of a mass determiner like *some*, the feature [γ SB] is resolved as [-SB]. The context here allows us to resolve underspecification and determine that *finishing* must be construed as concrete rather than abstract, referential rather than eventive, mass rather than count.

The LSF analysis predicts that nominalizations like *finishing* will be maximally polysemous, that is, as likely to be read as concrete or abstract, referential or eventive, count or mass. This framework allows us to derive any of the nuances of multiply polysemous nominalizations as easily as any other. Further, it leads us to expect that any combination of semantic attributes with morphological form is as likely as any other. Given our findings in Section 4, then, LSF, like DM, makes the wrong predictions. Whereas DM predicted a pattern of readings that was categorical, LSF predicts a pattern of readings that is completely free. What we find, however, is neither categorical nor free behavior, but a far more nuanced pattern of readings. Further, like DM, LSF has little to say about the role of the individual verbal base in the range of readings available to particular nominalizations.

[22] The skeleton for an adjective-forming suffix is indicated here with the features [-dynamic, β scalar], which indicate a stative predicate that is underspecified for gradeability.

[23] That is, since the R argument is optional, its relevance is triggered by the presence of a coreferential concrete noun.

5.3 *Analogical modeling*

Analogical modeling is usually conceptualized as an exemplar-based approach in which storage of individual occurrences of expressions plays a prominent role. Analogical models, more specifically, are computational algorithms that work on the basis of a lexicon in which forms are stored together with their properties, including the property in question ('outcome'), for instance the kind of morphological form they exhibit. Based on the similarity of a given form with the forms stored in the lexicon, the given form is assigned a probability of a particular outcome (for instance, 'conversion' as the morphological form). This approach differs radically from the previous ones in that it relies heavily on stored information, and higher levels of organization (encapsulated by rules or the like) are emergent and dynamic properties of the system, which is constantly updated with new information coming in.

Two analogical models that are often used in this way are TiMBL (Daelemans & van den Bosch 2005) and AM (Skousen & Stanford 2007). In word-formation research, analogical models have been used to model stress assignment in English compounds (Plag, Kunter & Lappe 2007, Arndt-Lappe 2011), the choice of linking morphemes in Dutch compounds (Krott, Baayen & Schreuder 2001, Krott, Schreuder & Baayen 2002), diminutives (Daelemans, Berck & Gillis 1997 on Dutch, Eddington 2002, 2004 on Spanish), negative prefixes (Chapman & Skousen 2005), and nominalizations (Eddington 2006 on Spanish, Arndt-Lappe 2014 on English). Arndt-Lappe (2011), for instance, used phonological features to predict the choice between the two nominal suffixes, but one could also use other kinds of features, e.g. semantic ones. Analogies (as based on similarities between words in the lexicon) come in different degrees of locality and generality. In this way, the models are able to make correct predictions for what looks like categorical behavior, as well for sub-regularities and word-specific idiosyncrasies.

Analogical models are thus well suited to treat the kind of behavior that we find for *-ing* and conversion nominals. They are in theory able to handle patterns that are probabilistic, rather than categorical (as DM is) or free (as LSF is), and – most importantly – they are suited to modeling the contribution of the individual verbal base to the patterns we find. To date no one has tried to use an analogical model to account for the complex sort of polysemy exhibited by nominalizations. However, future work may show how successful this approach is in coping with the variability in the conversion and *-ing* data.

5.4 *Distributional semantics*

Distributional Semantic models (DSM) are based on the premise that words with similar meanings will be found in similar contexts (Marelli & Baroni 2015, Pross et al. 2017, Boleda 2020). Instead of trying to characterize meanings directly, as is done in frameworks like LSF or Model Theoretic Semantics, DSM uses similarity of context as an approximation of meaning. In other words, it characterizes meaning indirectly, rather than directly. Starting with a corpus, the meanings of lexical items

can be modeled in terms of vectors that record and count the contexts in which the item occurs in the corpus. Very roughly, the closer in semantic space we find two vectors, the closer in meaning we assume the lexical items represented by those vectors to be.²⁴ By looking at the closest neighboring vectors of a particular item, we can infer something about the meaning of that item, although such models typically remain agnostic as to what that ‘something’ is. Boleda gives the example of the vector for the word *guy*, whose closest neighboring vectors are those for *bloke*, *chap*, *doofus*, *dude*, and *fella* (Boleda 2020: 4); we might infer from these vectors that *guy* shares characteristics like ‘male’ and ‘informal’ with its nearest neighbors, but semantic features of this sort play no direct role in DSM.²⁵ What is significant for the issue that we have raised here is that the distance between vectors is gradient.

Marelli & Baroni (2015) show that DSM can be applied in analyzing the semantic representation and processing not just of simplex words, but also of complex derived words. They attempt to model the meanings of affixes as functions that take the vectors of bases as input and produce other vectors as outputs (p. 489). These functions ‘capture systematic patterns linking two separate context distributions’ (p. 491), one of these being the base and the other the derived form: ‘affixes are to be considered as high-order associations between the distributional semantics of different words’ (pp. 291–292).

DSM offers an interesting take on polysemy. Presumably if a lexical item displays multiple senses, its vector should appear with different clusters of nearest neighboring vectors which help to identify and distinguish those senses. Marelli and Baroni illustrate this with the nominalizing affix *-ment* in English, which – like *-ing* and conversion – displays both eventive and referential readings (termed ‘process’ and ‘result’ by Marelli & Baroni 2015: 492). What they show is that the nearest neighbors of an eventive nominalization in *-ment*, such as *interment* are verbs and other eventive-flavored nominalizations, whereas for a referential nominalization such as *equipment* the nearest neighbors are words which themselves suggest referentiality (see also Lapesa et al. 2018 for a similar approach). Marelli & Baroni thus suggest that DSMs can make predictions with respect to the sorts of polysemy we ought to find with derived words and – importantly for our purposes – the extent to which that polysemy might be found for particular types of bases and particular derived forms.

An interesting study exploring these issues for competing nominalization processes in Italian is Varvara (2017). In this language, event nominalizations can be created by using nominalized infinitives or suffixes such as *-ita* or *-ione*. These two kinds of processes are not complementarily distributed, and the interpretations are also variable across and within categories. Using large amounts of data and DSMs Varvara demonstrates that Italian nominalized infinitives tend to be based on

[24] Closeness is represented as cosine similarity; in other words, ‘the closer the vectors, the larger the cosine similarity’ (Boleda 2020: 3).

[25] Boleda (2020) cites Baroni (2016) as the original source of the example.

a verb's metaphoric and abstract sense, while event-denoting nominalizations derived by suffixes tend to be based on a base's literal or concrete sense. Another important result is that nominalized infinitives are more predictable in their meaning, and that suffixed nominalizations are less semantically transparent and more prone to semantic drift. The competition of nominalizations in Italian is very similar to the one described for English in the present study, and Varvara's study shows that the tools of distributional semantics can do justice to the complexity of this situation.

At this point, there is no study available that tests a distributional semantic approach against the sort of patterns we have described in Section 4. However, because the framework is consistent with semantic patterns that are neither categorical nor free, it is potentially well suited to modeling the intricate polysemy and quantitative profile of conversion and *-ing* forms, as well as the contributions of individual verbal bases to this pattern.

6. CONCLUSIONS

We set out to determine the extent to which an association could be found between conversion, referential reading, and count quantification, and between *-ing* nominalization, eventive reading, and mass quantification. Furthermore, we wanted to find out whether there were any associations between the morphological form of a nominalization and the aspectual class of base verbs from which they are derived. What our statistical study has revealed is that there is no simple correlation. Sometimes the associations between morphological form, quantification, eventivity, and aspectual class of the base verb go in the predicted direction, but never in a categorical fashion. Some associations are stronger than others and some associations go in the opposite direction of that claimed in the literature. The pattern we find is complex and often it seems to be determined by the base verb of a particular nominalization.

While statistical murkiness might seem to some to be an unsatisfying result, it is nevertheless a theoretically significant one, as it leads us to question the majority of current analyses of nominalizations. Neither syntactic frameworks that predict categorical semantic behavior of particular kinds of nominalizations nor lexical semantic frameworks that predict free semantic behavior seem optimal, given our findings. But other frameworks such as analogical and distributed semantic models might be more suited to such findings. Although we cannot offer analyses within this framework here, our results suggest these as new avenues of analysis for nominalization that we hope will be pursued in the future.

REFERENCES

- Alexiadou, Artemis. 2001. *Functional structure in nominals: Nominalizations and ergativity*. Amsterdam: John Benjamins.
- Alexiadou, Artemis. 2010. Nominalizations: A probe into the architecture of grammar. *Language and Linguistics Compass* 4, 496–523.

- Alexiadou, Artemis. 2011. Statives and nominalization. *Recherches linguistiques de Vincennes* 40, 25–52.
- Alexiadou, Artemis, Gianina Iordachioaia & Elene Soare. 2010. Number/aspect interactions in the syntax of nominalizations: A distributed approach. *Journal of Linguistics* 46.3, 537–574.
- Andreou, Marios & Rochelle Lieber. 2020. Aspectual and quantificational properties of deverbal conversion and *ing* nominalizations: The power of context. *English Language and Linguistics* 24.2, 333–363.
- Arndt-Lappe, Sabine. 2011. Towards an exemplar-based model of stress in English noun–noun compounds. *Journal of Linguistics* 47.11, 549–585.
- Arndt-Lappe, Sabine. 2014. Analogy in suffix rivalry: The case of English *-ity* and *-ness*. *English Language and Linguistics* 18.3, 497–548.
- Baroni, Marco. 2016. Composes: An executive summary. Presented at the Composes Workshop.
- Barton, Kamil. 2009. MuMIn: Multi-model inference. Software package. <http://r-forge.r-project.org/projects/mumin>.
- Bates, Douglas M., Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models using *lme4*. *Journal of Statistical Software* 67.1, 1–48.
- Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
- Baunaz, Lena & Eric Lander. 2018. Nanosyntax: The basics. In Lena Baunaz, Karen DeKlerck, Liliane Haegeman & Eric Lander. *Exploring Nanosyntax*, 3–56. Oxford: Oxford University Press.
- Biese, Yrjoe. 1941. Origin and development of conversions in English. *Annales Academiae Scientiarum Fennicae B XLV-2*, 1–495.
- Boleda, Gemma. 2020. Distributional semantics and linguistic theory. *Annual Review of Linguistics* 6, 213–234. <https://doi.org/10.1146/annurev-linguistics-011619-030303>.
- Borer, Hagit. 2005. *Structuring sense*, vol. I: *In name only*. Oxford: Oxford University Press.
- Borer, Hagit. 2013. *Structuring sense*, vol. III: *Taking form*. Oxford: Oxford University Press.
- Brinton, Laurel. 1995. The aktionsart of deverbal nouns in English. In Pier Marco Bertinetto, Valentina Bianchi, James Higginbotham & Mario Squartini (eds.), *Temporal reference, aspect, and actuality*, vol. I: *Semantic and syntactic perspectives*, 27–42. Turin: Rosenberg and Sellier.
- Brinton, Laurel. 1998. Aspectuality and countability: A cross-categorical analogy. *English Language and Linguistics* 2.1, 37–63.
- Cetnarowska, Bożena. 1993. *The syntax, semantics and derivation of bare nominalisations in English*. Katowice: Uniwersytet Śląski.
- Chapman, Don & Royal Skousen. 2005. Analogical modeling and morphological change: The case of the adjectival negative prefix in English. *English Language and Linguistics* 9.2, 333–357.
- Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Croft, William. 2012. *Verbs: Aspect and causal structure*. Oxford: Oxford University Press.
- Daelemans, Walter, Peter Berck & Steven Gillis. 1997. Data mining as a method for linguistic analysis: Dutch diminutives. *Folia Linguistica XXXI*, 57–75.
- Daelemans, Walter & Antal van den Bosch. 2005. *Memory-based language processing*. Cambridge: Cambridge University Press.
- Eddington, David. 2002. Spanish diminutive formation without rules or constraints. *Linguistics* 40, 395–419.
- Eddington, David. 2004. *Spanish phonology and morphology: Experimental and quantitative perspectives*. Amsterdam & Philadelphia, PA: John Benjamins.
- Eddington, David. 2006. Look Ma, no rules: Applying Skousen’s analogical approach to Spanish nominals in *-ión*. In Grace Wiebe, Gary Libben, Tom Priestly, Ron Smyth & H. S. Wang (eds.), *Phonology, morphology, and the empirical imperative: Papers in Honour of Bruce L. Derwing*, 371–407. Taipei: Crane.
- Fábregas, Antonio & Rafael Marín. 2012. The role of Aktionsart in deverbal nouns: State nominalizations across languages. *Journal of Linguistics* 48.1, 35–70.
- Filip, Hana. 2012. Lexical aspect. In Robert I. Binnick (ed.), *The Oxford handbook of tense and aspect*, 721–51. New York: Oxford University Press.
- Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.
- Grimshaw, Jane. 2011. Deverbal nominalization. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 1, 1292–1313. Berlin & Boston, MA: DeGruyter.

- Harley, Heidi. 2009. The morphology of nominalizations and the syntax of vP. In Anastasia Giannakidou & Monika Rathert. (eds). *Quantification, definiteness, and nominalization*, 321–343. Oxford: Oxford University Press.
- Hothorn, Torsten & Achim Zeileis. 2009. partykit: A toolkit for recursive partytioning, <http://R-Forge.R-project.org/projects/partykit/>.
- Iordachioaia, Gianina & Martina Werner. 2019. Categorical shift via aspect and gender change in deverbal nouns. *Language Sciences* 73, 62–76.
- Jackendoff, Ray. 1991. Parts and boundaries. *Cognition* 41, 9–45.
- Krott, Andrea, R. Harald Baayen & Robert Schreuder. 2001. Analogy in morphology: Modeling the choice of linking morphemes in Dutch. *Linguistics* 39.1, 51–93.
- Krott, Andrea, Robert Schreuder & R. Harald Baayen. 2002. Linking elements in Dutch noun–noun compounds: Constituent families as predictors for response latencies. *Brain and Language* 81, 708–722.
- Langacker, Ronald. 1991. *Foundations of cognitive grammar*, vol. 2: *Descriptive applications*. Stanford, CA: Stanford University Press.
- Lapesa, Gabriella, Lea Kawalecz, Ingo Plag, Marios Andreou, Max Kisselew & Sebastian Pado. 2018. Disambiguation of newly derived nominalizations in context: A Distributional Semantics approach. *Word Structure* 11.3, 277–312.
- Lieber, Rochelle. 2004. *Morphology and Lexical Semantics*. Cambridge: Cambridge University Press.
- Lieber, Rochelle. 2016. *English nouns: The ecology of nominalization*. Cambridge: Cambridge University Press.
- Marelli, Marco & Marco Baroni. 2015. Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review* 122.3, 485–515.
- Mourelatos, Alexander. 1978. Events, processes and states. *Linguistics and Philosophy* 2, 415–434.
- Newmeyer, Frederick J. 2009. Current challenges to the lexicalist hypothesis: An overview and a critique. In William D. Lewis, Simin Karimi, Heidi Harley & Scott O. Farrar (eds.), *Time and again: Theoretical perspectives on formal linguistics. In honor of D. Terence Langendoen*, 91–117. Amsterdam: John Benjamins.
- Park, Chongwon & Bridget Park. 2017. Cognitive grammar and English nominalization: Event/result nominals and gerundives. *Cognitive Linguistics* 28.4, 711–756.
- Plag, Ingo, Gero Kunter & Sabine Lappe. 2007. Testing hypotheses about compound stress assignment in English: A corpus-based investigation. *Corpus Linguistics and Linguistic Theory* 3.2, 199–233.
- Pross, Tillman, Antje Rosseidscher, Sebastian Pado, Gabriella Lapesa & Max Kisselew. 2017. Integrating lexical-conceptual and distributional semantics: A case report. *Proceedings of the 21st Amsterdam Colloquium*, 75–84.
- R Core Team. 2020. R: A language and environment for statistical computing. Vienna, <https://www.R-project.org/>.
- Roeper, Thomas. 1993. Explicit syntax in the lexicon: The representation of nominalizations. In James Pustejovsky (ed.), *Semantics and the lexicon*, 185–220. Dordrecht: Kluwer.
- Roy, Isabelle & Elena Soare 2011. Nominalizations: New insights and theoretical implications. *Recherches Linguistiques de Vincennes* 40, 7–23.
- Skousen, Royal & Theron Stanford. 2007. AM: Parallel. available from <http://humanities.byu.edu/am/>.
- Smith, Carlota. 1991. *The parameter of aspect*. Dordrecht: Kluwer.
- Sutton, Peter R. & Hana Filip. 2019. Singular/plural contrasts: The case of informational object nouns. *Proceedings of the 22nd Amsterdam Colloquium*, 367–376.
- Varvara, Rosella. 2017. *Verbs as nouns: Empirical investigations on event-denoting nominalizations*. Ph.D. dissertation, Università degli Studi di Trento.
- Verkuyl, Henk. 1972. *On the compositional nature of the aspects*. Dordrecht: Reidel.

Authors' addresses: (Lieber)

*English Department, University of New Hampshire, Hamilton Smith Hall,
95 Main Street, Durham, NH 03824, USA
Rochelle.Lieber@unh.edu*

(Plag)

*Institut für Anglistik und Amerikanistik, English Language and Linguistics,
Heinrich-Heine-Universität Düsseldorf, 40204 Düsseldorf, Germany
ingo.plag@uni-duesseldorf.de*