

No ordered arguments needed for nouns^{*}

Scott Grimm and Louise McNally

Universitat Pompeu Fabra
{scott.grimm,louise.mcnally}@upf.edu

Abstract

Syntacticians have widely assumed since [11] that there is a fundamental difference between so-called argument structure nominals (AS-nominals, also called Complex Event Nominals), e.g. *destruction*, and non-AS-nominals, e.g. *book* ([1, 5], i.a.). Grimshaw provided a list of properties characterizing AS-nominals, most notably that they have obligatory arguments (e.g. *the destruction *(of Carthage) by the Romans*). She and others have associated having argument structure with having event structure, but it has never been clear what having or lacking such structures amounts to semantically. In this paper we present extensive corpus evidence that AS-nominals do not in fact exist as a distinct class. This result, we argue, removes an important challenge to [9]’s hypothesis that eventuality-denoting nouns systematically lack an ordered-argument semantics.

1 Introduction

Syntacticians have widely assumed since [11] that there is a fundamental difference between so-called argument structure nominals (hereafter, AS-nominals, also called Complex Event Nominals), e.g. nominalizations like *destruction*, and non-AS-nominals like *book* or *trip* ([1, 5], i.a.; see also [3, 12]). Grimshaw provided a list of properties characterizing AS-nominals, most notably that they have obligatory arguments (e.g. *the destruction *(of Carthage) by the Romans*). She and others have associated having argument structure with having event structure; this, in turn, has been crucial to motivating a fundamental grammatical distinction between *Complex Event Nominals*, on the one hand, and *Simple Event Nominals* and *Result Nominals*, on the other. However, it has never been clear in this literature what having or lacking AS or event structure amounts to semantically. In this paper we address precisely this issue. We begin by presenting extensive corpus evidence that AS-nominals do not in fact exist as a distinct class. This result, we argue, removes an important challenge to [9]’s hypothesis that eventuality-denoting nouns systematically lack an ordered-argument semantics.

[9], in a context outside the nominalization literature, distinguished two general sorts of analyses for any predicate describing an eventuality: (i) as an ordered n -tuple of individuals, where the participants must be fed to the predicate by function application in order to form a proposition, following classical Montague semantics (e.g. (1)); vs. (ii) as a neo-Davidsonian 1-place property of eventualities (e.g. (2-a)) with the participants added effectively as adjuncts by conjunction, mediated by thematic role predicates (e.g. (2-b); see also [6, 18]).

(1) $\lambda y \lambda x \lambda e [\text{destruction}(e, x, y)]$

(2) a. $\lambda e [\text{destruction}(e)]$ b. $\lambda e [\text{destruction}(e) \wedge \mathbf{Ag}(e, \text{romans}) \wedge \mathbf{Th}(e, \text{carthage})]$

^{*}We thank the anonymous reviewers for the Colloquium for comments. This research has been supported by Spanish Ministry of Education and Science grant FFI2010-15006 and an ICREA Foundation Academia award.

Dowty hypothesized that verbs have the former sort of semantics, while nouns have the latter, elaborating little on why this should be. However, he acknowledged (1989:90-91) that the existence of some cases of nominals with apparently obligatory PPs, AS-nominals in different terminology, posed a challenge for this hypothesis.

Though the intuitions supporting the existence of AS-nominals are initially convincing, counterexamples to the claims concerning their diagnostics are well known (see [5, 17, 14] references cited therein). We therefore decided to test the strength of these diagnostics on a broad empirical level. Our results show that (i) none of the putative properties in fact discriminate between nominalizations with PPs expressing participants and those without, and (ii) the presence of such PPs with a nominalization is instead predicatable based on the discourse context. We argue that the data can be explained with a single neo-Davidsonian semantics for nominals plus a general pragmatic principle that requires all token eventualities (and certain subeventualities) to be anchored to at least one token discourse referent on first mention, in the spirit of the Argument-Per-Subevent condition of [21].

2 An Empirical Study of the Diagnostics of AS-Nominals

[11] argued nominalizations are ambiguous between different readings, in turn influencing the presence of arguments. These different readings are exemplified in (3), reproduced from [2]. *Examination* has a “complex event interpretation” (involving participants) in (3-a), while *examination* in (3-b) is argued to refer to an event without an articulated event structure, and therefore no arguments. In (3-c), *examination* refers to a physical object, again not possessing event structure, thereby lacking arguments.

- | | | | |
|-----|----|---|------------------------|
| (3) | a. | The examination of the patients took a long time. | (AS-nominal) |
| | b. | The examination took a long time. | (Simple Event Nominal) |
| | c. | The examination was on the table. | (Result Nominal) |

[11] argued that a cluster of properties distinguished AS-nominals from non-AS-nominals, whether Simple Event Nominals or Result Nominals. AS-nominals 1) have obligatory arguments, 2) denote eventualities; 3) take agent-oriented modifiers such as *deliberate* or *intentional* which are banned for non-AS-nominals; 4) take Saxon genitives and *by*-phrases as arguments, not adjuncts; 5) allow implicit argument control; 6) take aspectual modifiers such as *in an hour* or *for an hour*; 7) allow modifiers such as *frequent* or *constant* in the singular, while non-AS-nominals require the plural; and 8) must be singular, while non-AS-nominals may be plural. We conducted an extensive corpus analysis in order to examine whether the ambiguity in nominalizations hypothesized by Grimshaw and others was observable. In particular, we examined properties 3), 6), 7), and 8), which are easily observable with corpus techniques.

Using the information on derivational morphology available in the CELEX database [4], we selected 150 nouns derived by *-(at)ion* or *-ment*, including all *-ment* or *-ation* nouns in the database as well as a random selection of frequent nouns derived by *-ion* (e.g. *destruction*). For each noun, we extracted up to 1000 random occurrences, or as many as present if less than 1000, from the COCA corpus (<http://corpus.byu.edu/coca/>).

We then processed the data in several steps, using a dependency parser and then extracting the information thereby obtained. First, we developed a Python script to standardize the search results from COCA, extracting the sentence where the search term occurred and discarding any surrounding fragments in the search result. These sentences were then passed to the Stanford Dependency Parser [7], which returned both a parse of the sentence and a list of

the dependencies. Another Python script used the information coded in the dependencies to automatically extract information relevant to the properties put forth by Grimshaw, including whether arguments were present (*of*-phrase, *by*-phrase, etc.), number (singular/plural), the presence of adjectival modifiers and any further PP modifiers. At this stage, instances which appeared as a part of a compound, as in *examination criteria*, were excluded as they clearly were not of interest for the hypotheses under investigation. Finally, an R script was developed which constructed a large data frame in which all the information for each nominalization was collected, tallied and submitted to the appropriate statistical tests where relevant.

Overall, the procedure was highly accurate for the questions we address. Although [7] estimate that the Stanford Dependency Parser has a per-dependency error rate of 80.3% correct, as we were only examining a subset of the dependencies, which were mostly very local, the number of errors that were relevant for our study were far fewer. An error analysis on 400 sentences resulted in an effective error rate of 97% correct. We also examined across 5 nouns how reliably the *of*-phrase extracted by this procedure corresponded to a true argumental *of*-phrase. This varied greatly according the meaning of the noun and potential for polysemy. Some nouns had nearly exclusively argumental *of*-phrases (*abandonment* 99%), while others had a minority (*adornment* 25%). As hand-correcting all of the occurrences is not currently feasible, we simply checked the data points which were relevant to the properties investigated to ensure reliability, e.g. the *of*-phrases for each occurrence where *deliberate* appeared as an adjectival modifier. We now discuss the results with respect to each of the properties.

Number: The overwhelming majority of deverbal nouns had nearly all of their occurrences in the singular, whether an *of*-phrase was present or not. While [11] claims that all AS-nominals only occur in the singular, 20% of AS-nominals in our data set occurred at least once in the plural. More generally, number does not distinguish between AS- and non-AS-nominals, since whether the nominal had an *of*-phrase or not had no effect on its occurrence in the plural. Nouns with high rates of singular (plural) occurrences when PPs were present also had high rates of singular (plural) occurrences when no PPs were present: e.g., 98% of the occurrences of *allegation*, both with and without an *of*-phrase, were singular, while for *observation*, 46% of the occurrences, with and without an *of*-phrase, were singular. For each nominal, we ran Fischer's Exact Test comparing the number of occurrences in singular and in plural of the AS-nominal with the number of occurrences in singular and in plural of the non-AS-nominal. For nearly all the nominalizations considered, there was no significant difference between nominals with and without an *of*-phrase. The only exceptions were *application*, *assessment*, and *examination*—all of which have substantive uses in a non-eventive result reading.

Adjectival Modifiers: [11] and subsequent authors claimed that (i) *constant* and *frequent* should occur with singular forms of AS-nominals, but with plural forms of nominalizations without an *of*-phrase and that (ii) *deliberate* and *intentional* only occur with AS-nominals as they crucially refer to an agent. Table 1 displays the distribution of these adjectival modifiers observed in the data. Neither claim is substantiated: *constant* and *frequent* appear happily with singular forms of nominalizations without an *of*-phrase and *deliberate* and *intentional* occur equally in nominalizations with and without an *of*-phrase. The behavior of these modifiers then, rather than distinguishing between AS-nominals and nominals without *of*-phrases, indicates that the nominalization behaves similarly regardless of the presence of an *of*-phrase.

Aspectual Modifiers: Aspectual modifiers of the type *for an hour/in an hour* were extremely rare. Out of 3363 data points which had *for*- or *in*-phrases, only 3 were aspectual modifiers, all of which were *for*-phrases and occurred with non-AS-nominals, contrary to the claim of [11]. An example is given in (4):

- (4) Rosenblatt and co-authors (1987) also reported that at least 40 mg/L ClOsub2 **gas**

Constant	Sing	Plur	Frequent	Sing	Plur
AS	16	0	AS	15	0
Non-AS	46	0	Non-AS	10	0
Deliberate	Sing	Plur	Intentional	Sing	Plur
AS	8	0	AS	6	0
Non-AS	21	0	Non-AS	8	0

Table 1: Distribution of Adjectival Modifiers *constant*, *frequent*, *deliberate* and *intentional*

treatment for one hour at 60 percent RH effectively sterilized Whatman 3-mm filter paper strips ... (COCA)

Obligatoriness of Arguments We then further tested some of the strongest evidence that has been used to argue for AS-nominals: (i) the *of*-phrase is obligatory when a *by*-phrase is present ([13]), and (ii) certain nominalizations, such as *sending* or *handing* appear to be banned without an *of*-phrase ([20]), e.g. *the sending* *(*of the letters*). We found that, in the right context, neither generalization holds, as the attested examples in (5) and (6) show. Interestingly, this behavior is not unique to deverbal or even deadjectival nominals (e.g. *awareness*): Even those derived from nouns, such as *friendship*, show similar behavior (see (7)).

- (5) How does a country recover from 40 years of **destruction by an unchallenged tyrant**? (Newsweek)
- (6) While the originator—or his or her computerized agent—purposefully sent the information item into Jurisdiction B, **the sending** occurred instantaneously in response to the address supplied by the client, without any advance arrangement by the content originator. (http://www.kentlaw.edu/perritt/publications/41%20VILL._L._REV._1.htm)
- (7) a. ??Bob’s friendship vs. ^{ok}Bob’s friendship with Sue
b. Sue has known Bob for years. Bob’s friendship means the world to her.

To conclude this study, the properties proposed by Grimshaw and others do not accord with the data that we examined. While the nominalization literature asserts a fundamental ambiguity between AS-nominals and eventive nominalizations without arguments (Simple Event Nominals), the examples in (6)–(7) indicate a different possibility: the optionality of PPs with a nominalization may be simply due to the overall discourse context, rather than to an inherent ambiguity between AS-nominals and non-AS-nominals.

3 The Discourse Behavior of AS-Nominals

We conducted a second study using the Brown Corpus to investigate the general relation between the presence of PPs and discourse context. The investigation centered on two related questions. First, if the nominalization is eventive, and no argument is present, is the argument recoverable? Second, what is the information status of the discourse referent specified by the *of*-phrase or that of the recovered argument? If the lack of an *of*-phrase corresponds with an interpretation where no internal argument is specifiable, this is strong evidence for an ambiguity account. If, however, the lack of an *of*-phrase corresponds with internal arguments that are recoverable from the discourse context, that supports a view upon which nominalizations simply do not need to specify the internal argument if its referent is sufficiently salient.

We used the Brown corpus for this study since the entire discourse context, e.g. an entire news article, is available to examine. We automatically searched for instances of a large subset of the *-ation* and *-ion* nominals from our first corpus study, extracting their occurrences along with the entire context. This resulted in 867 occurrences of 37 different noun types.

We ensured that we were comparing like interpretations of the nominalizations, namely eventive and full noun phrase uses of nominalizations which potentially would have an internal argument, by hand-checking each occurrence. We excluded all instances which referred to physical objects, i.e. “result nominal” interpretations (e.g. *All of the jackets carry a fairly technical and detailed explanation of this new recording program.*). We also excluded instances which were not full noun phrases, but occurred in light-verb constructions (*It would allow presentation to the public of a unified approach.*) or occurred bare with prepositions (*under occupation*). We also set aside nominals which selected for a different preposition than *of*, e.g. *invitation for*. Finally, we excluded nominals occurring as part of a compound (*a German recitation*) or with possessive pronouns which would specify the internal argument. 259 occurrences remained.

For the cases in which no *of*-phrase was present, to determine whether an internal argument was recoverable, we considered whether the occurrence was paraphrasable with an *of*-phrase. For instance, in the example in (8), *interpretation* occurs without an *of*-phrase, but it is clear that the interpretation is of something, namely ‘the information’ which is already present as the subject of the sentence, rather than a “simple event” of interpreting.

- (8) This information was accepted with **the frequent interpretation** that those persons who did not show arm-levitation must be preventing it.

In general, when no *of*-phrase was present, the internal argument was nearly always recoverable. Only in 4 instances did there seem to be a true “Simple Event” usage, and these were essentially limited to the noun *presentation*, e.g. *The presentation was made before several hundred persons at the annual meeting . . .*

We annotated the information status of the *of*-complement, if present, or if not present but recoverable, that of the recovered discourse referent. We followed the coding scheme in [8] which has three primary information structure categories: *given* if the discourse referent was previously mentioned, *accessible* if the discourse referent is generally known or inferrable in some manner from the discourse, and *new* if the discourse referent is not previously mentioned or otherwise inferrable. *Given* has two subcategories: *active*, when the discourse referent is present within the current or immediately preceding sentence, and *inactive*, if mentioned earlier than the preceding sentence. (*Accessible* also has subcategories which were coded for but are not relevant in what follows.)

If the *of*-complement was not present but the internal argument was recoverable, in 82% of the cases it was *given* and in 18% of the cases it was *accessible*. Most all occurrences classified as *given* were *given-active* (91%). Thus, when an eventive nominalization occurs without a PP, the content of the missing PP is nearly always recoverable and mostly very salient in the discourse. In contrast, for those nominalizations where the *of*-phrase was present, 47% of the referents in the *of*-phrase were classified as *new*. Of the remaining cases, 34% were *given* and 19% *accessible*. Those that were given were almost even split between *given-active* (51%) and *given-inactive* (49%). Thus, when an eventive nominalization occurs with a PP, the discourse referent is nearly half of the time new, and not frequently mentioned in the immediately preceding discourse (17% of the time). In summary, the presence or absence of a PP does not generally correspond to AS-nominal or Simple Event readings, respectively, although there may certainly be uses of certain nominalizations corresponding to a Simple Event Nominal, such as *presentation*. Instead, the driving force behind the occurrence of PPs appears to be the discourse salience of its referent.

4 Analysis

Based on the data, we see no argument for a distinct category of AS-nominals whose arguments are obligatorily expressed. Rather, the data indicate that nominalizations are simply relation-*entailing* nouns, as represented in (2b); the participant variables are free and can be contextually valued (see [22] for such a proposal). To this we add the condition on reference to a particular token eventuality in (9), which is a semantic counterpart of the Argument-Per-Subevent condition of [21], (p. 779) in (10):

- (9) **Event Instantiation Condition:** The introduction of a token discourse referent for an eventuality e requires that e be anchored to a discourse referent corresponding to at least one of its participants.
- (10) **The Argument-Per-Subevent Condition:** There must be at least one argument XP in the syntax per subevent in the event structure.

The crucial restriction to the *introduction* of a *token* eventuality is not a simple stipulation but rather is grounded in the deeper hypothesis that token eventualities can only be properly identified and apprehended if we identify at least one of their participants, whereas our identification and apprehension of *types* of eventualities involves other cognitive mechanisms for which the identification of participants is not essential (see [15] for related discussion).

We see various advantages to an account based on (9). First, it predicts the possibility of variability in the presence vs. absence of participant PPs. If we appealed to the Argument-Per-Subevent condition, we would predict no such variation because the subevent structure associated with a word should be the same across all of its uses. Second, (9) squares with the observation that AS-nominals tend to denote (unique) token eventualities, while non-AS-nominals are not so restricted. The Argument-Per-Subevent condition also fails to make this prediction because there is no reason to think that the event structure for kind- vs. token-level eventualities should differ.

Third, we predict a clear correlation between the anaphoric use of nominals and the omissibility of arguments. If a discourse referent for an eventuality is already introduced, anaphoric reference to it will not require re-anchoring it via some participant. In contrast, if the nominal is being used to introduce a token eventuality referent for the first time, as we would argue is the way in which the classic examples from the literature are most naturally interpreted, the participant anchor for the eventuality must be specified, either via a PP or in some cases via a Saxon genitive.

We can exemplify the process of specification via PPs with *of*-phrases. Following [19] i.a., we take (genitive) *of*-phrases to introduce an underspecified relation, e.g. $\lambda x[R(x, \text{carthage})]$, x of the entity or event sort. The value for this relation will be picked up from the relational noun. As is well known, *of*-phrases with nominalizations are ambiguous, for example between agent and theme interpretations. Conjoining a representation for nominalizations of the sort given in (2) with that for the *of*-phrase results in specifying either the agent or theme, depending on which role relation R picks up on. More “contentful” PPs will work analogously, but contribute additional lexical entailments. For example, *by*-phrases specify that the participant is the initiator of the event.

Fourth, an analysis grounded in (9) immediately carries over to deadjectival and denominal forms, since nothing in it hinges on verb-specific properties. Analyses grounded in event structure do not obviously predict this similarity. Fifth, it allows us to eliminate the otherwise unmotivated simple vs. complex event nominal distinction from the grammar. Finally, though space limitations preclude extensive discussion here, it avoids the difficulties faced by alter-

native accounts on which the noun's arguments are claimed to be present but simply elided or expressed by null pronominals or similar elements. The non-expression of the material in question simply does not appear to be sensitive to the structural restrictions that are typical of ellipsis in either nominals or VPs (the latter observation, by the way, further calls into question the use of verbal structure in the analysis of these nominals). An analysis in terms of null pronominal arguments is complicated by the need to deal with material corresponding to the missing preposition(s) and by the difficulties of integrating such an analysis into a theory of null pronominals in English more generally. On our account, there is no need to fit the facts into either the theory of ellipsis or null pronominals.

Rescuing Dowty's claim that nouns lack an ordered-argument semantics leaves us with the question of why nouns and verbs would differ in this regard, as he suggested. In fact, it would not surprise us if, contrary to what Dowty argues, [6] was right in defending a neo-Davidsonian analysis for verbs as well. Dowty based his claim concerning the differences between verbs and nouns on the fact that seemingly truth-conditionally equivalent verbs (e.g. those in (11)) differ in the obligatoriness of their arguments.

- (11) a. John ate (the meal). b. John devoured *(the meal).

However, these asymmetries might be explainable without appealing to an ordered-argument semantics.¹ The first crucial observation is that nonpronominal verb phrases, unlike noun phrases, are arguably never anaphoric, and they are almost always used for token reference. Neither characteristic applies to nominals. The participant realization properties of nouns and verbs will thus be systematically different, but we need not conclude that this is accompanied by the sort of distinction enshrined in the AS-nominal literature.

As for the contrast in (11), there are at least two possible explanations for it that do not depend on ordered arguments. One could appeal to the well-known aspectual and distributional differences between *eat* and *devour* to argue that *eat* is an activity predicate with no interesting subevent structure, while *devour* is necessarily a change of state predicate with activity and result state subevents (see e.g. the references cited in [16]). Alternatively, we could follow [10], who shows that *eat* and *devour* differ in that the latter occurs with a much wider variety of arguments; their presence may therefore be necessary to identify exactly what sort of eventuality is being described. Either way, the strong tendency for the object argument in (11-b) but not (11-a) could follow directly.

It remains to replicate this study with *-ing* nominals to see whether the observations reported here extend to them as well. For the time being, however, we tentatively conclude that no ordered arguments are needed for nouns.

References

- [1] Artemis Alexiadou. *Functional structure in nominals*. John Benjamins, Amsterdam, 2001.
- [2] Artemis Alexiadou and Jane Grimshaw. Verbs, nouns and affixation. In F. Schäfer, editor, *SinSpec (1): Working Papers of the SFB 732, University of Stuttgart*, pages 1–16, 2008.
- [3] Mona Anderson. Prenominal genitive NPs. *The Linguistic Review*, 3:1–24, 1983.
- [4] R. Harald Baayen, Richard Piepenbrock, and H. van Rijn. The CELEX lexical data base on CD-ROM. 1993.

¹We are optimistic that the account we sketch here for verbs can also address a concern raised by an anonymous reviewer as to whether (10) by itself is enough to predict all of the facts.

- [5] Hagit Borer. Exo-skeletal vs. endo-skeletal explanations. In John Moore and Maria Polinsky, editors, *The Nature of Explanation in Linguistic Theory*, pages 31–67. CSLI Publications, Stanford, CA, 2003.
- [6] Gregory Carlson. On the role of thematic roles in linguistic theory. *Linguistics*, 22:259–279, 1984.
- [7] Marie-Catherine de Marneffe, Bill MacCartney, and Christopher D. Manning. Generating typed dependency parses from phrase structure trees. In *LREC*, 2006.
- [8] Stefanie Dipper, Michael Götze, and Stavros Skopeteas, editors. *Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics and Information Structure*. Universitätsverlag Potsdam, 2007.
- [9] David Dowty. On the semantic content of the notion ‘thematic role’. In Gennaro Chierchia, Barbara Partee, and Raymond Turner, editors, *Properties, Types, and Meaning, II*, pages 69–129. Kluwer, Dordrecht, 1989.
- [10] Lelia Glass. If looks could kill: Honing a pragmatic account of English implicit objects. Talk presented at the 37th Penn Linguistics Colloquium, 2013.
- [11] Jane Grimshaw. *Argument Structure*. MIT Press, 1990.
- [12] James Higginbotham. Logical form, binding, and nominals. *Linguistic Inquiry*, 14:679–708, 1983.
- [13] N. Hornstein. S and X’ convention. *Linguistic Analysis*, 3:137–176, 1977.
- [14] Andrew McIntyre. Argument linking and complex predicates in English nominalizations. Talk presented at JeNom5, UPF, 2013.
- [15] Louise McNally, Gemma Boleda, and Marco Baroni. Conceptual vs. referential affordance in concept composition. Talk presented at the Workshop on Concept Composition and Experimental Semantics/Pragmatics, U. Utrecht, 2013.
- [16] Åshild Næss. The grammar of eating and drinking verbs. *Language and Linguistic Compass*, 4/5:413–423, 2011.
- [17] Frederick Newmeyer. Current challenges to the lexicalist hypothesis. In William D. Lewis, Simin Karimi, Heidi Harley, and Scott O. Farrer, editors, *Time and Again: Theoretical Perspectives on Formal Linguistics*, pages 91–117. John Benjamins, Amsterdam, 2009.
- [18] Terence Parsons. *Events in the Semantics of English: A Study in Subatomic Semantics*. Current Studies in Linguistics Series 19. MIT Press, Cambridge, MA, 1990.
- [19] Barbara Partee. Uniformity vs. versatility: the genitive, a case study. In Johan van Benthem and Alice ter Meulen, editors, *The Handbook of Logic and Language*, pages 464–470. Elsevier, Amsterdam, 1983/1997.
- [20] Malka Rappaport. On the nature of derived nominals. In B. Levin, M. Rappaport, and A. Zaenen, editors, *Papers in Lexical-Functional Grammar*, pages 113–142. Indiana University Linguistics Club, 1983.
- [21] Malka Rappaport Hovav and Beth Levin. An event structure account of English resultatives. *Language*, 77:766–797, 2001.
- [22] Alessandro Zucchi. *The Language of Propositions and Events*. Kluwer, Dordrecht, 1993.