

Heuristic Disambiguation of Deverbal Nominals in Greek

Vassiliki Rentoumi*, Stasinou Konstantopoulos*

*Institute of Informatics and Telecommunications,
NCSR ‘Demokritos’, Ag. Paraskevi 15310, Athens, Greece
{vrentoumi, konstant}@iit.demokritos.gr

Abstract

Deverbal nouns exhibit both nominal and verbal syntactic behaviour: they operate as concrete nouns, but also participate in verbal constructions where they require arguments and accept aspectual modification. We first describe the phenomenon and discuss theoretical analyses proposed in the literature, and then propose a methodology for applying the theoretical results to parsing of nominal phrases of Greek.

1. Introduction

Deverbal nouns are derived from verbs or verb phrases and they operate as concrete nouns, but also participate in verbal constructions where they require arguments and accept aspectual modification. This poses an interesting problem, theoretically as well as computationally.

There is extensive theoretical linguistics literature investigating the derivation of deverbal nouns and proposing explanations and analyses of their syntactic idiosyncrasies. We review the literature in Section 2., while describing the phenomenon and presenting examples from Greek. We focus on Greek data and present a treatment of Greek nominal phrases, although the phenomenon appears in various other languages as well, including English and German. Although the exact treatment presented here does not directly apply cross-linguistically, some of the analysis and computational methodology carries over to other languages.

The most striking of these idiosyncrasies is that deverbal nouns participate in constructions that are—on the surface—identical to genitive and prepositional-phrase complementization, but have completely different syntactic structure and semantics. Both analyses are sometimes tenable, but truly ambiguous phrases are rare and it is usually possible to disambiguate on the basis of semantics and pragmatics. From a computational point of view, this gives rise to an interesting problem of selecting the right structure to assign on the basis of semantic and contextual ‘hints’, discussed in Section 3..

We close with a presentation of empirical results, some concluding remarks, and an outlook on future research on the subject.

2. Deverbal Nominals

Lebaux (1986) and Grimshaw (1990) identify two types of deverbal nominals, *complex event nominals*¹ and *result nominals*. The former exhibit a strong verbal character and support verbal arguments and aspectual modification, where the latter share the characteristics of concrete nouns and do not support any arguments or aspectual modifiers. This distinction is a very important one, as it gives rise to radically different syntactic structures and semantics.

To demonstrate the phenomenon, let us observe and compare the following examples:

- (1) *H koupa tou Yianni apo ksylo*
The cup the-GEN John from wood
John’s wooden cup
- (2) *i syllogi tou Yianni apo palious diskous*
the collection the-GEN John from old records-ACC-PL
John’s collection of records
- (3) *i syllogi ton stoixeion apo eidikeumenous epitheorites*
the collection the-GEN evidence from specialist inspectors-ACC
The collection of the evidence by specialist inspectors

In Example 1 the theme, ‘cup’, is a concrete noun, and ‘John’ and ‘wooden’ are adjuncts. Example 2 is very similar, in that it again denotes a concrete collection, with adjuncts that denote the possessor and the kind of items collected. In these two examples we see nominals complemented by *thematic adjuncts* (Markantonatou, 1993, 1995) that specify *optional* complements. In this reading, ‘collection’ is a result nominal which behaves like a concrete noun and denotes the outcome of the collecting process; that is, the set of things collected.

Example 3, on the other hand, has a radically different reading, despite its superficial resemblance. In this example the deverbal nominal *syllogi* is an event nominal and denotes the process (and not the result) of building a collection, the event of collecting evidence. Its analysis yields a syntactic tree with the same positions for thematic roles as the verb *syllago*. By contrast to the analysis of the same noun as a result nominal, this tree includes a position for an obligatory theme argument, and, naturally, the two analyses also yield different semantics.

Alexiadou and Stavrou (1999) formalize the verbal character of deverbal nouns, by proposing the presence of verbal projection nodes inside nominalization. The appearance of adverbs in event nominal phrases is offered as evidence in favour of this analysis, as it is well known that

¹Or, here, simply *event nominals*.

adverbial modification is only consistent with verbal structures.

Moreover, they argue that nominalizations should be derived from a verbal node, otherwise the presence of an adverb could not be accounted for. The presence of the adverbs is not simply allowed because of the process interpretation of the nominal, but it is rather syntactically conditioned. Nouns with the above eventive reading are assumed to project a voice phrase (vP) and an aspect phrase, in addition to the standard Number/Agreement phrase (FP) stipulated for results nouns.

Building on this, they unify the analysis of deverbal nominals and their verbal counterparts, by adopting Marantz' (1999) assumption that all word formation, including nominalization, takes place in the syntax. Under this assumption, they argue that deverbal nominals are not derived from their verbal correlates, but rather that both nouns and verbs emerge from common abstract roots that surface as nouns or verbs once inserted into nominal or verbal (functional) structures.

The functional categories vP and AspectP can occur under either tense or determiner projections (TP and DP, resp.). In the latter case, the embedding of TP and AspectP categories under a DP results in a morphological noun.

Under this analysis, nominals with an eventive reading produce a syntactic tree which is similar to that of a verbal structure. Consider the following examples:

- (4) *i syllogi stoixeion adiakopa*
the collection evidence-GEN constantly
the constant collection of evidence
- (5) *ta stoixeia syllegontai adiakopa*
the evidence is collected constantly
evidence is being constantly collected
- (6) * *i syllogi adiakopa*
the collection constantly

All eventive deverbal nominals have a passive character. Thus, in Example 4, the event nominal *syllogi* must appear with its subject (*stoixeion*) and may have an aspectual modifier (*adiakopa*). Figures 2 and 3 show the syntactic trees for the deverbal nominal phrase in Example 4 and the corresponding verbal phrase of Example 5. These examples are contrasted with the result nominal phrase shown in Figure 1. Note that Example 6 is ungrammatical since there is no filler for the genitive argument position.

In this manner, event nominal phrases denote the internal organization of an event, the unfolding of an event which seems to be syntactically more verb-like—and thus argument-supporting. Result nominals, on the other hand, denote the outcome of an event, behave like concrete nouns, and are analysed using only nominal-projection categories.

3. Disambiguation Criteria

Let us revisit Examples 1–3. In these examples we show strings of words where word order, part-of-speech, and inflectional information alone is not sufficient in order

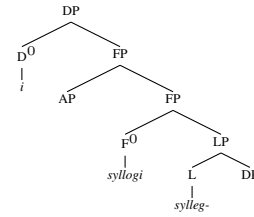


Figure 1: Tree structure of the result deverbal-nominal phrase *i syllogi*.

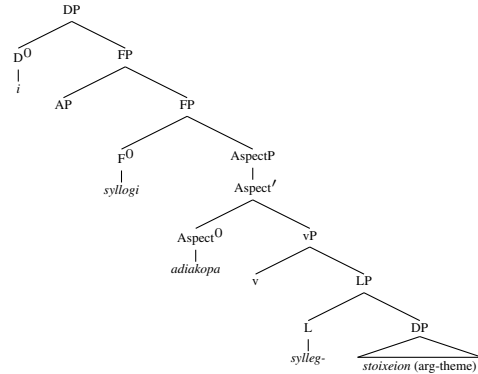


Figure 2: Tree structure of the event deverbal-nominal phrase *i syllogi stoixeion adiakopa*.

to disambiguate between two radically distinct readings. Lexicographically or morphologically identifying the class of deverbal nominals will sort out concrete nouns like the one in Example 1, but it will not help us disambiguate between result and event readings of deverbal nominals, since the same word-form might appear under both a result and an event lexical entry.

Markantonatou (1993), following Grimshaw (1990), proposes a series of criteria for disambiguating between the event and result reading of deverbal nominals, which are presented here.

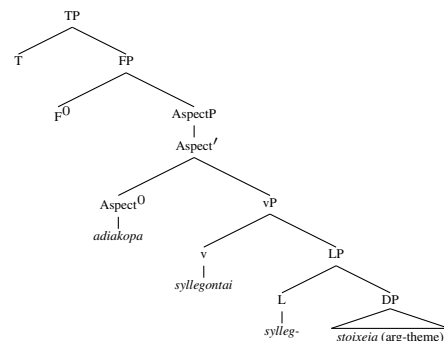


Figure 3: Tree structure of the verbal phrase *syllegontai stoixeia adiakopa*.

3.1. Aspectual Adjuncts

In this context the term aspect means the lexical aspect of a verb. The lexical aspect, denotes the internal temporal structure of an event, or event type; that is, the relationship between the verb, its arguments and the type of event that the two, taken together, denote.

As a consequence of lexical aspect's being a verbal property, it can only characterize event deverbal nominals which project an Aspect Phrase in the syntactic tree.

Concerning the first aspectual adjunct criterion, we already have an insight from the examples presented above that the eventive reading of a nominal admits the presence of aspectual adjuncts (Example 4), while the result reading precludes aspectual modification (Example 6).

In a broader sense we use as a diagnostic of the eventive reading of deverbal nominals the presence of aspectual modifiers (Alexiadou, 1994) or *aktionsart* modifiers (Moens, 1987) in nominalizations.

Aspect interacts with aspectual adverbials, which point to the internal temporal properties of a situation, e.g. duration, frequency. (Alexiadou, 1994) classifies aspectual adverbials as follows:

- Durative: *gia mia ora* 'for an hour'
- Indefinite frequency: *kathimerina* 'daily', *kathe toso* 'every so often'
- Definite frequency (cardinal count): *mia fora* 'once'
- Point adverbials: *amesos* 'immediately', *ton Iounio* 'in June'
- Completive adverbials: *se mia ora* 'in an hour'

Furthermore, as far as the distribution of adverbs is concerned, in derived nominals we observe that the type of adverbs that are acceptable inside nominalizations with eventive reading are manner adverbs (Example 7), *aktionsart* adverbs (Example 8), and temporal adverbs (Example 9):

(7) *i katastrofi ton stoixeion toso*
the destruction the evidence-GEN so
prosektika
carefully
The destruction of the evidence so carefully

(8) *i katastrofi ton stoixeion entelos*
the destruction the-GEN evidence completely
The destruction of the evidence completely

(9) *i katastrofi ton stoixeion xthes*
the destruction the-GEN evidence yesterday
the destruction of the evidence yesterday

Additionally only aspectual adjectives can modify event nouns, which are related with the lexical aspect of the nominal. Moreover it is claimed that as the presence of aspectual adjuncts and complements constitute a strong evidence for the eventive reading of a nominal, modal and temporal adjectives are not compatible with the eventive

reading of a derived noun, as they denote the result character of a noun. The above are shown in the following examples:

(10) *stin synexi apokrypsi*
to the-ACC continuous concealment
stoixeion apo tin kyvernisi
evidence-GEN from the government
in the continuous/repetitive concealment of evidence by the government

* *stin synexi apokrypsi apo*
to the-ACC continuous concealment from
tin kyvernisi
the government

(11) * *i persini/pithani syllogi*
the last year's/probable collection
ton stoixeion adiakopa
the-GEN-PL evidence constantly

Note that the presence of adverbs, inside nominalizations with an eventive reading, is not present in English or German, but is not unique to Greek either, as it is also observed in Hebrew (Alexiadou and Stavrou, 1999).

3.2. Agentivity

Grimshaw (1990) first pointed out that event deverbal nominals accept agent-oriented modifiers, while result nominals do not.

Agent oriented or *voice modifiers*, are some of the participles in *-menos*. They are derived from verbs that denote the intention of an agent, e.g. *thelo-ithelimenos* 'want', *skopo-eskemenos* 'have purpose', *prosexo-prosegmenos* 'attend'. Such participles denote the implication of an event and bear agentivity.

Additionally there exist adjectives as well as adverbs which bear the notion of an agent, called *agent-oriented* adjectives and adverbs, e.g. *prosektikos-prosektika* 'careful-carefully'.

As they imply an event, agent-oriented modifiers are more likely to modify deverbal nominals that yield an eventive reading. Thus, their presence before a deverbal nominal constitutes an indication in favour of an eventive reading.

(12) *i ithelimeni aposiopisi simantikon*
the intentional withholding important
stoixeion epi tosa xronia dimiourgise
evidence-GEN for all these years created
asximi entyposi
ugly impression
The intentional withholding of important evidence for all these years, gave a bad impression

(13) * *i ithelimeni aposiopisi epi tosa*
the intentional withholding for all these
xronia dimiourgise asximi entyposi.
years created ugly impression.

3.3. Iteration

Markantonatou (1993) claims that in some cases the eventive reading of nominals in Greek is present only when nominals are inflected for plural number. This happens for nominals that correspond to verb predicates that do not have internal temporal structure. It is essential to observe this phenomenon in nominals derived from unaccusative predicates that have inherently telic reading, e.g. *afiknoumai* ‘arrive’. In this way, derived nominals obtain the durative/iterative interpretation when the reference of their theme argument changes from a quantized to a cumulative one. When the theme is quantized the example becomes ungrammatical. When these nominals as well as their themes are inflected for plural, their aspectual interpretation changes from perfective to imperfective, thus an eventive reading is obtained.

(14) * *i afiksi touriston oli ti nyxta*
the arrival tourists-GEN all the night

(15) * *i afiksi tourista oli ti nyxta*
the arrival tourist-GEN all the night

(16) *oi afikseis touriston oli ti nyxta*
the arrivals tourists-GEN all the night
the arrivals of tourists during the whole night

3.4. Thematic Adjectives

Thematic adjectives are typically derived from ethnic nouns, and are also called *ethnic* (Kayne, 1994) or *referential* (Longobardi, 1991) adjectives. A comparison of Examples 17 and 19 suggests that thematic adjectives can only assume the agent role.

Alexiadou and Stavrou (1999), on the other hand, doubt the above assumption and propose a revised formulation of this restriction: thematic adjectives are not genuinely thematic, in the sense that they behave as classifiers, cannot absorb a thematic role from the nominal, and are always mapped to the specifier position. Consider, for example:

(17) *i germaniki katoxi*
the German occupation
The German occupation

(18) * *i germaniki katoxi tis Elladas*
the German occupation the-GEN Greece

(19) * *i elliniki katoxi apo tous germanous*
the Greek occupation by the Germans

Relying on the passive character of the deverbal nominals with eventive reading, deverbal nominals—like their verbal counterparts—do suppress the agent role; in this way, in Example 18 above, the adjective *germaniki* violates the θ -criterion. In this case the agent role should be expressed optionally by a prepositional phrase. As claimed above, thematic adjectives are mapped in a specifier position of an NP-shell. Moreover event nominals that are derived from transitive verbs require their internal arguments

to be realized into complement position. It is obvious that in Example 19 the Projection Principle is violated as the complement—theme—is missing.

Under this analysis, thematic adjectives are not compatible with the eventive reading of deverbal nominals and force a result reading upon the noun they modify.

Furthermore, the above analysis predicts a result meaning for *katoxi* in Example 17, rendering aspectual modification ungrammatical. Indeed, if we compare Example 17 against:

(20) * *i germaniki katoxi mesa se liges*
the German occupation within few
meres
days

we see that Example 17 resists aspectual modification.

4. Corpus Evaluation

In order to get some idea of the applicability of the proposed methodology, we used a manually Part-of-Speech tagged corpus of Greek (Petasis et al., 1999). The corpus comprises text from *AdWeek*, an on-line newsletter for the advertisement industry, and from *Europarl*, the transcripts of the sessions of the European Parliament. Its size is 5530 sentences (around 136k words).

The Part-of-Speech tags in the corpus provide gender and number information for nominals. Combining the tags with the endings in the word-forms we were able to retrieve partial case information. More precisely, we were able to retrieve the distinction between genitive and the other cases². The most interesting cases were, naturally, the sentences where nominals are succeeded by genitive nominals. The corpus contained 6427 instances of this pattern, found in 3071 sentences.

We applied the adverbial adjunct, aspectual adjective adjunct, and agentivity criteria to these sentences, and retrieved 179 instances were one of the criteria recognized an event reading. We manually checked these 179 instances, and found 84 event deverbal nominals, 58 result deverbal nominals, and 37 concrete (non-deverbal) nouns (Table 1). This gives us an overall accuracy of 47% if no lexical information is available, and 59% if a lexicon is used to distinguish between concrete nouns and deverbal nominals. This discrepancy is due to the aspectual adjective modification criterion, which is very accurate in distinguishing between eventive and result readings, but cannot distinguish between deverbal nominals and concrete nouns, which might very well be modified by aspectual adjectives.

With respect to the result reading falsely retrieved, it should be noted that the single instance falsely accepted by the aspectual modification criterion is a plural noun, and thus, this result would be correctly classified by the subsequent application of the singular definite nominal criterion. The (many more) false positives due to the adverbial adjunct criterion are cases of adverbs that fit into

²Because of the suffix overlaps in Greek nominal inflection, only this coarse distinction can be made even in the presence of grammatical gender and number tags.

Criterion Triggered	Deverbal Nouns		Concrete Nouns	Sum
	Event	Result		
Adverb	59	57	0	116
Aspectual Adj.	17	1	37	55
Agentivity	8	0	0	8
Sum	84	58	37	179

Table 1: The correct reading of the 179 nominal phrases that triggered the criteria described in Sect. 3.

the nominal-genitive nominal-adverb pattern, but, in fact, modify the verb to their right rather than the nominal phrase to their left. This phenomenon makes the adverbial adjunct criterion weaker, as it cannot be applied in a shallow parse before phrase structure is fully determined.

In the opposite direction, we tried the thematic adjective criterion to identify result deverbal nominals. We looked at 285 sentences where thematic adjectives appear, and found no instances whatsoever of event deverbal nouns.

5. Conclusions

The empirical results presented here are very promising, and appear to verify the theoretical results. They also provide valuable insight with respect to the order of application, as, for example, we have seen the singular definite nominal criterion correct a mis-classification by the aspectual adjective criterion.

The other interesting observation is that the most wide-coverage criterion (namely, the adverb adjunct criterion) is also the least accurate one, at least in the shallow manner of application used here. The aspectual adjective and agentivity criteria are much more accurate, but have a much smaller domain of application and also require access to the relevant lexical information about adjectives and participles.

These results prompt us to pursue this line of research further.

6. Future Work

The work discussed in this paper is part of an on-going project to develop a *Combinatory Categorical Grammar* (CCG, Steedman 1996) account of Greek, for use with the OpenCCG parser and realizer³. At its current state of development, the Greek grammar features an almost complete account of Greek nominal phrases, following the analysis of Alexiadou and Anagnostopoulou (2000), and a fragmentary account of verbal phrases and clauses.

First, and most importantly, in order to be able to more accurately assess the proposed criteria we intend to try them on a treebank of Greek, from where we can retrieve all deverbal nominals and their correct readings. Furthermore, we are planning to collect and implement a cross-linguistic set of criteria, and then propose a methodology

for selecting the ones that perform best for each given language.

Another possible improvement that can be tried is enriching the lexical information that is taken into account. One lexical classification that might prove useful is the lexical aspect of the deverbal nominal, and especially the way that it interacts with aspectual modification.

7. Acknowledgments

The work described here was carried out in the context of a project for implementing a CCG grammar of Greek, for the OpenCCG parser and realizer. The project was subcontracted to NCSR ‘Demokritos’ by the University of Edinburgh.

References

- Alexiadou, Artemis, 1994. *Issues in the Syntax of Adverbs*. Ph.D. thesis, University of Potsdam.
- Alexiadou, Artemis and Elena Anagnostopoulou, 2000. Greek syntax: a Principles and Parameters perspective. *Journal of Greek Linguistics*, 1:169–221.
- Alexiadou, Artemis and Melita Stavrou, 1999. Derived nominals in Greek. *Linguistic Review*, 1(2):22–25.
- Grimshaw, Jane, 1990. *Argument Structure*. Cambridge, Mass.: MIT Press.
- Kayne, Richard, 1994. *The Antisymmetry of Syntax*. Cambridge, Mass: MIT Press.
- Leboux, David, 1986. The interpretation of derived nominals. In *Proceeding of Chicago Linguistics Society*, number 22.
- Longobardi, Giorgi, 1991. *The Syntax of Noun Phrases: Configuration Parameters and Empty Categories*. Cambridge: Cambridge University Press.
- Marantz, Alec, 1999. *Creating Words Above and Below Little*. Cambridge, Mass.: MIT Press.
- Markantonatou, Stella, 1993. *The Syntax of the Modern Greek Noun Phrases with a Deverbal Nominal Head*. Ph.D. thesis, University of Essex.
- Markantonatou, Stella, 1995. Modern Greek deverbal nominals: An LMT approach. *Journal of Linguistics*, 31:267–299.
- Moens, Marc, 1987. *Tense, Aspect and Temporal Reference*. Ph.D. thesis, Centre for Cognitive Science, University of Edinburgh.
- Petasis, Giorgos, Giorgos Paliouras, Vangelis Karkaletsis, C. D. Spyropoulos, and Ion Androutsopoulos, 1999. Using machine learning techniques for Part-of-Speech tagging in the Greek language. In *Proceedings of the 7th Hellenic Conference on Informatics, Ioannina, Greece*.
- Steedman, Mark, 1996. *Surface Structure and Interpretation*. Cambridge, Mass.: MIT Press. Linguistic Inquiry Monograph No. 30.

³OpenCCG: The OpenNLP CCG Library. See <http://openccg.sourceforge.net/>