

# **A syntactic analysis of synthetic and phrasal compound formation in Greek<sup>1</sup>**

Dimitrios Ntelitheos  
United Arab Emirates University  
Dimitrios\_N@uaeu.ac.ae

## **1. Introduction**

This paper argues for a syntactic approach to deverbal synthetic compound formation in Greek. Synthetic (or verbal) compounds have been at the forefront of the discussion of the morphology-syntax interface because of their mixed morphosyntactic properties. The decision of whether they are formed exclusively within the morphological or the syntactic component of the grammar is important for deciding the status of morphology as an independent module of grammar.

The work presented here shows that recent research insights from syntactic theory provide a number of tools that can adequately answer questions usually swept under the carpet in syntactic approaches to morphological derivations. In particular I show that the ‘split-DP’ hypothesis, defended in Sportiche (1999, 2005) and the “phases-within-words” hypothesis in Marantz (2001, 2006) can shed new light on why so-called “lexical integrity effects” appear in synthetic compounds. In addition, the differences in the distribution and prosodic properties of various compound forms, including synthetic and phrasal compounds, are shown to be derived from properties of the lower verbal domain and the variable positional and selectional properties of nominalizing affixes. In fact, it is shown that the merger of nominalizing affixes, at least for languages of the Greek and English type, is essential for the successful derivation of such compounds.

The domain of synthetic compound formation is the root domain, i.e. the domain before a head determining categorical status is merged. The nominalizer marks the boundary of the root domain, and thus determines the level at which non-transparent semantics and other idiosyncrasies arise (Marantz 2001, 2006). The root domain contains the base and a single argument that has not been quantized (i.e. become referential yet) (Sportiche 1999). Attaching the nominalizer turns the domain into a ‘phase’ and provides it with phase properties, i.e. island-hood and prosodic autonomy, the two properties connected with lexical integrity effects and stress assignment in synthetic compounds.

The paper is organized as follows: section 2 presents some of the basic questions concerning synthetic compounds and a summary of the literature that addresses these questions. It also presents some of the well-established properties of synthetic compound formation and the lexical-integrity effects that the latter exhibit with reference mainly to Greek compound deverbal nominals; section 3 presents the main proposal of a phase-based analysis of synthetic compound formation and provides an explanation of these effects; section 4 discusses phrasal compounds and shows how their frequent non-transparent readings form an important problem for strictly morphological and lexicalist accounts of word formation; ; and finally section 5 presents concluding remarks.

---

<sup>1</sup> I would like to thank the participants at the Mediterranean Syntax Meeting-II, 16-18 October 2008, Boğaziçi University, Istanbul, Turkey, for their questions and comments. I would also like to thank an anonymous reviewer for raising a number of issues concerning the analysis proposed here and for valuable comments. I have tried to address these concerns in the final version. Any errors in the work presented here are of course my own.

## 2. Morphology or Syntax

### 2.1 Properties of Synthetic Compounds

Synthetic compounding has been at the forefront of syntactically oriented morphological research since the work of Roeper and Siegel (1978) and extending through Williams (1981); Selkirk (1982); Lieber (1983); Pesetsky (1985); Di Sciullo and Williams (1987); Booij (1988; 2005); Roeper 1988; Di Sciullo (2005); Di Sciullo & Ralli (1999).

Roeper and Siegel (1978) provide an analysis which, even though lexicalist in spirit, encompasses a transformational aspect in postulating a movement rule. In this kind of ‘lexical transformation’ the internal argument of the verb moves from its postverbal position in the verb’s lexical subcategorization frame to a preverbal ‘empty’ position. The movement must obey a First Sister Principle (Roeper & Siegel, 1978:208) which states that only a ‘word’ that is the first sister of the head can move to the empty slot. Restriction to words rules out phrasal compounds, while the first-sister restriction rules out verb-goal synthesis in the presence of a theme.

In many ways this first post-‘Remarks’ (Chomsky 1970) approach to synthetic compounding, incorporates aspects of both of the two main approaches that have dominated the relevant literature since then: the lexicalist approach and the syntactic approach. The first assumes a separate grammatical component for the operations that give rise to the derivation of synthetic compounds and similar structures (see among others Williams (1981); Selkirk (1982); Di Sciullo and Williams (1987); Booij (1988); Di Sciullo (2005). The second assumes a syntactic analysis of synthesis which adopting a number of different tools from the syntactic inventory, claims that synthetic compounds are formed from parts that are syntactic “atoms” and through processes that are independently motivated by purely syntactic mechanisms (see Fabb 1984; Sproat 1985; Roeper 1987, 1988).

Synthetic compounds are formed by at least three basic morphosyntactic units: the verbal head, a verbal argument (in most cases the internal argument but see below), and a derivational morpheme (usually a nominalizer) as in the following example from Greek:

- (1)      pliroforio - dho -tis  
             information - give -er  
             ‘informer’

The structure immediately poses a bracketing problem: the verb forms a morphological constituent with its argument before the nominalizer applies, as in (2a), or the compound is formed by two nominal elements, the verb plus nominalizer complex and the nominal argument, as in (2b):

- (2)
- a.

```

graph TD
    N1[N] --- V1[V]
    N1 --- N2[N]
    V1 --- N3[N]
    V1 --- V2[V]
    N3 --- pliroforio[pliroforio]
    V2 --- dho[dho]
    N2 --- tis[tis]
          
```

b.

```

graph TD
    N1[N] --- N2[N]
    N1 --- N3[N]
    N2 --- pliroforio[pliroforio]
    N3 --- V[V]
    N3 --- N4[N]
    V --- dho[dho]
    N4 --- tis[tis]
          
```

The analysis of synthetic compounds as N-N root compounds (c.f. (2b)) has been proposed in a number of studies, most notably Selkirk 1982; Di Sciullo & Williams 1987; Ralli 1989; Booij 1988; 1992; Di Sciullo and Ralli 1999, and others. There are a number of facts that seem to support the analysis. Languages like English allow formation of N-N compounds freely and such formations are very productive in the language. On the other hand N-V compounds, where the nominal element is the internal argument of the verb are very rare (Ackema & Neeleman 2004:55). Thus, for proponents of the analysis in (2b), the analysis in (2a) cannot be maintained, as the first constituent is not a possible independent structure in the language. Similarly, the nominalizing suffix *-tis* merges productively with verbs to produce nominals and therefore, both ingredients of the analysis in (2b) are independently motivated.

The only potential problem for this analysis would be explaining the presence of a verbal internal argument in the configuration, i.e. the fact that the verbal base has one of its arguments saturated inside the compound:

- (3) \* O Giorgos ine pliroforiodhotis hrisimon pliroforion.  
 D Giorgos is information.giver useful information.GEN  
 ‘Giorgos is an informer of useful information.’

In order to explain this, proponents of the analysis in (2b) have to assume some sort of inheritance whereby the internal argument of the verbal base is inherited by the nominalization. However, in cases where the argument is assumed to have been inherited by the verbal base, idiomatic readings are usually excluded (Ackema & Neeleman 2004:56):

- (4) a. John has always made trouble.  
 b. \* John has always been a maker of trouble.

On the other hand, synthetic compounds clearly allow for idiomatic readings to be maintained:

- (5) John has always been a trouble-maker.

Thus, an argument-inheritance analysis is problematic. Finally, the analysis in (2b) has nothing to say about the unavailability of N-V compounds, where N is the internal argument of V.

The presence of an internal-argument and the availability of idiomatic readings are not problematic for the proponents of the analysis in (2a) (see Lieber 1983; Fabb 1984; Sproat 1985; Ackema and Neeleman 2004 among others). In these proposals the nominal argument merges directly with a verbal base and thus argument saturation is triggered by the independently motivated selectional restrictions on the verb. Locality of selection at that level would also explain the presence of idiomatic readings. But this looks more and more like a syntactic configuration and this has led some of the relevant research to the conclusion that the processes involved in synthetic compound formation are independently motivated syntactic processes and there is no need for the involvement of a separate morphological component (Fabb 1984; Sproat 1985; Roeper 1987; 1988). In most of these analyses the internal structure of synthetic compounds is assumed to be syntactic in nature, involving

elements of the basic head level ( $X^0$ ), which is an atom for the purposes of syntactic operations.

Greek synthetic compounds exhibit the same word order as English synthetic compounds with the nominal argument being the leftmost sister of the compound, while the root appears on the right, followed by derivational and inflectional (gender/number) morphology. A difference between English and Greek synthetic compounds lies in the presence of a ‘linking’ morpheme *-o-* that appears between the verbal domain and the nominal argument. Thus, (1) is better represented as (6):

- (6)
- |             |     |        |      |
|-------------|-----|--------|------|
| plirofori   | -o  | - dho  | -tis |
| information | LNK | - give | -er  |
| ‘informer’  |     |        |      |

The existence of the linking morpheme seems to formally spell-out the compounding process, and its presence or not in a language may be related to the morphological and/or phonological properties of the language (Ralli 2007). In other approaches (c.f. DiSciullo 2005) it is assumed to head a compound-internal functional projection that imposes an asymmetry in the way the compound elements relate to each other. I will not pursue this issue here as it does not directly bear on the analysis presented but I will follow Di Sciullo’s proposal as it complies with the antisymmetric approach to syntax (Kayne 1994) which provides a principled explanation of the surface word order in synthetic compounds (See section 3 for details).

As far as I know the only attempt to propose a syntactic analysis for a process in Greek that was traditionally assumed to be morphological in nature is Rivero (1992). Based on the general process of syntactic incorporation (via head-movement) as detailed in Baker (1988), Rivero proposed that Greek synthetic compounds whose first element is either a verbal argument or a manner adverbial involve nominal or adverbial incorporation to the verb. This analysis met with strong criticism (see for example Kakouriotes et al 1997; Smirniotopoulos & Joseph 1997, 1998). The criticism against the analysis was based on a series of lexicalist arguments, including lexical integrity effects, semantic transparency, and productivity. In particular, the case of lexical integrity has been at the forefront of the polemic against syntactic analyses of synthetic compounding, most notably in the case of Greek synthetic compounds, in the work of Ralli (1989; 1992; 1999; 2003; 2007); Ralli and Stavrou (1998); Di Sciullo and Ralli (1999). In the remainder of this paper I will address this issue in detail but will also provide an explanation of semantic opacity phenomena, in the spirit of Marantz (2001, 2006).

## 2.2 Lexical Integrity

Lexical Integrity refers to the fact that ‘words’ resist syntactic operations that apply to ‘sub-lexical’ (i.e. word-internal) elements (c.f. Lapointe 1980; Selkirk 1992; Bresnan & Mchombo 1995; Di Sciullo and Williams 1987; Anderson 1992; Boij 2005; Spencer 2005 and others). Bresnan & Mchombo (1995:181f) define lexical integrity as “*syntactic principles do not apply to morphemic structures. Morpheme order is fixed, even when syntactic word order is free; the directionality of ‘headedness’ of sublexical structures may differ from supralexical structures; and the internal structure of words is opaque to certain syntactic processes.*” Thus, lexical integrity is an umbrella term for a number of different phenomena that involve sublexical elements, including inbound anaphoric islands, availability of phrasal recursivity,

conjoinability, gapping, and displacement. Let us examine some of these cases with respect to Greek synthetic compounds.

Word parts in general are not possible links in chains, ruling out movement into and out of words (Chomsky 1970, Di Sciullo & Williams 1987, Bresnan & Mchombo 1995).

- (7) a. O Giannis ine plirofori-o-dotis.  
The Giannis is information- LNK-giver  
'Giannis is an informer.'
- b. \* Ti ine o Giannis dotis?  
What is the Giannis giver?  
'What is Giannis giver (of)?'
- c. \* Ti ine o Giannis plirofori(as)?  
What is the Giannis information(GEN)?  
'What is Giannis (of) information?'

As we can see in (7b), trying to question the leftmost element of a synthetic compound results in ungrammaticality, and thus, no *wh*-movement operation can displace a compound-internal element. A similar type of ungrammaticality is observed in attempting to *wh*-question the rightmost element in the compound (7c).

A stronger lexical-integrity-based argument comes from the fact that words seem to be anaphoric islands – sub-lexical elements cannot be referential. Binding relations or coreference cannot relate elements below the word level to material external to the word (Postal 1969, Di Sciullo & Williams 1987):

- (8) a. \* O Giannis ine plirofori<sub>1</sub>-o-dotis                      alla den tin<sub>1</sub> gnorizo.  
the Giannis is information- LNK-giver                      but not it know.1SG  
'Giannis is an informer but I don't know it (the information).'
- b. \* O kapn<sub>1</sub>-o-kalliergitis      dhen katafere      na ton<sub>1</sub> poulisi fetos.  
The tobacco- LNK-grower neg manage.pst to it sell this.year  
'The tobacco-grower didn't manage to sell it (the tobacco) this year.'
- c. O Giannis edhose [mia pliroforia]<sub>1</sub> ston Giorgio alla den tin<sub>1</sub> gnorizo.  
The Giannis gave.3SG an information to.the Giorgio but not it know.1SG  
'Giannis gave some information to Giorgio but I don't know what it is.'

In (8a) the leftmost part of the compound, *plirofori-* (information) is supposed to be coreferential with the 3SG.FEM clitic pronoun *tin*, but the sentence is completely ungrammatical in Greek. Compare the grammatical sentence in (8c) where the antecedent of the clitic is now a fully independent DP. A second ungrammatical example is provided in (8b) where the compound-internal nominal argument *kapn-* (tobacco) is supposed to bind the 3SG.MSC clitic pronoun *ton*.

A further argument from lexical integrity comes from the fact that co-ordination or gapping word-internally is not allowed (Bresnan & Mchombo 1995). Coordination of compound-internal elements would allow, for example, two or more nominal arguments to appear conjoined as the leftmost element in a compound:

- (9) \* O Giannis ine plirofori-o (ke) em-o-dotis.  
 The Giannis is information- LNK- (and) blood- LNK- giver  
 ‘Giannis is an information-(and)-blood-giver.’

However, as (9) illustrates this is not possible. One can argue here that the impossibility of (9) stems from the inability of the conjunction *ke* (and) to appear word-internally. However, Greek has so-called copulative compounds that seem to encode coordination:

- (10) alat-o-pipero  
 salt-LNK-pepper  
 ‘salt and pepper’

Given this property of Greek grammar, it could be possible to form a compound whose leftmost element is itself a copulative compound. However, as (11) illustrates, this is also impossible:

- (11) \* O Giannis ine plirofori-o-em-o-dotis.  
 The Giannis is information-LNK-blood-LNK- giver  
 ‘Giannis is an information-blood-giver.’

Similarly, gapping constructions where the leftmost or rightmost elements of the synthetic compound are elided because of the presence of an antecedent in the immediately preceding environment are ruled out. Compare (12a) to (12b):

- (12) a. O Giannis kalliergi kapno ke i Maria ~~kalliergi~~ sitira.  
 the Giannis grows tobacco and the Maria ~~grows~~ wheat  
 ‘Giannis grows tobacco and Mary wheat.’  
 b. \* O Giannis ine ele-o-kaliergitis (ke) i Maria sitir-o- ~~kaliergitis~~.  
 The Giannis is olive- LNK-grower (and) the Mary wheat- LNK- ~~grower~~.  
 ‘Giannis is an olive-grower and Mary a wheat-grower.’

In (12a) the verb in the second conjunct can be deleted (presumably due to the fact that the verb is ‘given’ information), but no such deletion of the verbal base is available in the compound of (12b) even though the context (and therefore the informational make-up of the sentence) remains the same.

A final issue arises with the distribution of definite determiners and other elements that arguably make use of the D-projection of the noun phrase (pronouns, proper names, and so on). Definite determiners and referential nouns (pronouns and proper names) are in general excluded in synthetic compound formation (c.f. Postal 1969; Di Sciullo and Williams 1987):

- (13) a. \* [o kapno-] kalierjia  
 [D tobacco-] cultivation  
 b. \* Giorgo-thavmastis  
 Giorgo-admirer

Finally, outside lexical integrity effects, a further restriction in the formation of synthetic compounds that has been noticed is that in most cases additional arguments (i.e. goals, benefactors, and so on) are excluded when the internal argument is present (c.f. Grimshaw 1990):

- (14) a.      emo-o-dotis  
              blood- LNK-giver
- b.      \* nosokomi-o-emo-o-dotis<sup>2</sup>  
              hospital- LNK -blood - LNK-giver  
              ‘a hospital-blood-donor’

For a purely morphological analysis of compound formation this is problematic, as in most cases N-N compound formation is freely recursive, at least in languages like English. In the following section recent proposals in syntactic theory will be implemented in order to explain these facts.

### 3.      A phase-based syntactic analysis of synthetic compounds

In recent years, work within the minimalist program (Chomsky 1994 and later work) has assumed that syntactic derivations proceed in steps that interface with the phonological and interpretive components (‘phases’ in Chomsky 2001; 2005). In its initial implementation Phase Theory assumed that the only possible phases are CPs and *v*Ps (and possibly DPs although this is left open in Chomsky 2001). Marantz (2001) extends the phase inventory to include every case where category-changing morphology attaches to a syntactic structure, changing the extended projection. Thus, every time a nominalizer attaches to a verbal string, it defines a new phase. In other work the term phase is applied to each domain of the verb in which an argument is added (Sportiche 1999, 2005; Hallman 1997; Barss & Carnie 2003). Even though the notion of ‘phase’ and what are the possible domains in which it applies is not completely clear yet, ‘phase theory’ has helped explain a number of problems within the syntactic component and has been applied to the analysis of morphological derivations with similar success, most notably in the work of Marantz (2001, 2006).

Furthermore, Sportiche (1999, 2005) puts forward a novel account of how verbal arguments acquire referential properties. The account is based on the assumption that selection is ‘strictly’ local. This means that predicates must select for bare NPs and that subsequent nominal layers (case, number, quantification) project outside the thematic domain and trigger movement of the argument NP to VP-external positions. Thus a VP-internal argument is selected by the verb as an NP. It subsequently raises to number, case and D projections outside the VP shell. The evidence that Sportiche (2005) provides for such a claim is drawn from reconstruction effects. Consider for example the following:

- (15)      In 1986, no integer had been proved to falsify Fermat’s theorem

Under current assumptions the underlying structure for (15) would be something like (16):

- (16)      In 1986, had been proved [no integer falsify Fermat’s theorem]

---

<sup>2</sup>Ralli (2003) mentions the attested form ‘agroto-danio-dhotisi’ (farmer-loan-giving) but I think that the agentive ‘*agrotodaniiodotis*’ is not attested and this type of synthetic compounds are extremely rare – in most cases (as in (14b)) the form is ungrammatical. ‘agroto-danio-dhotisi’ may be based on a back formation of *daniodhoto*, the latter being reinterpreted as a simplex verb.

This structure should give rise to two different interpretations (depending on the scope of the determiner *no* with respect to the main predicate):

- (17) a. In 1986, no integer *x*, it had been proved that *x* falsifies Fermat's theorem  
 b. In 1986, it had been proved that no integer falsifies Fermat's theorem

However, the second interpretation is not possible, which means that the quantifier does not reconstruct in its base position. Assuming that there is always reconstruction when there is a movement operation, Sportiche (2005) concludes that (16) is not an accurate underlying representation for (15), and it should change to (18):

- (18) No .....prove ...[EMBEDDED CLAUSE integer falsify...]

Thus surface structure is derived by movement of the NP *integer* to the projection that hosts the quantifier in order to be quantified. Since this is not DP movement, reconstruction is not possible and the paradox is explained straightforwardly.

What does this predict with respect to the lexical integrity effects which were discussed in the previous section? If the analysis is on the right track then D-elements merge outside the verbal domain of 'first merge' between the verb and an un-quantified nominal head. I propose that the domain of synthetic-compound formation (and possibly most 'derivational' processes) is exactly this 'first merge' domain (see for example 'first phase syntax' Ramchand 2003; see also Sportiche 1999 for a discussion of synthetic compounds). In other words, the domain of synthetic compound formation includes the lower VP but no additional verbal functional layers.

Nominalizers (e.g. the agentive *-er* in English or *-tis* in Greek) merge at different levels in the syntactic spine, changing the categorial status of the projection from verbal to nominal (see Alexiadou 2001; Nteli-theos 2006 for different approaches on how this is achieved). In the case of synthetic compounds the nominalizer merges directly above the *v*P (the projection where the external argument is licensed). Let us see how the proposed derivation is implemented.

Consider the formation of *thiriodamastis*/ 'beast-tamer'. The verb enters the derivation selecting its internal (THEME) argument in the lower VP [VP *damas-* *thir-*]. The nominal has not been quantified in that number and definiteness projections are not available inside this lower domain. Thus this is a case of two heads merging. In terms of tree-relations this defines a symmetrical structure where each head c-commands the other since there is no functional material to create antisymmetry. Following Kayne's Linear Correspondence Axiom (Kayne 1994), this predicts that the resulting structure cannot be linearized. Therefore, the internal argument is forced to move to a higher functional projection in order to be linearized – thus the movement of the internal argument in a synthetic compound to a pre-verbal position is triggered by phonological requirements (see Barrie (2006) for a detailed discussion of these issues).

Thus, the internal argument inverts over the verb to some licensing projection (presumably headed by the linker *-o-* (see Ralli 1992 for the status of the linking morpheme in Greek compounds):



- (19) [VP thiri-o- [damas- \_\_\_\_]].

At this point the nominalizer merges changing the projection from verbal to nominal. The nominalizer in the case of the agentive nominal *thiriodamastis*/ ‘beast-tamer’ expresses the external theta-role of the verb. This is because the nominalizer merges in the projection where the external argument is licensed (i.e. the *v*P projection). This means that *v*P and NMLP are the same projections at this level in the derivation, defining a single phasal domain (for detailed discussion of this process see Ntelitheos 2006):

- (20) [nP tis [VP thiri-o- [damas- \_\_\_\_]]]

Movement of the VP to spec-nP is also phonologically driven (the nominalizer being a bound morpheme, and thus entering the derivation with the relevant morpho-phonological selection information). The important point to be noticed here is that the nominalizer has changed the properties of the projection from verbal to nominal and the distribution of the final string is that of an NP. What predictions does this derivation make?

This being a syntactic derivation, the fact that the leftmost element in the compound is interpreted as an argument of the verbal base of the compound is a direct consequence of the derivation: *thiri-* is the internal argument of the verb and is selected directly by it within the VP. This predicts the exclusion of additional internal arguments in the structure:

- (21) \* O Giannis ine thiriodamastis liontarion.  
The Giannis is beast-tamer lions.GEN  
‘Giannis is a beast-tamer of lions’

More importantly, most lexical integrity effects follow directly. As has been noted in the previous section, most lexical integrity effects are directly related to the lack of referential properties of the internal (“sublexical”) argument. But if Sportiche’s (1999, 2005) proposal is on the right track and the internal argument must enter the derivation without any referential properties then these effects disappear.

Let us revisit the relevant examples. We have seen that synthetic compounds seem to be anaphoric islands ((8), repeated here as (22)):

- (22) \* O Giannis ine plirofori<sup>1</sup>-o-dotis alla den tin<sub>1</sub> gnorizo.  
The Giannis is information- LNK-giver but not it know.1SG  
‘Giannis is an informer but I don’t know it (the information).’

It is well known in the relevant literature that only DPs (i.e. “quanticized” nominals) can be referential, NPs never can (cf. Stowell 1991; Longobardi 1994). If the nominal inside the synthetic compound has not been quanticized (due to lack of the appropriate projections) then its inability to bind/co-refer with a compound-external referential expression follows straightforwardly.

The fact that the internal argument cannot be quanticized is further supported by its inability to host determiners or proper names and pronoun internal arguments ((13) repeated here as (23)):

- (23) a. \* [o kapno-] kalierjia

[D tobacco-] cultivation

- b. \* Giorgio-thavmastis  
Giorgio-admirer

There are certain derivational processes, including compounding, that seem to allow for some proper names to be included in the resulting forms (see Lees 1960; Lieber 1983; Di Sciullo and Williams 1987; and others):

- (24) a. a Nixon admirer  
b. the Euler number  
c. my computer is an IBM machine

However, as Di Sciullo and Williams (1987) show, this is not possible with any proper name:

- (25) \* a Bill admirer

It is therefore possible that the proper name inside the compound is not fully referential – it does not refer to the respective entity in the real world but rather to a specific property of that entity (e.g. in the case of ‘Nixon admirer’ to the policies of Richard Nixon).

Moving to other types of integrity effects, we have noted that no movement is allowed out of the compound ((7), repeated here as (26)):

- (26) a. O Giannis ine plirofori-o-dotis.  
The Giannis is information- LNK-giver  
‘Giannis is an informer.’  
  
b. \* Ti ine o Giannis dotis?  
What is the Giannis giver?  
‘What is Giannis giver (of)?’

In Bresnan and Mchombo (1995) this inability is taken as one of the main lexical integrity effects. However, the problem here is that there are numerous, clearly syntactic structures that disallow such movement. The effect observed in (26) is a case of the well-known, left-branch extraction (Ross 1967) that the following example also exhibits:

- (27) a. I like green apples.  
b. \* Which do you like apples?

Other types of extraction can also be accounted for because of the phasal status of the synthetic compound. The Phase-Impenetrability Condition (Chomsky 2001, 2005) assumes that at each step of the derivation, the only elements available to subsequent syntactic operations are the head and specifier of the phase. This means that for elements inside the compound the rightmost element is not available for subsequent computations as it resides below the phase head (i.e. the nominalizer). Thus, neither the leftmost nor the rightmost elements of the compound can move for independent reasons and thus the effect disappears.

The unavailability of coordination and gapping (examples (9) and (12b), repeated here as (28a -28b)) follows straightforwardly from well-known restrictions on what can be

coordinated or gapped. In particular, it has been noted in various sources (c.f. van Oirsouw 1987 and references therein, also Johnson 2001) that only constituents can coordinate or gap. However, in most cases of synthetic compound formation, under current assumptions, what is attempted to be coordinated or gapped is not a constituent at all.

- (28) a. \* O Giannis ine plirofori-o (ke) em-o-dotis.  
 The Giannis is information- LNK- (and) blood- LNK- giver  
 ‘Giannis is an information-(and)-blood-giver.’
- b. \* O Giannis ine ele-o-kaliergitis (ke) i Maria sitir-o- ~~kaliergitis~~.  
 The Giannis is olive- LNK-grower (and) the Mary wheet- LNK- ~~grower~~.  
 ‘Giannis is an olive-grower and Mary a wheet-grower.’

In the case of coordination in (28a) what is coordinated is the functional head/linker *-o* plus material in its specifier. One could test coordination of just the nominal specifiers without the linking morpheme, but that would violate the fact that *-o* is an obligatory marker of syntactic relations at this level – i.e. coordination, modification, or predicate argument licensing is not possible without the functional morpheme, for linearization reasons. Thus, coordination of the nominal arguments cannot succeed. A similar reasoning applies to the exclusion of gapping in these cases. It is possible for gapping structures to appear as eliding non-constituents but this has been attributed to prior movement of material to left-peripheral projections (c.f. Jayaseelan 1990; Johnson 2001). However, the phase defined by the nominalizer is reduced in that it does not contain left peripheral projections (see Ntelitheos 2006, Ch. 5, for a discussion of the lack of left-peripheral positions in nominalised strings). Thus no gapping of any sort can be possible in the structure involved in synthetic compounds.

A final issue has to do with the inability of further arguments to appear within synthetic compounds when the internal argument is present ((14b) repeated here as (29b)):

- (29) a. emo-o-dotis  
 blood- LNK-giver
- b.\* nosokomi-o-emo-o-dotis  
 hospital- LNK -blood - LNK-giver  
 ‘a hospital-blood-donor’

As was emphasized earlier, each verbal argument enters the derivation as a bare NP without any further functional material. This material enters the derivation in subsequent functional projections that define their own phasal domain (i.e. merger of each argument corresponds to a new phase (see Sportiche 1999, 2005; Hallman 1997 and Barss & Carnie 2003 for a discussion on why this follows from minimalist assumptions). But if this is on the right track then only one argument can appear as a bare nominal adjacent to the root base. Adjunction of a second argument is prohibited by the fact that either the functional material of the first argument or the functional projections introduced by the nominalizer would intervene between the two bare arguments. Thus, the only possibility for further arguments to materialize is as full DPs/PPs, as in the case of normal noun or prepositional phrases:

- (30) emo-o-dotis se nosokomia  
 blood- LNK-giver to hospitals

We turn now to the examination of an additional type of structures that seems to have some of the properties of synthetic compounds but have predominately been categorized as purely syntactic in nature within the lexicalist-morphology-oriented tradition.

#### 4 Phrasal Compounds

Greek exhibits a type of phrasal structures that seem to have a limited number of compound-like properties. In these structures the internal argument appears following the deverbal nominal in genitive case (Anastasiasi-Symeonidi 1986; Horrocks & Stavrou 1989; Ralli 1992, 2003, 2007):

- (31) a.    kalliergitis    kapnou  
              grower        tobacco.sg.gen  
              ‘grower of tobacco’
- b.    damastis        thirion  
              tamer         beast.PL.GEN  
              ‘beast tamer’

Phrasal compounds exhibit mixed properties: they share with synthetic compounds the properties of not allowing D-elements (32a- 32b) and occasional non-compositional semantics (33):

- (32) a.    \* kalliergitis tou kapnou  
              grower    D    tobacco.SG.GEN  
              ‘grower (of) the tobacco’
- b.    \* damastis        ton        thirion  
              tamer            D        beast.PL.GEN  
              ‘tamer (of) the beasts’
- (33)        epexergastis dhedhomenon  
              processor    data.PL.GEN  
              ‘processor (of) data (*can only be a computer*)’

However, contrary to synthetic compounds, phrasal compounds of this sort allow for recursion (34)), exhibit default syntactic word order (31)-(34), the internal argument appears with morphological genitive case as well as number and gender morphology (31)-(34), and the structure defines two distinct prosodic domains for stress assignment.

- (34) a.    epexergastis neon dhedhomenon  
              processor    new    data.PL.GEN  
              ‘processor (of) new data’
- b.    ekpobi    dhilitiriodhon        aerion  
              release    poisonous.PL.GEN    gas.PL.GEN  
              ‘release (of) poisonous gasses’

I propose that these properties can be explained if the nominalizer attaches higher up in the verbal structure, to include all referential projections but crucially not case:

(35) [**damas2-tis3** [CASEP **thirion1** [NML t2 [t3 [DP t1 [ NUMP t1 [ *vP* [*VP* t2 t1]]]]]]]]]

The nominal argument raises to NumP and DP to check its number and definiteness features. The *vP* subsequently moves to spec-NMLP where the structure's properties change from verbal to nominal (see Ntelitheos 2006 for details on the mechanisms involved in nominalization processes of this type). The case projection thus merges outside the verbal domain and within the nominal extended projection and therefore is morphologically realized as genitive case. The verb-nominalizer complex moves to Spec-CaseP - the mechanics of the movement operation parallel those of *of*-insertion in similar structures in English. Greek genitive case acts as a 'linker' in the sense of DenDikken (2006), i.e. links the predicate to its argument through predicate-inversion (see Kayne 2000 for a similar structure for prepositional complementizers).

The proposal explains the word order without further stipulations. Final word order is achieved following the same mechanisms that apply to more straightforward nominal structures, bringing the predicate to the left of its argument through a linking morpheme. The final structure contains some of the functional material that is available in the verbal domain, namely NUMP and DP, and thus exhibits more "syntactic" properties in that it provides the space for adjectival modification or syntactic coordination (i.e. recursive structures of any sort).

The fact that the nominal has moved to a number projection before nominalization has taken place, with subsequent movement to the projections headed by the nominalizer and case, explains the presence of inflectional morphology on the noun. The internal argument appears as a generic plural as it does in the corresponding verbal structure:

(36) a.      kallierg-o      kapn-a  
              grow-1SG.PRS tobacco.PL.ACC  
              'I grow tobacco'

A definite argument would imply a specific type of tobacco and would normally be followed by a relative clause further specifying the argument:

(37) a.      kallierg-o      ta kapn-a      pou hrisimopii o ASSOS  
              grow-1SG.PRS tobacco.PL.ACC that USE.3SG.PRS the ASSOS  
              'I grow the tobacco that ASSOS uses (in its cigarettes)'

Therefore, the existence of an indefinite plural in these cases is not a property of the phrasal compound but of the more general syntactico-semantic properties of the construction. Thus, in general there is nothing that distinguishes these strings from common noun phrases and this also explains the stress-patterns observed in these strings. The nominalizer dominates two phases: the DP domain of the nominal argument and the higher nominalised domain of the full structure (the *nP*). Marantz (2001, 2005) proposes that category changing morphology defines a Spell-out domain (a phase). Marvin (2003) examines the prosodic properties of words and puts forward the idea that each phase (inside a word) defines its own prosodic domain for purposes of stress assignment.

Following these proposals stress in Greek synthetic and phrasal compounds is predicted to differ. Synthetic compounds define a unique phase (the nominalizer merges immediately above the root domain and there is no DP argument). This predicts that synthetic compounds will have a single primary stress:

- (38) a.      kapnokaliérjia  
              'tobacco cultivation'
- b.      thiriodamastís  
              'beast tamer'

Phrasal compounds on the other hand define two phases: the phase of the argument (DP), and the higher phase dominated by the nominalizer *n* (*nP*). Thus, two prosodic domains are defined and this predicts two primary stresses:

- (39) a.      kaliérjia          kapnoú  
              cultivation      tobacco.GEN
- b.      damastís        thiríon  
              tamer            beast.PL.GEN

The significance of 'phrasal' compounds of this sort lies in their strong tendency to appear with non-transparent, idiomatic meanings. Ralli (2007:245-246) accepts that this is problematic for her proposal of a distinct morphological component of the grammar and seems to indicate rather a grammatical continuum where synthetic compounds reside at the lowest, strictly morphological level, followed by "loose" compounds (in her terminology) that have a limited number of syntactic properties, followed by semantically non-transparent "phrasal" compounds, with syntactic phrases occupying the other extreme). The problem with such an approach is that it aims to maintain two distinct grammatical components for what seems to be a clear continuum of structures that exhibit gradient syntactic properties. But the argument for two distinct components is weakened because the boundaries between these components become fuzzy. Capturing these gradient properties within a single syntactic component, as the present proposal attempts to do, clearly results in a simpler and therefore preferable theory of morphosyntactic structure.

The problem for proponents of the lexicalist approach is of course a lot more serious. One of the strongest arguments against syntactic analysis of previously thought to be morphological processes was the semantic opacity of the derived strings (see Chomsky 1970 and most work in the "lexicalist tradition" that followed). However, if there are clearly syntactic strings that exhibit semantic opacity to a great extent, then at least this argument against syntactic analysis of derivational processes such as compounding disappear.

On the other hand, the analysis proposed here does not face a problem. The continuum that Ralli (2007) acknowledges is encoded in the proposed structures. Synthetic compounds are formed at the 'first merge' domain, the most basic available domain that contains no functional projections of any sort. This is the domain of lexical idiosyncrasies, which Marantz (2005) calls the domain of "inner word formation" as contrasting with the domain of "outer word formation", the contrast defined by the cyclic interpretation imposed by phases, given that each category-determining head defines a phase. The domain of synthetic compound formation or root domain is closed by the first category-determining head and

constitutes what Ramchand (2006) calls the “first-phase syntax”, or Hale & Keyser’s “l-syntax” (2002).

Thus, the inner domain is the domain of potential special forms and idiosyncratic meaning and syntactically corresponds to the domain that attaches inside morphology determining the lexical category of the resulting structure. The idiosyncratic properties of synthetic compounds and “phrasal” compounds follow straightforwardly: in both cases the first category defining morpheme (the nominalizer) attaches above the domain where compositional meaning is established. In the derivation of normal syntactic phrases on the other hand, all additional functional projections merge on top of different category assigning morphemes for the verb (*vP*) and the nominal argument (*nP*) and thus the meaning is predictable from the structure.

## Conclusion

I have argued for a purely syntactic approach to deverbal compound formation in Greek. Following independently motivated syntactic principles such as the ‘split-DP hypothesis, in Sportiche 1999-2005, and the implementation of phase theory (Chomsky 2001, 2005) in the morphological domain (Marantz 2001, 2006), I proposed that lexical integrity effects can be captured if we assume that the domain of synthetic compound formation is the root level, before categorical morphology applies. In this level the base merges with a bare nominal argument which is quantized (i.e. acquires referential properties) in a series of projections outside the ‘first-merge’ domain. The fact that the argument is not referential explains a number of lexical-integrity effects while some of the remaining effects are captured by the phasal status of the projection (i.e. the base or the argument cannot be displaced).

The differences in the distribution of different compound forms, including synthetic and phrasal compounds, were shown to be derived from properties of the lower verbal domain and the idiosyncratic selectional properties of nominalizing affixes. The proposal also explains the distribution of stress assignment in these compounds as a result of matching phases with prosodic/stress domains.

A number of issues clearly remain unexplained. Some of the integrity effects remain to be accounted for, as are issues of productivity and the exact properties of semantic opacity. However, the issues tackled here seem to indicate that despite the accumulated evidence against a purely syntactic analysis of morphological derivational processes such as compounding, such an analysis is still tenable (and still preferable) under current assumptions in syntactic theory.

## References

- Ackema P. & A. Neeleman. 2005. *Beyond Morphology. Interface Conditions on Word Formation*. Oxford: Oxford University Press.
- Alexiadou, A. 2001. *Functional Structure in Nominals: Nominalization, and Ergativity*. Amsterdam: John Benjamins.
- Anastasiadi-Symeonidi, A. 1983. La Composition en Grec Moderne d’un Point de Vue Diachronique. *Lalies* 2:77-90.
- Anderson, S. R. 1992. *A-morphous Morphology*. Cambridge: Cambridge University Press.

- Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Barrie, M.G.M. 2006. Dynamic Antisymmetry and the Syntax of Noun Incorporation. PhD Dissertation, University of Toronto.
- Barss, A. and A. Carnie. 2003. Phases and Nominal Interpretation. Ms. University of Arizona.
- Booij, G.E. (1988). The relation between inheritance and argument linking. In Everaert et al. (eds) *Morphology & Modularity*, Foris, Dordrecht, 57-75.
- Booij, G. E. 1992. Compounding in Dutch. *Rivista di Linguistica* 4/1:37-61.
- Booij, G. E. 2005. *The grammar of words. An introduction to linguistic morphology*. Oxford: Oxford University Press.
- Bresnan, J. & S. Mchombo. 1995. The lexical integrity principle: evidence from Bantu. *Natural Language and Linguistic Theory* 13, 181-254.
- Chomsky, N.. 1970. Remarks on Nominalizations, in R. Jacobs and P. Rosenbaum, P. (eds.) *Readings in English Transformational Grammar*. Waltham, MA: Blaisdell, 184-221.
- Chomsky, N.. 2001. Derivation by Phase. In M. Kenstowicz (ed.), *Ken Hale: a life in language*. Cambridge, MA: MIT Press, 1-52.
- Chomsky, N.. 2005. On Phases. Ms. MIT.
- den Dikken, M. 2006. *Relators and Linkers. The Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, MA: MIT Press.
- Di Sciullo, A. M. & E. Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.
- Di Sciullo, A. M. 2005. *Asymmetry in Morphology*. Cambridge, MA: MIT Press.
- Di Sciullo, A. M. & A. Ralli. 1999. Theta-role saturation in Greek deverbal compounds. In A. Alexiadou et al. (eds.) *Issues of Greek generative syntax*. Dordrecht: Kluwer.
- Fabb, N. 1984. Syntactic Affixation. PhD Dissertation.
- Grimshaw, J. 1990. *Argument Structure*. Cambridge MA: MIT Press.
- Hale, K. & J. S. Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Hale, K. & S. J. Keyser (eds.) *The view from Building 20*, Cambridge, MIT Press.
- Hallman, P. 1997. Reiterative Syntax. In J. R. Black & V. Motapanyane (eds.) *Clitics, Pronouns and Movement*, Amsterdam: John Benjamins, 87-131.
- Horrocks, G. & M. Stavrou. 1989. Clitics and Demonstratives in Noun Phrases. *Studies in Greek Linguistics*, Thessaloniki: Kyriakidis.
- Jayaseelan, K.A. (1990) Incomplete VP deletion and gapping, *Linguistic Analysis* 20, 64-81.
- Johnson, K. 2001. What VP ellipsis can do, what it can't, but not why, in Mark Baltin and Chris Collins (eds.) *The handbook of contemporary syntactic theory*, Blackwell Publishers, pp. 439-479.
- Kakouriotes, A., M. Papastathi & A. Tsangalides. 1997. Incorporation in Modern Greek: Lexical or Syntactic? In G. Drachman, A. Malikouti-Drachman, C. Klidi & J. Fykias (eds.) *Proceedings of the Second International Conference of Greek Linguistics*. Graz: Neubauer Verlag, 78-86.
- Kayne, R. 1994. *The Antisymmetry of Syntax*, Cambridge, MA: MIT Press.
- Kayne, R. 2000. *Parameters and Universals*, Oxford University Press.
- Lapointe, S. 1980. A theory of grammatical agreement. Doctoral dissertation, University of Massachusetts, Amherst, Mass.
- Lieber, R. 1983. Argument Linking and Compounding in English. *Linguistic Inquiry* 14: 251-86.
- Longobardi, G. 1994. Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form. *Linguistic Inquiry* 25: 609-665.



- Marantz, A. 2001. Words. Ms., MIT.
- Marantz, A. 2006. Phases and Words. Ms., NYU.
- Marvin, T. 2003. Topics in the Stress and Syntax of Words. PhD Dissertation. MIT.
- McNay, A. 2006. Information structure at the phase level. In Lisa Mackie and Anna McNay (eds.), *Online Proceedings of LingO - the first University of Oxford postgraduate conference in linguistics*.
- Ntelitheos, D. 2006. The Morphosyntax of Nominalizations: A Case Study. PhD Dissertation. UCLA.
- van Oirsouw, R. R. 1987. *The Syntax of Coordination*. Groom-Helm.
- Pesetsky, D. 1985. Morphology and Logical Form, *Linguistic Inquiry* 16, 193-246.
- Postal, P. 1969. Anaphoric Islands. In R. Binnick, A. Davison, G. Green, and J. Morgan (eds.) *Proceedings of the Fifth Regional Meeting of the Chicago Linguistic Society*, Chicago: CLS, 205-219.
- Ralli, A. 1989. Ta Rimatika Sintheta tis Neas Ellinikis. *Meletes gia tin Elliniki Glossa* 1988: 205-221.
- Ralli, A. 1992. Compounds in Modern Greek. *Rivista di Linguistica* 4:1, 143-174.
- Ralli, A. 1999. To phenomeno tis synthesis sti Nea Elliniki: Perigrafi ke Analis. *Parousia IA-IB*:183-185.
- Ralli, A. 2003. Morphology in Greek Linguistics: The State of the Art. *Journal of Greek Linguistics* 4: 77-129.
- Ralli, A. 2007. *I Sinthesi Lexion*. Athens, Greece: Patakis.
- Ralli A. and M. Stavrou. 1998. Morphology-Syntax Interface: A-N Compounds vs. A-N Constructs in Modern Greek. In G. Booij & J. Van Marle (eds.) *Yearbook of Morphology 1997*: 243-264.
- Ramchand, G. 2003. First Phase Syntax. Oxford University ms.
- Rivero, M. L. 1992. Adverb Incorporation and the Syntax of Adverbs in Modern Greek. *Linguistics and Philosophy* 15: 289-331.
- Roeper, T. 1987. Implicit arguments and the head-complement relation, *Linguistic Inquiry* 18, 267-310.
- Roeper, T. 1988. Compound syntax and head movement, *Yearbook of morphology* 1, 187-228.
- Roeper, T. and D. Siegel. 1978. A Lexical Transformation for Verbal Compounds, *Linguistic Inquiry* 9, 199-260.
- Selkirk, E. 1982. *The Syntax of Words*. Cambridge, MA: MIT Press.
- Sminiortopoulos, J. and B. D. Joseph. 1997. On the so-called Adverb Incorporation in Modern Greek. In G. Drachman, A. Malikouti-Drachman, C. Klidi & J. Fykias (eds.) *Proceedings of the Second International Conference of Greek Linguistics*. Graz: Neubauer Verlag, 117-128.
- Smirniotopoulos, J. And B. D. Joseph. 1998. Syntax vs. The Lexicon: Incorporation and Compounding in Modern Greek. *Journal of Linguistics* 34: 447-488.
- Spencer, A. 2005. *Morphological Theory*. Oxford: Blackwell.
- Sportiche, D. 1999. Reconstruction, Constituency, and Morphology. GLOW Newsletter 42: 56-57.
- Sportiche, D. 2005. Division of Labor between Merge and Move: Strict Locality of Selection and Apparent Reconstruction Paradoxes. LingBuzz,
- Sproat, R. 1985. On Deriving the Lexicon. PhD Dissertation, MIT.
- Stowell, Tim. 1991. Determiners in NP and DP. In: Katherine Leffel and Denis Bouchard (eds.), *Views on Phrase Structure*. Amsterdam: Kluwer Academic Press, 37-56.
- Travis, L.. 1994. Event phrase and a theory of functional categories. P. Koskinen (ed.) *Proceedings of the 1994 annual conference of the Canadian Linguistic Society*, 559-570.

Williams, E. 1981. Argument Structure and Morphology, *The Linguistic Review* 1, 81-114.