

WPFG

Web Papers in Functional Grammar

wpfg no 80
February 2007

English 'NOUN+NOUN' sequences: their place in Functional Discourse Grammar
Carmen Portero Muñoz
Universidad de Córdoba, Spain

English ‘NOUN+NOUN’sequences: their place in Functional Discourse Grammar

0. Introduction¹

The choice of the topic of ‘noun + noun’ compounds for my paper has been motivated by the realization that there are relevant aspects to which the FG conception of morphology pays very little attention. One of these is compounding. To my knowledge, with the exception of the account on noun incorporation (Dik 1980, 1997), not much is said on the nature of compounds, i.e. the types of processes involved in the generation and interpretation of compounds and the formal mechanisms to account for them, as well as of the place of compounding in grammar or its interactions with the different components. Unfortunately, the attention paid to compounding processes is even more scarce in the new version of the model (FDG) (Hengeveld 2004a, 2004b).

The focus of the present paper is certain types of compounds, specifically so-called *primary* ‘noun + noun’ compounds like *university man* (Selkirk 1981, 1982). In contrast with *synthetic* compounds (like *grammar teacher*), where an element satisfies an argument of the deverbal head constituent, the non-heads of primary ‘noun + noun’ compounds add something other than a thematic specification to the head. These compounds raise a number of questions, such as the place they occupy in the grammar. To this, we must add the problems related to their interpretation.

I will mention only some of these problems and I will focus on one main question: what sort of mechanisms are there to account for the creation and interpretation of these compounds within the new conception of FDG? In answering this question, the starting point will be discourse, which is seen as participating in the choices made at each level of generation of linguistic structures. In my view, this influence should be extended to apply to the predicate formation component in a more explicit way, since contextual factors may be relevant for a given word formation process, in this case compound creation, to take place.

This paper will be organized as follows. In the first section I will summarize common assumptions about noun + noun compounds, concerning both their connection with

¹ A preliminary version of this paper was delivered at the 11th International Conference on Functional Grammar (Gijón, September 2004). Thanks are due to Daniel García Velasco and the editor Evelien Keizer for helpful comments and suggestions.

syntax and the problems with their interpretation. Afterwards I will offer an overview of the treatment of compounding in FG. The last section constitutes the main part of this paper; here I will suggest a possible way to account for noun + noun compounds in a way that is more in accordance with the new version of the model, i.e. FDG.

1. Common assumptions about compounds

1.1 The syntax of ‘Noun + Noun’ compounds

Compounding processes form a central topic when dealing with the syntax-morphology interface. ‘Noun + noun’ compounds are even more in between syntax and morphology than other types, since sequences of ‘noun + noun’ can be treated either as compounds or as noun phrases with modifiers which happen to be nouns. There is, in fact, considerable disagreement regarding the treatment of ‘noun + noun’ combinations and many consider as compounds collocations that others regard as phrases (see Carstairs-McCarthy 2002: 69). A reason for this seems to be that some linguists are reluctant to analyse nouns as modifiers within a phrase. Therefore, they don’t distinguish structurally between a compound like \cap *toy factory* (“a factory that makes toys”), and the phrase *toy* \cap *factory* (“a factory that is a toy”). The parallel between ‘noun + noun’ combinations like *landing craft* or *surface craft* and *hovercraft*, generally recognized as a compound, is such as to regard the former as compounds as well. On the other hand, one of the properties characterizing compounds is recursion (exemplified in (1)), which points toward their syntactic status.

- (1) *film society*
student film society
student film society committee
student film society committee scandal
student film society committee scandal inquiry

Several of the syntactic criteria which are frequently used to distinguish compounds from free combinations, such as

- a. that *Adjectival modification of the first member cannot occur*
- b. that *Compounds do not permit co-ordination*
- c. that they are *Referentially opaque*
- d. that *Replacement of the head noun by one is not possible*

have been shown to fail, as the following examples illustrate:

- (2) *public lending right* (as a counterexample to a.)
- (3) *cat and dog shows* (as a counterexample to b.)
- (4) So, I hear you're a real *cat-lover*. *How many* do you have now?
(as a counterexample to c.)
- (5) He wanted a riding *horse*, as neither of the carriage *ones* would suffice
(Bauer 1998: 77) (as a counterexample to d.)

In view of these facts, I agree with Bauer (1998) that the criteria which are usually assumed to distinguish between these compounds and phrases do not allow us to draw a clear and consistent distinction between a morphological and a syntactic construction. The process by which a noun phrase consisting of a head noun and a premodifying noun is formed is probably similar to the process of 'noun + noun' compound formation, although 'noun + noun' compounds must be subject to more strict semantic restrictions in order to count as lexemes and become part of the lexicon.

1.2 The interpretation of 'noun + noun' compounds

Another problem linguists have been faced with when dealing with these compounds is related to their interpretation. There is agreement among scholars about the large variation in the interpretation of noun-noun compounds. One of the factors on which the meaning of these compounds might be taken to depend is the argument structure of the head. A second factor is the semantics of the nouns and their semantic relationships.

1.2.1. *The argument structure of the head*

The argument structure of the head is relevant to the interpretation of ‘noun + noun’ compounds. However, the argument structure of the head is useless in the case of primary ‘noun + noun’ compounds. In contrast with synthetic compounds, which display a rather specific range of semantic interpretations, for ‘noun + noun’ primary compounds it seems that virtually any relation between head and non-head is possible (see Selkirk 1981: 249). For example, a compound like *water-mill* (Allen 1978: 92) is open to multiple interpretations, so that in appropriate contexts it could have any of the meanings in (6):

- (6) “mill powered by water”
- “mill which produces water”
- “mill located near the water”
- “mill for analysing the content of water”
- “mill where the employees drink water, etc”

1.2.2 *The semantics of the nouns and their semantic relationships*

A different factor on which the meaning of these compounds might be taken to depend is the semantics of the nouns and the semantic relation between them. A classification in terms of semantic categories may cast some light on the generation and interpretation of new ‘noun + noun’ compounds. Different scholars have proposed taxonomic lists of types of semantic relations between compound members (Jespersen 1942: 142-145, Marchand 1969: 45-52, Adams 1973, Downing 1977, Levi 1978, Warren 1978, Adams 2001: 82-86, Plag 2003: 148-152).

For example, Downing (1977: 828) suggests that any inventory of the most common underlying relationships should include at least the list in (7), which is not, however, exhaustive:²

² Other scholars have provided similar semantic classifications: Marchand (1969: 45-52), for example, establishes different groups on a syntactic rather than semantic basis, although each set covers different semantic groups: (i) The Subject type: e.g. *bulldog* (‘B is like A’), *sugar loaf* (‘B consists, is made up of A’), *arm chair* (‘B has, possesses, contains A’), etc; (ii) The Object type: Affected object, e.g. *steamboat* (‘A operates B’), etc, Effected object, e.g. *beet sugar* (‘A produces B’), etc; (iii) The Adverbial Complement type, e.g. *corn belt* (‘A grows in b’), etc. Adams (1973: 61) offers a taxonomic list of types of semantic relations obtaining between compound members, like Appositional, Associative, Instrumental, Locative, Resemblance, Composition/Form, Contents, Adjective-Noun, Names, and Other. Levi (1978) identifies nine relationships which any ‘noun + noun’

- (7) Whole-part (*duck foot*)
 Half-half (*giraffe-cow*)
 Part-whole (*pendulum clock*)
 Composition (*stone furniture*)
 Comparison (*pumpkin bus*)
 Time (*summer dust*)
 Place (*Eastern Oregon meal*)
 Source (*vulture shit*)
 Product (*honey glands*)
 User (*flee wheel barrow*)
 Purpose (*hedge hatchet*)
 Occupation (*coffee man*)

Nevertheless, this attempt at classification relies on rather loosely defined distinctions. For example, the concept ‘source’ may have different meanings depending on the entities involved.

Several scholars (see Downing 1977, Levi 1978) agree that the fact that a given compound expresses one of the typical relationships is not sufficient to guarantee the right interpretation of a compound. Instead, Downing suggests that there are certain pragmatic limits which play a role in the creation and interpretation of ‘noun + noun’ compounds. For her, on most occasions, compounds are easy to create and understand because they refer to entities ‘in culturally relevant ways’ and in spite of the variety of possible semantic relationships between the two nouns that can be expressed, some of them will be favoured because they reflect the way speakers refer to entities. For Levi, certain relations seen in compounds are privileged.

compound can potentially embody: Cause (*tear gas*), Have (*picture book*), Make (*honey bee*), Use (*steam iron*), Be (*soldier ant*), In (*field mouse*), For (*horse doctor*), From (*olive oil*), About (*tax law*). Adams (2001: 83) draws up a number of different groups on the basis of the different semantic specifications added to the referent of the head by the modifier. They coincide to a large extent with those established by Levi (1978). For example, he distinguishes the following sets: Function (*traffic lights*), Location (*amusement park*), Cause (*horror film*), Origin/ Source (*mountain ash*), Possessor (*student loan*), Resemblance (*bell jar*), What something is about (*tax law*). To these Adams adds further sets: Content (*cheque book*), Composition (*butter mountain*), Material (*ivory tower*), How something functions (*computer game*).

They are ‘recoverable’ and need not be expressed. Although ambiguity is to be expected *context* dictates which interpretation is more appropriate.

2. Compounds in FG

The formation of compounds is briefly handled in FG as one of the processes involving valency reduction, namely, *incorporation* (Dik 1980; 1997: 10). A particularly frequent type is Goal incorporation, illustrated in (8) (in what Dik 1997 calls ‘quasi-English’):

(8) *John bird-catches*

This type of construction has a number of recurrent properties:

- a. the output predicate is typically intransitive.
- b. the incorporated nominal cannot be modified or determined in any way, and is non-referential.
- c. The resulting complex verb often gets a “de-actualized” (generic, habitual, virtual, dispositional) interpretation.

The relevant predicate formation rule for Goal incorporation is then formulated as shown in (9) (Dik 1997: 11):

(9) Goal incorporation
input: (a) pred1 [V] (x₁) (x₂:<sel>)_{Go}
(b) pred2 [N], sel(x₂)
output: {Inc pred1-pred2} [V] (x₁)

In this rule, the symbol “sel” indicates that the incorporated nominal predicate should match the selection restriction imposed on the second argument position of the verbal input predicate.

Dik (1997: 11) points out that across languages, nominal predicates corresponding to different semantic functions may be incorporated into the verbal predicate, not only semantic functions typical of arguments, as in (8), where the

incorporated nominal is a Goal, but also semantic functions typical of a satellite of level 1, like (10):

(10) *John was chair-sitting,*

where the incorporated nominal is a Location (Dik 1997: 11 provides a list of different semantic functions corresponding to satellites that can be incorporated).

This poses a problem to standard FG since satellites are by definition not included within predicate frames so that this implies either an extension of the domain of predicate frames or allowing predicate formation rules to operate outside the Fund on core predications.

Since the non-head elements correspond to either the verbal argument or to a first level satellite, not all of these compounds could be considered as synthetic (if we followed Selkirk's definition strictly). However, they are like most synthetic compounds in that they can be easily interpreted. This is due to the fact that the head is a verbal or a verb-derived element, in contrast with 'noun + noun' compounds.

However, this is the only type of compounding processes handled in FG, and nothing is said about the formation of other types, that is, compounds which do not contain a deverbal element and in which the first component bears no argument relation to the head, that is, *primary* 'noun + noun' compounds. These compounds constitute a productive type not only in English but also in other languages. For example, a large number of Mandarin compounds are of the type *ga#ng-b&i* ('steel-pen': (fountain) pen), in which the first element of the compound is a modifier of the second, limiting or making more specific the reference of the element they modify (see Anderson 1985: 46).

3. 'Noun + noun' compounds in FDG

3.1 Introduction

We now get to the question of how the FDG framework might contribute to provide an account of the process of 'noun + noun' formation. Any theoretical model should reflect the following facts:

- a. Compounding, and particularly ‘noun + noun’ compounds, is a productive, though constrained, process. Therefore, in spite of the existence of some lexicalized compounds, we cannot deal with them within the lexicon, or at least not exclusively there (needed: A PREDICATE FORMATION RULE).
- b. ‘Noun + noun’ compounds bear some resemblance to syntactic constructions, though their formation is more constrained than that of syntactic constructions. This therefore raises the question whether they could be accounted for by the same mechanisms that generate/produce phrases? (needed: TERM FORMATION RULES?).
- c. The semantic relationships appropriate for compound formation cannot be characterized in terms of a finite list. However, the compounding process is not totally ungoverned and there are some basic semantic relationships that should be taken as point of departure (part of CONCEPTUAL COMPONENT).
- d. In many cases, compounds are interpretable in context. Any discussion of them should consider their role in discourse, and the way in which context contributes to their interpretation (part of CONTEXTUAL COMPONENT).

In view of the claim in (a), it appears that the creation of ‘noun + noun’ compounds should be explained by means of a predicate formation rule. This rule already exists in the FG model (see Section 2). However, (c) and (d), suggest that we should try to formulate a revised version of this predicate formation rule or think of a new way to account for these facts about the generation and interpretation of ‘noun + noun’ compounds.

An important basic property of FDG is that it systematically interacts with a conceptual, a contextual and an output component (Hengeveld 2005). In Hengeveld's words “the conceptual component is not part of the grammar but is the driving force behind the grammatical component.” (Hengeveld 2005: 57) The contextual component, on the other hand, contains a description of the discourse domain, not only of the content and form of the *preceding discourse*, but also of the *actual perceivable setting* in which the speech event takes place. In this way the contextual component provides information relevant to the form that subsequent utterances may take.

Another important difference between FDG and FG as far as predicate formation is concerned is that the fund, where (in the standard FG version) predicate

formation used to be handled, is now taken care of in separate components belonging to the various levels (interpersonal, representational, structural and acoustic) (Hengeveld 2004b). The basic idea is that for every level within the model the fund contains the set of basic units which are used to build up that level.

In the following sections I will show how the conceptual and contextual components are relevant for the formation and interpretation of ‘noun + noun’ compounds. In particular I will question the appropriateness of a predicate formation rule for this process, suggesting a revision of the standard formulation of this compounding process. In the last section I will suggest a proposal for a new treatment of these compounds that makes use of the new modular approach of FDG.

3.2 ‘Noun + noun’ compounds and the conceptual component

If we consider Dik’s proposal for noun incorporation (section 2), the possible semantic functions of the incorporated nominal (see (10)) resembles the list of typical semantic relations suggested by other authors (see (7)). On the other hand, selection restrictions could account for a more specific type of compatibility required by the two predicates. Nevertheless, it seems that some other type of information, apart from selection restrictions and semantic functions, is required for the correct creation and interpretation of many ‘noun + noun’ compounds.

In exploring this idea, we should bear in mind a number of facts:

- a. There is *no explicit predicative relation* between the two nouns.

In primary ‘noun + noun’ compounds there is no explicit predicative relation between the two nouns.³ Thus, in contrast with a synthetic compound like *taxi-driver*, where the first noun satisfies one argument position of the second noun and which can only mean “a person who drives taxis”, a compound like *water-mill* may have different meanings, as we have seen before (see (6) in 1.2.1).

³ See Marchand (1969: 18), for whom primary compounds, which he calls *non-verbal nexus compounds*, are implicit syntagmas. He claims that all compounds are explainable on the basis of syntactic relations underlying them in sentences, but whereas verbal nexus compounds (i.e. synthetic compounds) are reduced from full sentences and represent the entire verbal nexus in an overt form, primary compounds do contain the syntactic relations, but only implicitly, while the semantic element of the verb is not expressed. The semantic relations between the component element are more in evidence than the syntactic relations in the case of primary compounds.

Therefore, selection restrictions and semantic functions only are not sufficient to capture the meaning of the primary compound since there still remains the interpretation of the semantic relation between the two nouns.

Now, although selection restrictions are not sufficient, the identification of this verbal relation relies partly on the semantic features of the nouns, that is, we have to depart from these semantic features. This, however, brings us to the second fact to bear in mind:

b. There is *a range of possible semantic features*

The number of semantic features characterizing the nouns is not limited to one. This could be represented by extending the number of selection restrictions of the first input noun, as shown in (a) in (11):

- (11) ‘Noun + noun’ compound formation rule
input: (a) pred1 [N] (x_i :<sel₁>, <sel₂>, <sel₃>, etc)
(b) pred2 [N], sel \mathbf{R} (x_i)
output: {pred1-pred2} [N] (x_i)

This, however, means that there are different possible matches between the two nouns, that is, the relationship between them is not a constant, expressible by a single verb, though in real use there will only be one actual meaning. This range of possible matches that pred2 may have with pred1 could be represented by means of the symbol R in (b) (Allen 1978: 91). This variable establishes a range of possible meanings derived from the interaction of the sets of semantic features of the two nouns, and predicts that the meaning of the second input noun matches one of the available feature slots of the first input noun. The question now is which of the range of features the second input noun should match. And this leads to the third point to bear in mind:

c. There are *hierarchies* of semantic features of the two nouns.

There are features that are *cognitively more important*, as they represent the defining aspects of things. For example, two dominant features of a *mill* are that it is “powered

by something” and that “it produces or makes something” (ibid. cit).⁴ Allen then goes on to claim that the meaning of the second input noun must match a semantic feature with a dominant position in the hierarchical feature arrangement of the first noun (Allen 1978: 91-111).

The crucial factor is, therefore, the nature of the dominant features of the first input noun. We could even establish sets of cognitively important features for each semantic class of nouns:

- (12) second order entities
- PARTICIPANTS
 - TIME
 - DURATION
- first order entities
- naturally existing entities (<+ human> nouns)
 - OCCUPATION
 - IDENTITY (SEX, RACE)
 - naturally existing entities (<+animate> nouns, i.e. animals, plants)
 - APPEARANCE
 - HABITAT
 - LOCATION
 - natural objects (<-animate> nouns)
 - COMPOSITION
 - ORIGIN
 - LOCATION
 - synthetic objects (<-animate> nouns)
 - PURPOSE/ USE

The correct interpretation of a ‘noun + noun’ compound thus depends crucially on conceptual information about the nouns involved. What makes this proposal different from Dik’s (1980, also in 1997) is the role of the conceptual component in the process.

⁴ The idea that there are cognitively more relevant features the possession of which makes examples of a category more central members comes from prototype theory (Rosch 1973, 1975).

The next step would be to look for a way to represent this fact. If we were to use a revised version of the standard formulation for noun incorporation, the formation of noun-noun compounds would look roughly as in (13):

- (13) input: (a) pred1 [N] (x_i :<-animate>, <+synthetic>, <PURPOSE>, etc)
(b) pred2 [N], selR(x_i)
output: {pred1-pred2} [N] (x_i)

The string of semantic features in (a) indicates that whenever we encounter a noun referring to a non-animate synthetic entity, the feature PURPOSE is activated as a possible candidate to match one semantic feature of pred2; if the entity is human, the features <OCCUPATION>, <IDENTITY> will be activated instead. This way it will be possible to keep the FG standard proposal of Predicate Formation Rules.

However, it is not clear to me whether such a predicate formation rule makes sense in the new model of FDG, if we bear in mind that

- a. the process by which these compounds are created is not very different from that of free ‘noun + noun’ combinations. The only difference is that in the case of compounds, the combination is semantically more restricted and that a ‘noun + noun’ compound may become part of the lexicon.
- b. predicate frames are no longer stored as part of the predicates in the fund (see García Velasco & Hengeveld 2002).

An alternative to the standard FG formulation more in accord with FDG would be to use a sort of (conceptual) *frames*. When a speaker decides to refer to an entity he starts by constructing a mental representation or finding a conceptual frame and then he makes a decision on how to express this. For a compound to be formed it must be possible to establish a correspondence between both nouns by finding a schema connected with both words. This is somehow similar to the standard FG formulation, where it is required that the selection restrictions of the second input noun match the selection restrictions of the first input noun. However, whereas in standard FG the starting point is lexemes, which are taken from the fund with fixed predicate frames,

in FDG the starting point would be frames where two lexemes must be able to enter to form a compound.⁵

In a study on noun-noun compounds, Ryder (1994) has proposed the use of such frames, which she calls *linguistic templates*. She defines these templates as the speaker's stock of constructions for established compounds, and the schemas developed from them. These templates can be selected from a number of different types that form a continuum of increasing abstractness and generality:

- a. individual conventional expressions

(14) *doghouse*

- b. groups of compounds sharing a common first or second element (a core word)

(15) e.g. *X + box = a box intended to contain/store X (bread box, cash box, coin-box, lunch box, sandwich box, tool box)*

- c. patterns found in groups of individual compounds composed of varying nouns:

(16) e.g. *Substance that can be contained + Container = Container to hold/store the substance* (a pattern seen in many compounds with different component nouns: *cigar box, coke bottle, grocery bag*).

According to Ryder, the number of linguistic templates may vary depending on which is the head noun (i.e. the first input noun). Thus, if the input noun to be made more specific by another noun is *man*, there will be a number of different possible templates depending on X:

⁵ Many discussions of compounds make the explicit or implicit claim that given the starting point for the generation of compound is some sort of underlying semantic, syntactic or semi-syntactic structure of an appropriate form, that is, some sort of frame. The difference here is that the frame I propose would be conceptual, that is, existing at a prelinguistic level. Compounding is an optional process that speakers can choose to encode that structure or frame, but they can equally well choose to encode it in sentential form.

- (17) X (*location*) + *man* = man who lives and perhaps works in X
(e.g. *mountain man*)
- X (*instrument*) + *man* = man who works using X (e.g. *plowman*)
- X (*animal*) + *man* = man who works with/ tends X
(e.g. *cowman*)
- X (*formless, basic substance*) + *man* = man-shaped thing made of X
(e.g. *snow man*)
- X (*a produced object*) + *man* = man who produces/ sells/ transports X
(e.g. *milkman*)

My suggestion is that similar templates or frames should be introduced in FDG, though with some differences:

- a. Ryder's templates also include the phonological characteristics of both component and composite structures (Ryder 1994: 97) only mentions this when she describes the characteristics of the compounds she proposes). However, phonological features would be assigned at the acoustic level in FDG.
- b. Secondly, it is not clear to me how much specificity should be allowed, i.e. whether individual instantiations (i.e. conventional compounds already existing in the fund) and cases with a fixed lexeme should be considered frames. For me, frames would rather be conceptual.

Now, postulating conceptual frames for the creation of 'noun + noun' compounds poses a problem to the model: if the base for the creation of these compounds is conceptual and the conceptual component is separate from the grammar component, we would have to conclude that this process is not properly linguistic. An alternative solution would be to introduce conceptual information as part of lexical entries in a similar way to Pustejovsky's (1995: 101) *Qualia Structures* (Qs), illustrated in (20), that is, to try to integrate conceptual information as part of grammar or, at least, to make their interaction more explicit.

(18)

$$\left[\begin{array}{l} \mathbf{book} \\ \text{ARGSTR} \\ \\ \text{QUALIA} \end{array} \right] = \left[\begin{array}{l} \left[\begin{array}{l} \text{ARG1} = \mathbf{x: information} \\ \text{ARG2} = \mathbf{y: phys_obj} \end{array} \right] \\ \\ \left[\begin{array}{l} \mathbf{information.pis_obj_lcp} \\ \text{FORMAL} = \mathbf{hold(y,x)} \\ \text{TELIC} = \mathbf{read(e,w,x,y)} \\ \text{AGENT} = \mathbf{write(e',v,x,y)} \end{array} \right] \end{array} \right]$$

Pustejovsky QSs enable nouns to encode information about particular properties and activities associated with them. The sort of formalization proposed by him is probably the most accurate one available to represent the conceptual knowledge about the nouns involved which is relevant for the process.

Given the developmental stage of FDG, it is not my intention to come to any definite conclusions about what would be the most convenient way to represent the process of noun-noun formation, but it seems to me that a revision of the standard proposal is required and that the use of some sort of frames would be more in accord with the new version of the model.

3.3 ‘Noun + noun’ compounds and the contextual component

In the previous section we have seen that the conceptual component is the starting point for the formation and interpretation of ‘noun + noun’ compounds. In addition, however, the correct formation and interpretation of a noun-noun compound might be dependent on the particular context.

3.3.1 *The dynamic construction and interpretation of ‘noun + noun’ compounds*

Firstly, FDG is expected to account for the dynamic construction of linguistic expressions and the formation of ‘noun + noun’ compounds is a good example of this dynamism: although a ‘noun + noun’ compound may *normally* have a particular meaning (especially in the case of well-established compounds) which is derived from the particular frame and the semantics of the components, it might have a different one *under certain conditions*.

For example, example (19a) may have the meaning given in (19b) in a context where the topic is the building material of public buildings, but in an article about an exhibition on marble sculptures the interpretation will be the one given in (19b') (Plag 2003: 150)

- (19) a. *marble museum*
 b. 'museum building made of marble'
 b'. 'museum where marble objects are on display'

This means that the decision of which features will match the two nouns may occur in an online fashion since the interpretation of *marble museum* largely depends on the surrounding discourse. The addressee will have to check whether the linguistic frame he has chosen based on the concepts denoted by both nouns fits into the particular context in which the compound appears. For example, when facing the compound *marble museum*, the addressee might construct the frame

- a. $x2$ (*Material*) + $x1$ (*Building*) = $x1$ is built out of $x2$.

But the presence in the following context of one or more clues that marble is shown within the museum will make him look for a new frame like

- b. $x2$ (*Content*) + $x1$ (*Container*) = $x1$ has/ contains $x2$.

If these clues appeared in the previous co-text, the addressee would most probably get to the right interpretation at the first try.

3.3.2 Referential 'noun + noun' compounds

Furthermore, there are cases where the primary function of a nominal compound is a discourse referential one. A 'noun + noun' compound can be used to refer back to some explicit part of the co-text, taking it up again in condensed form (*anaphoric reference*).⁶ Consider the following example, cited by Adams (2001: 87):

⁶ Kastovsky (1982, 1986) points out that the function of compounds (and of other cases of word-formation) in these examples is that of *syntactic recategorization* rather than a lexical, *labelling* or *naming* function.

- (20) It seems to be that time of year again for the Guardian leader to make its ritual misguided suggestion that tinkering with the clocks could save lives on the road...what about the many people who are currently being killed in broad daylight?...Your clock-tinkering would not have saved them. (Guardian: 1993).

The meaning of a ‘noun + noun’ compound may also be explained further by the following co-text (*cataphoric reference*, or rather, *forward pointing compounds*). This is particularly important for writing headlines or advertisement texts. Consider in this respect example (21), where it is only the following context that reveals the meaning of the compound.

- (21) *Snake girl's record*

Two Chinese girls set record living for 12 days in a room with 888 snakes. (Guardian 1993, cited by Adams 2001)

Finally, a nominal compound can be used to refer to some piece of extralinguistic reality, where the compound is a sort of *deictic* device. Consider example (22), which is used as a deictic device to distinguish one seat from others in its proximity (Downing 1977). Many of these are online creations and do not survive beyond the situation in which they are created.

- (22) the apple-juice seat

In the case of discourse referential compounds like these contextual dependence is crucial. Turning to the previous examples: in (20) we need to resort to the previous discourse to get to the right interpretation “tinkering with the clocks” and, more importantly, to produce such an online compound. In (21) we have to look at the following co-text to get at the interpretation “girls living with snakes”, while in (22) the presence of a given referent in the situation that must be distinguished from others licenses the creation of the compound.

In all these cases there has to be a suitable referent in the context for the compound to be correctly interpreted. In Hengeveld’s FDG (Hengeveld 2004a: 3), the

communicative component is assumed to interact with the representational level “in order to enable later reference to earlier acts and expressions”. More recently, Hengeveld (2005: 62) has pointed out that

the various levels of representation within the grammar feed into the contextual component, thus enabling subsequent reference to the various kinds of entity relevant at each of these levels once they are introduced into the discourse. The contextual component feeds into the operation of formulation, so that the availability of antecedents and visible referents may influence the composition of (subsequent) discourse acts.

In cases like (20) we could say that the speaker makes a referential act (R) in which reference is made to a previous event (expressed as a phrase). However, this referential act is now expressed not as a sentence but as a compound lexeme. This is represented in (23):

(23)

(A ₁ : [DECL (P ₁) _S (P ₂) _A (C ₁ : [...	(R ₁)	...] (C ₁))](A ₁))
	(e ₁)	
	(Lex ₁)	

Since the function of ‘noun + noun’ compounds in these cases is limited to that of discourse reference, the representation of this type of compounds should include an operator to express the anaphoric relation with either the preceding discourse or the communicative setting. The co-indexing of the nouns forming the compound with previous or following elements in the co-text makes it possible for the compound to adopt the semantic relationship holding between both elements. For instance, in the case of *clock-tinkering*, the instrumental relation of *clock* with the head noun is copied from the semantic role of the satellite *with the clocks*, which has appeared in the previous discourse, as shown in (24):

(24) Past e_i: [*tinker* [V] (x_i)_{Ag}] (*the clocks*_{Instr}]
 [Ae_i] [[*clock*_N]_{ModP N}]_{RefP}

In addition, the process by which these compounds are formed is different from that of non-discourse-referential ‘noun + noun’ compounds. The two nouns involved are no longer subject (or not necessarily) to the kind of conceptual relationships that hold between the two nouns in the formation of ‘noun + noun’ compounds since the kind of relation that holds between the two nouns is not habitual, i.e. conventional or typical (see Zimmer 1971, 1972). In other words, in many cases no frame could be found to allow the interpretation of the compound since this interpretation relies only on the surrounding discourse or on the speech situation. The context overrides everything else.

3.4 ‘Noun + noun’ compounds and the modular approach

In Section 3.2 I have suggested a revision of the way to account for processes of compounding. If we decide to get rid of predicate formation rules for the process of ‘noun + noun’ compound formation and account for this process by means of a mechanism that takes into account the new modular conception of FDG, we would have to bear in mind different points at different levels.

3.4.1 Prelinguistic conceptual level

First of all, at the *prelinguistic conceptual level* the choice of a ‘noun + noun’ compound, particularly in the case of online creations, starts with the intention of the speaker to present his information in a packaged format and with the corresponding mental representations (i.e. a conceptual frame). Thus, if I want to refer to one specific type of surgery as ‘surgery with the help of a computer’, I may choose to present this in an extended format and this decision will result in the formation of a modified term phrase, e.g. ‘surgery which uses a computer’ or ‘computer-assisted surgery’. I can, however, also choose an even more condensed form. If I decide to present this in the (so to speak) “packaged” format, my decision will result in the creation of a compound. i.e. ‘computer surgery’.

Besides, in joining two nouns to create a name for a particular referent, the speaker must integrate two conventional nouns in such a way as to allow the listener to pick out the referent, i.e. in creating a compound the speaker must bear in mind the (assumed) knowledge of the listener. The speaker knows that the listener expects

there to be a correspondence between the meanings of the two nouns. He must try to construct a mental/conceptual representation, or rather to find one semantic frame from the stock of frames he has. This frame might look like (25):

$$(25) \quad x_2 \text{ (Synthetic object)} + e_1 \text{ (activity)} = \text{'}e_1 \text{ uses } x_2 \text{ as instrument'}$$

The speaker will then check whether the semantic schemas of the two lexemes *computer* and *surgery* fit into this frame. From our conceptual knowledge about *surgery* and *computer* we know that surgery is a second order entity and that computers are synthetic objects (i.e. first order entities characterized as <-animate>). We know that a cognitively important feature of synthetic objects is the purpose or use to which they are put, e.g. synthetic objects can be used as instruments. On the other hand, activities (i.e. second order entities) involve participants who may play different roles. Therefore, there can be assumed to be semantic compatibility between the two concepts, since a possible logical relationship between both can be that of Activity-Instrument.

3.4.2 Representational level

The conceptual representations are translated into semantic representations at the *representational level*. The representational level will contain the appropriate (compatible) semantic frames of the two nouns, illustrated in (26):

$$(26) \quad (e_i: [\text{surgery}]) (e_i: [\text{computer}]_{N\text{Instr}} (e_i))_N$$

At the representational level, we will also have to represent some further semantic information. For example, it has often been observed that when a noun is incorporated to create a verbal predicate, this involves some kind of semantic change in the incorporated noun. For instance, the noun is said to be non-individuated or to have lost its individual salience both semantically and syntactically (Dik 1980: 38f.; 1997).

3.4.3 Morphosyntactic level

The representation at the representational level is translated into a morphosyntactic representation at the structural level through the operation of morphosyntactic encoding. At the *morphosyntactic level* two lexically filled units at the

representational level are inserted into a single complex word template (that is, a morphological template), as shown in (27) :

$$(27) \quad \begin{array}{l} [[\text{lexeme}_N]_{\text{ModP}} \text{lexeme}_N]_{\text{RefP}} \\ [[\text{computer}_N]_{\text{ModP}} \text{surgery}_N]_{\text{RefP}} \end{array}$$

The morphosyntactic level will have to account for different structural/morphosyntactic facts:

- a. Firstly, if the formation of the compound is licensed given the description provided at the representational level, an appropriate template will dictate the order in which the two nouns will appear, so that the second input noun (i.e. the modifying noun) will appear (at least in this case) in first position.
- b. In addition, morphologically speaking, English compounds are distinguished in that their first member cannot be inflected on its own. Thus, plural markers can typically not be attached to the first element. This is just an effect of the ‘Right-hand head’ rule (Williams 1981).⁷

3.4.4 *Acoustic level*

Finally, at the *acoustic level* a template will dictate the new phonological patterning (especially stress pattern). As a general rule, compounds have stress on the first word, phrases on the second (Bloomfield 1933: 228; Chomsky & Halle 1968: 16f.).

In order to assign the correct phonological pattern, the acoustic level must be provided with the information that the combination in question is a compound and not a nominal phrase. Since both types of linguistic expressions are structurally similar, the speaker probably moves directly from the interpersonal to the acoustic level.⁸ Once the speaker decides to express a given content in a packaged way, he may select

⁷ The right-hand head rule refers to an important systematic property of English compounds, namely that their head normally occurs on the right-hand side. The compound as a whole inherits most of its semantic and syntactic information from its head. Thus, if the head is a noun, the compound will be a noun, if it has feminine gender, the compound will have feminine gender, and if the compound is pluralized the plural marking occurs on the head. The inheritance of features from the head is also known as *feature percolation*.

⁸ That is to say, the first step the speaker takes when he decides to express something in a “packaged” way, i.e. as a compound, is to assign this compound a particular stress pattern.

a combination of two combined lexical items (i.e. a compound) which serve this purpose and this decision is passed on to the acoustic level:

(28) / kɒmpju:tə sɜ:dʒəri /

4. Concluding remarks

To sum up, in this paper I hope to have been able to show that the creation of ‘noun + noun’ compounds is a productive process that involves both conceptual and contextual information. My aim has been to question the place of this type of compounds within the model of FDG, analysing its interactions with the conceptual and contextual components, and to suggest the need of a new way to account for these processes more in accord with the new conception of the model. Thanks to the inclusion of the conceptual and contextual components, the new proposal of FDG is expected to be able to account for such processes in an adequate way.

References:

- Adams, V. (1973) *An Introduction to Modern English Word Formation*. London: Longman.
Adams, V. (2001) *Complex Words in English*. London: Pearson Education.
Allen, M. (1978) *Morphological Investigations*. University of Connecticut.
Anderson, S.R. (1985) ‘Typological distinctions in word formation’. In: Shopen, T. (ed.) *Language Typology and Syntactic Description*. Vol. III. Cambridge: Cambridge University Press.
Bauer, L. (1998) ‘When is a sequence of two nouns a compound?’. *English Language and Linguistics* 2: 65-86.
Bloomfield, L. (1933) *Language*. New York: Holt.
Carstairs-McCarthy, A. (2002) *An Introduction to English Morphology*. Edinburgh: Edinburgh University Press.
Chomsky, N. & M. Halle (1968) *The Sound Pattern of English*. New York: Harper and Row.
Dik, S.C. (1980) *Studies in Functional Grammar*. London-New York: Academic Press.
Dik, S.C. (1997) *The Theory of Functional Grammar, Vol. II: Complex and derived constructions*. Edited by Kees Hengeveld. Berlin: Mouton.
Downing, P. (1977) ‘On the creation and use of English compound noun’. *Language* 53.4: 810-842.
García Velasco, D. & K. Hengeveld (2002) ‘Do we need predicate frames?’. In: Mairal, R. & Pérez Quintero, M. J. (eds) *New Perspectives on Argument Structure*. Berlin: Mouton, pp. 95-123.
Hengeveld, K. (2004a) ‘The Architecture of a Functional Discourse Grammar’. In Mackenzie, J.L. & M.A. Gómez-González (eds) *A New Architecture for Functional Grammar*. Berlin: Mouton, pp. 1-21.

- Hengeveld, K. (2004b) 'Epilogue'. In Mackenzie, J. L. & M.A. Gómez-González (eds) *A New Architecture for Functional Grammar*. Berlin: Mouton. pp. 365-378.
- Hengeveld, K. (2005) 'Dynamic expression in Functional Discourse Grammar'. In: Casper, de G. & K. Hengeveld (eds) *Morphosyntactic expression in Functional Grammar*. Berlin: Mouton. pp. 53-86.
- Jespersen, O. (1961) [1942] *A Modern English Grammar on Historical Principles* (Vol. VI: Morphology). George Allen & Unwin: London.
- Kastovsky, D. (1982) 'Word formation: a functional view'. *Folia Linguistica* 16: 181- 188.
- Kastovsky, D. (1986) 'The problem of productivity in word-formation'. *Linguistics* 24: 585-600.
- Levi, J.N. (1978) *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.
- Marchand, H. (1969) *The Categories and Types of Present Day English Word Formation*. München: C.H. Beck.
- Plag, I. (2003) *Word-formation in English*. Cambridge: University Press.
- Pustejovsky, J. (1995) *The Generative Lexicon*. Cambridge MA: MIT Press.
- Rosch, E.H. (1973) 'Natural categories'. *Cognitive Psychology* 4: 328-350.
- Rosch, E.H. (1978) 'Principles of categorization'. In: E. Rosch & B. Lloyd (eds) *Cognition and Categorization*, Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 27-48.
- Ryder, M.E. (1994) *Ordered Chaos: the interpretation of English noun-noun Compounds*. Berkeley and Los Angeles: University of California Press.
- Selkirk, E.O. (1981) 'English compounding and the theory of word structure'. In: Moorgat, M., H. Hultst & T. Hoekstra (eds) *The Scope of Lexical Rules*. Dordrecht: Foris, pp.229-277.
- Selkirk, E.O. (1982) *The Syntax of Words*. Cambridge MA: MIT Press.
- Warren, B. (1978) *Semantic Patterns of Noun-Noun Compounds*. Gothenburg Studies in English 41.
- Williams, E. (1981) 'On the notions "lexically related" and "head of a word"'. *Linguistic Inquiry* 12: 245-274.
- Zimmer, K.E. (1971) 'Some general observations about nominal compounds'. *Working Papers on language Universals*. Stanford University, no. 5, C1-21.
- Zimmer, K.E. (1972) 'Appropriateness conditions for nominal compounds'. *Working Papers on Language Universals*, no. 8, 3-20.