Categorial Deficiency and Phrase Structure

ABSTRACT: This paper sets out to derive the co-occurrence restrictions between functional and lexical elements as an upshot of categorial deficiency, the fact that functional items encode unintepretable versions of lexical heads. Coupled with a theory of syntactically active and LF interpretable categorial features, such as the one in Baker (2003), the hypothesis can derive the properties of mixed projections – such as gerunds, the motivation behind head movement up a projection, as well as hitherto unexplained but fundamental characteristics of phrase structure.

1 Introduction

Although selection in syntax has been investigated in depth, perhaps in a depth comparable to that achieved in the investigation of non-local dependencies, the (selectional) relation of lexical heads, such as N and V, to functional heads dominating them, such as D and T, has been inadequately researched into.

Before proceeding to offer an account of what regulates the relation between lexical and functional elements, I wish to bring two cases to the readers' attention as to why this issue is of importance. The first is of mainly conceptual and the second of descriptive nature.

In the earlier generative grammar, this contrast between lexical and functional categories was partially captured as a differentiation between projecting categories (lexical categories) and non-projecting functional categories (as in Jackendoff, 1977; Emonds, 1985 and chapters in Muysken & Van Riemsdijk, 1986). The selectional relation between a lexical (i.e. projecting) and a functional (i.e. non-projecting) functional category could be quite straightforwardly captured as one of c-selection. The subsequent realisation that a model in which functional categories also project proved to

be of greater explanatory capacity (Fukui & Speas 1986) and the ensuing revolution led to a unified X' schema for all categories and the abandonment of non-endocentric projections like S (now IP) and S' (now CP). At the same time it led to three undesirable consequences:

a. The intuitive idea that the head of a phrase is 'what this phrase is about' (as originally in Chomsky, 1970 and as recently as in Pinker, 1994: 106-7) cannot be captured any longer in an uncomplicated fashion. For instance, the phrase *the keys of heaven* (ultimately about keys, a sort of keys) was formerly described as an NP headed by the N *keys* and complemented by the PP *of heaven* with non-projecting Det *the* as its specifier. After the DP hypothesis, the phrase is described as a projection of *the*. Obviously, though, the phrase is not about 'the's'

b. The empirical observation that the majority of functional heads in human languages are affixal or phonologically dependent as well as semantically dependent on a lexical one becomes harder to capture. Consider the following simple example from Italian.

(1) legevano la gazetta
read-PAST-3RDPLU the newspaper

The above was formerly readily analysable as a VP with an Infl specifier and an NP complement (the Theme object); it has now to be accounted for in terms of its being an Agr_SP (Chomsky, 1993) or TP (Chomsky, 1995). The affixal status of the T head (say, '*-va-*') and the Agr_S head (say, *-no*) is reduced to a curiosity of morphology independent from the fact that these affect / instantiate the event of reading.

As implicit above, if lexical projections are in the complement of functional ones, they apparently have to be selected by them. Even conceding to this, if all selection is local, Ds select Nums that will select NPs. So, the relation between D and N is indirect. In (1),

Agr_SP *-no* will not select the VP, but the TP *-va-* which will select the vP which will select the VP. Hence, a C, which is less related to the actual verb of a clause than T is, will 'select' this verb in the same manner as this verb's Tense and as indirectly.

Let us again consider the functional head T(ense).[1] It has been argued that T must select and c-command an event at LF (cf. Laka, 1990), or – alternatively – that it expresses the event argument of a predicate (cf. Stowell, 1996). Let's now take a stem expressing an 'inherently' eventive concept, like *walk* and see whether a Tense head can (ultimately)[2] select it.

(2) [Our walk(***ed*) home by the river] had a therapeutic effect

(3) We walk*ed* home by the river

Morphological intricacies of English aside, we see that (3) is grammatical: not only did tense morphology show up affixed to the stem *walk*, but also the specifier of a T head – the subject *we* – is present. Neither of these is the case in (2), where no T occurs in the bracketed constituent, a DP. What this illustrates is that selection relations between functional and lexical elements are not completely reducible to semantics. We will return to this in section 6.

To summarise, something needs to be said about the status of the lexical head at the end of a projection line (e.g. C…T…(Neg)…v…V) as the 'semantic head' (Abney 1987) of the projection in question. This is almost impossible to capture in the current Principles and Parameters / Minimalist framework. At the same time, something needs to be said about the fact that T cannot select (a constituent containing) a noun in (2). In this paper I will attempt to do both by hypothesising that it is lexical heads that select the functional

[1] Chomsky (2000; 2001a) considers it a lexical head, apparently for reasons relating to the definition of phases. We will adhere to the standard assumption of T being a functional head here.

[2] The caveat "ultimately" will henceforth accommodate the fact that other functional heads (can) intervene between T (or D) and the lexical verb (or noun).

ones further up their projection line by virtue of the latter being specified for *uninterpretable* versions of the lexical head's categorical features. I hope that this analysis will offer not only a coherent account of the intuitive notion of 'semantic head' but that it will also explain selection between functional and lexical items by shedding some light on the category the former are specified for. Finally it will provide unified motivation for Head Movement within a projection line and capture mixed projections, such as gerunds, and some simple but elusive facts about phrase structure.

The paper will be organised as follows: in section 2 I will pose the question of the direction of selection in a projection line; in section 3 I will present and review the literature on the matter. Section 4 will review the analyses offered by recent frameworks positing category-neutral roots being active in syntax. Section 5 will present and elaborate an alternative based on the uninterpretability of the categorial features of functional heads, Categorial Deficiency, whereas Section 7 will put the emerging theory to the test by reviewing the predictions it can make. The subsequent section discusses Phrase Structure and derives properties thereof from the uninterpretability of such features. Section 9 concludes the paper.

2 What selects what, why and how.

Before discussing the problem of selection relations between lexical and functional categories we need to clarify four issues:

1. *The syntactic versus semantic nature of selection*, that is whether (some) selection relations between lexical and functional heads are extra-syntactic and can be reduced to purely semantic requirements. This could probably be true for some of these relations. For instance, Baker (2003) argues that D takes an NP as / in its complement because of semantics: nouns bear referential indices and

determiners, being an essentially referential category as well, can only match with them for interpretive reasons.

2. *The exclusive* (*or not*) *character of selection*. There are functional categories that appear to combine only with verbs (e.g., T) or nouns (e.g. Num); this situation Panagiotidis (2002) liberally calls *biuniqueness*, a term I will take up here for mnemonic purposes.[3]
3. *The direction of selection*: whether it is the functional head that selects the related lexical one (i.e. D $\rightarrow$ N) or the lexical head that selects the functional one (i.e. N $\rightarrow$ D).
4. *The mechanism of selection*. This question is an extension of the dilemma of question 2 above. Evidently, it is necessary to describe how selection comes about. Is it instantiated via purely selectional features, as is the claim in part of the literature? Or does it reflect deeper realities regarding the nature of categories and clause structure?

In this paper I will attempt to answer the above questions. In order to do so, I am going to present a new theory of grammatical category of *functional* elements, remaining less specific on the issue of the categorial features of lexical elements. The latter topic is definitely a most exciting one and, fortunately, one recently brought to spotlight because of Baker's (2003) book. For the purposes of this paper, then, Baker's theory of lexical categorial features will be adopted here as a working hypothesis.

Before anything else, a review of previous accounts of the categorial nature of functional elements is in order.

[3] It appears that some functional elements felicitously combine with both: C / ('functional') P could be an instance of this more promiscuous kind, following analyses by Emonds (1985), Starke (1995), Kayne (1998) and others. I will not real with adpositions here.

3 Ouhalla (1991) and Extended Projection.

Grimshaw (1991) and Ouhalla (1991) are the best known and most detailed attempts towards an account of what the grammatical category of functional elements might be.

Ouhalla (1991) hypothesises that functional heads simply bear the categorial feature of the lexical head inside their complement. In other words, D will have an [N] feature, T a [V] feature and so on. Such a formulation appears to be a sound basis to start from if one wants to account for cases of biuniqueness, as presented in point 2 of section 2 above. To illustrate Ouhalla's idea, if *walk* in (2) encodes an [N] feature, it will be unavailable for selection by T, which in turn is marked for a [V] feature, because feature mismatch will occur – to recast Ouhalla's analysis in more up-to-date terms. Feature mismatch will cancel the derivation, therefore *–ed* cannot select / be selected (by) a noun:

(4) the [*little town* that was *savagely bombed* by its *liberators*] was called …

The complement of the D *the* above, quite uncontroversially, contains five lexical heads: two nouns (*town* and *liberators*), an adjective (*little*), a verb (*bombed*) and an adverb (*savagely*). Which of these does D have the categorial specification of? The clear intuition of course is that *town* is the relevant lexical head here. In a similar context, Abney (1987) discusses the notion of the *semantic* head of a projection: *the* plus its complement are, ultimately, about a town.

Grimshaw (1991) puts forward the notion of *Extended Projection* in order to capture this intuition. At the same time, giving her Extended Projection a purely syntactic character, she tries to obliquely address point 1 of section 2: particular functional heads combine with particular lexical ones because of selectional features they encode, rather

because of semantic needs.[4] An Extended Projection *y* of a lexical head *x* is defined as follows:

(5) *y* is an Extended Projection of *x* iff:

a. *y* dominates *x*

b. *y* and *x* share all categorial features

c. all nodes between *x* and *y* share all categorial features

d. For $F_n(x)$ and $F_m(y)$, $n \leq m$

Her feature $F_{0...2}$ is supposed to be encoded on every entry coming from the lexicon and its purpose is to define this entry's position in the Extended Projection it belongs to: lexical heads, which head Extended Projections, are specified as F_0, Infl and Det as F_1, Comp and P as F_2. Crucially, functional heads are specified for the categorial feature of the head of their Extended Projection, like in Ouhalla and as expressed in (5) above. Biuniqueness is thus automatically derived: T heads will always be inside the Extended Projection of a verb, and Ds inside that of a noun because they are specified for the categorial feature of a verb and a noun respectively. In (4), *town* is the head of the Extended Projection that has the D *the* as its syntactic head. An added attraction of (5) is that the relative ordering of the various functional elements inside the Extended Projection is also catered for. Before closing this subsection, let me just briefly mention the updated version of the concept of Extended Projection by Grimshaw (2003) and, crucially, van Riemsdijk's (1998) reworking thereof in order to cater for more than two functional elements per Extended Projection, as well as the existence of semilexical elements of verbal and nominal nature. Van Riemsdijk's analysis has the added virtue of eliminating the $F_{0...n}$ feature and appealing to two general principles in order to derive

[4] I am grateful to an anonymous reviewer for raising this issue with me.

the way (Extended) Projections, his M(acro)-Projections are built. He still has to appeal to a F(unctional) feature, though, as the one defining functional heads.

It should have emerged by now that in order to understand the categorial nature of functional heads, we need a theory for the categorial nature of lexical ones. Most theories of grammatical category, whether functionalist or formalist in their inspiration, restrict themselves to laying out descriptive part-of-speech systems. Because these systems are *just* that, i.e. descriptive, they cannot predict the syntactic behaviour or the semantic interpretation of an element *x* belonging to a category *Y*. Stowell's (1981) reinterpretation of the Chomsky-Jackendoff [±N] [±V] system is a bright exception, even if superseded by research itself. Stowell attempted to link the values of the [N] and the [V] features to syntactic properties of the lexical heads encoding them, e.g. only [-N] can assign Case. Hence, in the majority of generative proposals about categorial features, [N] and [V] are perceived as syntactic / classificatory features with no reasonable LF interpretation. Déchaine's (1993) system is an important step to recasting categorial features as LF interpretable features. For instance, Déchaine perceives [V] to mean 'predicative'. Other systems, such as Chomsky's and Jackendoff's make no predictions, whereas Givón and Croft work with a theory of categories that takes them to be prototypes along a continuum of temporal stability, in the vein of Ross (1973): verbs are prototypically the least time-stable elements, nouns the most time-stable ones and adjectives fall somewhere in the middle. Of course a prototypical account a priori fails to explain anything: the fact that a prototypical noun has time stability, is comfortable with *tachyon* being a noun: it is simply a *non-prototypical* noun.

Baker (2003) advances an explanatory theory of category. In his system, [V] and [N] are privative features. [V] entails that the lexical projection can project a specifier, [N] entails the presence of a referential index, the absence of either entails a categorially

unspecified category, Adjective.[5] Extending the account, verbs are specified as [V] and are interpreted at LF as predicates, either by virtue of [V] or by being in a predicative relation with their specifier, or both. Nouns are specified as [N] and are interpreted at LF as referential either by virtue of [N] or by virtue of their referential index, or both. Adjectives are unspecified and are interpreted at LF solely as a bundle of descriptive features, a fact making them ideal modifiers. Now, Baker (2003) wishes to derive biuniqueness on a semantic basis. He consequently convincingly argues that Quantifier and Number can only combine with Noun, exactly because only Noun is sortal. Nevertheless, there are instances where biuniqueness can be explained away on morphosyntactic terms: hence the exclusive combination of Tense with Verb in the overwhelming majority of languages is attributed to the affixal nature of Tense in these languages and the blocking effect of functional categories intervening between T and V. It seems that if one embraces Baker's explanatory system (putting it to work he is able to derive most of the special properties nouns and verbs display as well as the lack of any special properties of adjectives), it is not necessary to argue for matching categorial features on functional heads: the functional-lexical selection can be morphosyntactic (T-V) or semantic (Num-N), depending on the situation. In fact Baker explicitly argues that his system does exactly this: void the need for matching categorial features between lexical and functional heads in order to derive their co-occurrence restrictions (2003: 324). Nevertheless, as we will see in section 5, not only is it possible to reconcile the two systems but it is also necessary to do so.

[5] Accordingly, nouns are *sortal* in the sense of Geach (1962): they are the only category it is coherent to ask whether a particular instance thereof is *different* or the *same* to another. The interested reader is referred to Chapter 3 of Baker (2003). Baker considers v lexical, a part of the V-shell. I will gloss over the matter here for ease of exposition.

4 Do functional categories 'make' lexical ones?

In the mid-nineties, a new account of the selection relations between functional and lexical elements emerged. Within the framework of Distributed Morphology, consensus has arisen that lexical categories such as nouns, verbs and adjectives are morphological epiphenomenona and that categorial features are extraneous to syntax (Halle & Marantz 1993; 1994, Marantz 1997; 2000, Harley & Noyer 1998, Embick 2000, Bale & Barner 2002). What *is* relevant to syntax is lexical roots which are unspecified for category such as 'noun', 'verb' or 'adjective'. When lexical roots are inserted in a nominal environment, say an Extended Projection consisting of, *D*, *Num* and *n*, they surface as nouns; when they are inserted in a verbal environment, say an Extended Projection consisting of *C, T*, *Aux* and *v*, they surface as verbs. An example of this approach is as follows: a root like √*destroy* will become a noun in (6) and a verb in (7). This is so because in (6) the root √*destroy* is embedded within a 'nominal' Extended Projection, and because in (7) it is embedded within a 'verbal' Extended Projection.

(6) The army's destruction of the settlement (was a humane act).

(7) The army will have destroyed the settlement (by tomorrow).

The problem of selection between functional and lexical elements is thus explained away: D does not 'co-occur' with nouns; rather the projection D (Num, n…) belongs to, the nominal Extended Projection, *makes* the lexical root a noun. Similarly, it is only natural that T and Asp 'co-occur' with verbs, as it is the Extended Projection they belong to that makes the root morphologically surface as a 'verb'. In later work, Marantz (2000) and Chomsky (2001b: 9) identify *v* and *n* as the actual verbaliser and nominaliser of roots. This makes the account even more attractive for two reasons: first, category assignment to the root from the Extended Projection becomes more straight-forward, as it is now carried out by a single functional head, v or n, not a group of

functional heads in unison. Second, and in relation to the first point, mixed projections, in the sense of Borsley & Kornfilt (2000), can now be better captured. To illustrate on English nominal gerunds (as realised in some dialects at least), suppose they somehow contain *both* a D *'s* and a v. The D is responsible for the genitive case on the gerund's Agent *Jebediah's*, the Agent itself appearing courtesy of a v, which also assigns accusative to the Theme argument *her*:

(8) [Jebediah's discovering her in the bar] triggered a chain of events.

If v, rather than a 'nominal' or 'verbal' Extended Projection in unison assigns category to the lexical root (√*discover* in this case), then we can explain the essentially verbal nature of the gerund, in (8) and in general: we cannot add *–ing* to a noun.

The attractiveness of an account whereby unspecified roots are 'nominalised' and 'verbalised' by n and v respectively is evident, at least in terms of conceptual simplicity. Nevertheless, the class of analyses invoking category-neutral roots that are nominalised or verbalised according to their respectively being inserted in a nominal or a verbal functional complex, stumble upon two major problems, already discussed in detail in Chomsky (1970): the issue of (*un-*)*predictability* and that of *productivity*.

If a single root, say √*discover*, underlies phonologically similar words belonging to different categories like *discovery* and *discover*, then the only difference among them should be that of category. More precisely, their lexical semantics should be identical, pace argument structure restrictions imposed by the different syntactic environments (nominal, verbal, adjectival): in any case, the outcome should be semantically predictable. Moreover, if the process that inserts roots into nominal, verbal and adjectival environments is syntactic, then we would expect it to be fully productive and produce a noun, a verb and an adjective with any given root. Briefly put, every root would be expected each to become a noun, a verb and an adjective and with very similar

meanings. The first prediction, predictability, is borne out with √*discover*: both the noun and the verb have identical meanings, pace the different environments they are inserted in. As for the second prediction, we already run into problems: there is a noun *discovery*, a noun *discover* but no adjective. Turning to other roots, such as √*water*, we have roughly the opposite problem: although a noun *water*, a verb *water* and an adjective *watery* all exist, their meanings are not predictable from each other and, moreover, they are quite distinct. Actually, dozens of examples of roots that fail to become nouns, verbs or adjectives exist, as well as of roots that take up unpredictable and idiosyncratic meanings as nouns, verbs and adjectives: the interested reader is referred to Chomsky (1970: 189-190; ibid: endnote 11).

Nevertheless, this is not really the place for a detailed criticism of the research programme that seeks to reduce the combination of functional and lexical elements to a combination of functional elements with category-neutral roots. Suffice it here to say that such a programme requires syntax both to dramatically tamper with the semantics of roots it manipulates (recall $_{N}$*water*, $_{V}$*water* and *watery*) and to overgenerate elements such as *$_{A}$ *discover*.

5 The category of functional elements.

Let me now proceed into presenting an alternative analysis of the issues raised in section 2. Although it is indebted to Ouhalla (1991) and Grimshaw (1991), it runs counter to them for two reasons:

Both Ouhalla's and Grimshaw's analyses proliferate the number of identical categorial features in an Extended Projection, to use Grimshaw's term. Take a concrete example like the one below:

(9) They will probably not have finished it.

Taking *will* to be a Tense head here, *probably* to be a specifier of a dedicated projection (Cinque, 1999), *not* to head a Neg projection, *have* to be an Aux and *finished* to be a V in a v position, we count six [V] features by both Ouhalla's and Grimshaw's analyses. Whether the surface order reflects the order these functional heads are merged in is not important here, what *is* important is the fact that, according to both, all five functional and the lexical head V are specified for a categorial [V] feature. What is role of these six instances of [V] at LF? What is the LF interpretation of categorial features? Obviously none, [V] is treated here as a purely selectional feature. In fact *all* generative accounts of how many and which categorial features exist, except Déchaine (1993) and Baker (2003), fail exactly for this reason: in all of them categorial features are not interpretable on *either* interface. Nevertheless, if we suppose them to be purely selectional features, then we are better off with a Marantz-style account (as discussed in the previous section): otherwise categorial features will have the fate of Case features, which must be completely eliminated by LF.

Furthermore, Ouhalla and Grimshaw capture biuniqueness and the syntactic nature of functional-lexical relations at the cost of not being able to distinguish between functional and lexical items, a distinction robust both in psycholinguistic and semantic terms. This inability stems from the fact that in both accounts lexical as well as functional categories bear identical categorial features: [N] and [V]. In Ouhalla's analysis they can only be distinguished according to their position: lexical categories occupy the bottom of a projection line. Similarly, in Grimshaw's work, a differently valued feature F stipulates the different position of the head in an Extended Projection: a 0 value places it at the bottom, making it lexical and the head of the Extended Projection, a 1 value places it higher and a 2 value at the top of the Extended Projection. Once more, the role of the feature $F_0 \ldots F_n$ at LF is at best unclear.

Wishing to capture Ouhalla and Grimshaw's key intuition that the functional-lexical relation is syntactic and sharing with them the understanding that functional heads like D, Num and n are essentially nominal whereas T, Asp and v verbal, I propose the hypothesis below:

(10) *Categorial status of functional heads* (preliminary): Functional heads are specified for [*X], the uninterpretable version of the categorial feature [X] of the lexical head X at the bottom of their projection line.

The hypothesis in (10) captures exactly the intuition that the functional members of the Extended Projection of the verb (noun, adjective) are somehow verbal (nominal, adjectival), without seeking recourse to a notion of Extended Projection. It also foregrounds the syntactic nature of the co-occurrence between functional and lexical heads, while suggesting it may reflect or force particular *semantic* relations. Before going on to explain why uninterpretable categorial features [*X] are necessary and what they explain, it is necessary to clarify which of the issues stated in section 2 are answered by (10), while elaborating it. I will turn to this right below.

6 The syntactic nature of the functional-lexical selection

6.1 The semantics of grammatical category.

If what bonds a functional head (e.g. D) together with its related lexical one (e.g. N) is the drive for the elimination of LF uninterpretable features, then we have a fruitful way to account for the intuition that, although captured syntactically, this is a relation that either *reflects* certain semantic requirements or, alternatively, *forces* particular interpretive options. In order to make this claim more substantial, it would be worth

looking at what kind of LF property I understand (interpretable) categorial features to encode.

Recall that, according to Baker, [V] entails that verbs, the lexical category encoding it, can project a specifier; [N] entails that nouns, the lexical category encoding it, will bear a referential index. Nevertheless, as Baker himself acknowledges, there is at least one *functional* category that also projects a specifier and acts predicatively: Pred. Pred is by hypothesis the category responsible for the predicative versions of nouns and adjectives:

(11) a. Fatima [$_{VP}$ ~~Fatima~~ likes spaghetti].

b. Fatima is [$_{PredP}$ ~~Fatima~~ a lawyer].

c. Fatima is [$_{PredP}$ ~~Fatima~~ young].

In this way, Pred can mediate between the inherently referential Noun or the drastically underspecified Adjective and a subject, in SpecPredP, hence emulating a Verb, so to speak. Similarly, C can also bear a referential index: this is most strikingly illustrated in Modern Greek, where a CP, whether marked [+wh] (13) or not (12), can be the complement of a D, without the involvement of a (null) noun (Roussou, 1991):

(12) gnorizo (to) [$_{CP}$ oti agonizeste]

know.1STSG the that fight.2NDPL

'I know that you struggle.'

(13) gnorizo (to) [$_{CP}$ pos agonizeste]

know.1STSG the how fight.2NDPL

'I know how you struggle.'

What is it then that distinguishes between V and Pred or between N and C? Obviously the fact that V and N are lexical, whereas Pred and C functional, that much Baker himself acknowledges (2003: 325). Given that Baker's, as well as any other coherent theory of categorial features, capitalises on this necessary distinction between lexical

and functional, how is it to be captured? Incorporating Baker's account into (10), we can now revise it as follows:

(14) *Categorial status of functional heads* (revised): Functional heads are specified for [*N] or [*V], the unintepretable version of the categorial feature [N] or [V] of the noun N or the verb V at the bottom of their projection line.

Does the above entail that functional-lexical co-occurrence restrictions are an internal affair of syntax? Certainly not; nobody, I would think, would wish to consider such restrictions on a par with, say island effects. What (14) brings about, however, is the necessity to capture functional-lexical co-occurrence restrictions in a unified way: as feature (mis)match between unintepretable and interpretable features, not as a labour to be divided between morphosyntax and semantics, depending on the case. This (mis)match might very well reflect properties of the Conceptual-Intentional systems, for instance the incompatibility between a non-sortal category like adjective or verb and Number, or it might be a genuine idiosyncrasy of the syntactic component, like the T-N mismatch seems to be.[6] The essential thing is that both kinds are syntactically encoded in terms of an unintepretable-interpretable pair. That much is enough, for the time being, as regards the nature of the lexical-functional co-occurrence and biuniqueness; interestingly (14) also entails that in an intuitively very clear way, it is lexical heads that categorially 'select' functional ones by virtue of their interpretable categorial feature.

6.2 Why are functional heads categorially deficient?

The short answer to the question heading this subsection could be 'because they cannot be semantic heads'. In more detail, functional heads have a number of properties that

[6] More on the T-N mismatch in Panagiotidis (in preparation).

distinguish them from lexical heads. Following Abney (1987: 64-65), let us review the properties relevant for the discussion here:

a. Functional elements are generally phonologically and morphologically dependent. They are generally stressless, often clitics or affixes, and sometimes phonologically null.

b. Functional elements lack […] "descriptive content". Their semantic contribution is second order, regulating or contributing to the interpretation of their complement. […]

Moreover, Abney (1987: 285) discussing the oddity of the concept of 'intransitive determiners' notices that

> "[…]the appearance of functional elements as "intransitives" […] constitutes a systematic exception to the otherwise general requirement that [functional elements] take an obligatory complement."

Leaving morphological dependence behind for the moment, we observe that functional elements lack descriptive content and that they take an obligatory complement, whereas lexical elements may do so but *not as part of their categorial definition*: they may or may not take complements for reasons of subcategorisation.

Descriptive content is a notion notoriously difficult to pin down. First of all, grammatical nouns and verbs (Emonds, 1985; Panagiotidis, 2002) also known as semi-lexical heads (see van Riemsdijk, 1998 and chapters in Corver & van Riemsdijk, 2002), despite being lexical, have probably much less semantic content than, say, a Quantifier or a Tense head. Moreover, next to the lexical N, the functional D and C can also be referential – cf. (12) and (13) above – by virtue of a referential index: some languages take N arguments, some take D and C arguments and some all three kinds. Undeniably, distinguishing between lexical and functional elements cannot be done on the basis of

'semantic weight'. As Baker (2003: 311), once more, aptly notes, "no counting of features will do". Still, functional elements are semantically deficient, their semantic contribution being second order. This deficiency we capture here with (14): it is not the number of LF interpretable features a particular head carries that makes it functional or lexical, it is whether it marks an uninterpretable categorial feature. Given that uninterpretable features cannot survive at LF, they attract a matching interpretable one: thus functional heads are always semantically subservient to lexical ones.[7]

From the uninterpretability of a functional element's categorial feature, we can also automatically derive Abney's observation that they *must* have a complement.[8] Actually, there must exist a matching lexical head within the complement. So, there must be an N somewhere in the complement of D, so as the [N] feature of the former can eliminate D's [*N] feature by LF.

Finally, recalling the problematics laid out in the opening section, we now have a way of defining the "semantic head" of a projection without seeking recourse to Grimshaw's F_0 feature: the element in the projection line bearing (only) an interpretable categorial feature *is* the semantic head: *the two keys of heaven* is 'about' keys and not about 'the'. This is because key bears the categorial feature [N], not the Num projection or the D the. Similarly, *we walked home* is about walking, not home, a completed action or an action in the past. This is not because of some ad hoc privilege of the interpretable categorial features [N] and [V], and their hosts, but, rather, because they normally are the only ones in their projection that are interpreted as bearing a referential index or predicates.

---

[7] This new account opens up the possibility that there exist functional elements with descriptive features. This fits very smoothly with van Riemsdijk's (1998) account of semilexical heads, maybe epithets could be analysed this way: the field is open to research.

[8] Of course, there are no intransitive Ds: Corver & Delfitto (1999), Panagiotidis (2002).

7 Biuniqueness and mixed projections

As stated above, by applying (14) we can straightforwardly account for the fact that a D...V projection is illicit: the [*N] feature of D cannot be checked by the [V] feature of V, as they do not match, leading the derivation to crash. That much can be predicted by both Ouhalla's (1991) and Grimshaw's (1991) accounts, as well as any account reducing biuniqueness to semantic or morphological factors, including Baker (2003). Biuniqueness is usually not the problem, instead there is a more interesting and as yet unresolved question with respect to category matching: the status of *mixed projections*, in the sense of Borsley & Kornfilt (2000). I will dedicate this section to consider this crucial but elusive matter.

In a considerable number of languages there are constructions that combine nominal and verbal characteristics. The best studied among them, although certainly not a one-off oddity cross-linguistically, is English (verbal) gerunds.[9] Since Chomsky's (1970) *Remarks on Nominalization*, gerunds have been an intensively studied topic (Hazout 1994 and Hudson 2003 provide overviews as well as references). The general conclusion seems to be that

(15) The lower part of a projection containing a gerund is verbal and the higher one nominal

The generalisation for the lower part as verbal can be justified as follows (Hudson, 2003; Borsley & Kornfilt 2000, 104-5):

[9] I will unexceptionally take this term as equivalent to both '*Poss-ing*' (with a genitive subject and an accusative direct object), '*Acc-ing*' (with an accusative subject and an accusative direct object) and 'verbal' (with a PRO subject and an accusative direct object) gerunds here, glossing over dialectal differences. See (at least) Reuland (1983) and Hudson (2003) for some discussion on dialectal differences regarding the status of the above types of gerundival constructions.

a. Only verbs have a gerundive form: *draw-ing*, *book-ing*, *collid-ing* but **poem-ing*, **collision-ing*.

b. Gerunds can take an internal argument marked for accusative: [*Nina*(*'s*) *pulling him out of the water*] *saved Knut*. This is currently standardly taken to be a reflex of a transitive v head.

c. In some dialects gerundival forms can be dominated by an Aspect projection: [*Nina*(*'s*) *having pulled him out of the water*] *saved Knut*. Aspect is typically associated with verbal projections.

d. Gerunds take adverbs: [*Nina quickly pulling him out of the water*] *saved Knut*. Adverbs, of course, are part of verbal projections.

The generalization for the higher part as nominal can be justified as follows:

a. The external argument of the gerund may appear in genitive:[10] [*Nina*(*'s*) *pulling him out of the water*] *saved Knut*. Genitive is a reflex of / assigned by a D head.

b. Projections containing gerunds can be used as arguments. In the example [*Nina's pulling him out of the water*] *saved Knut*, the bracketed phrase is the subject of *saved*. Only nominal projections can be used as arguments. Note that this is also true even if the D layer (i.e. *'s*) is missing: the bracketed constituent in [**He* / *Him pulling him out of the water*] *saved Nina* cannot be a TP, as it cannot take a nominative subject *she* and has to go by a default accusative *her* instead.[11]

[10] Not so in all dialects. See also Hudson (2003) for discussion of some idiosyncrasies.

[11] This is the line I will be following here, but see Reuland (1983), Johnson (1988) and Dubinsky & Williams (1995) for discussion of (external) Case assignment to subjects of gerunds.

What is the solution to his puzzle? We either need to relax biuniqueness claiming that, for instance, D can occasionally select an AspP, or we need to come up with a special 'Ger' hermaphrodite head.

Relaxing biuniqueness in general is rather undesirable, as we end up with the possibility of a host of unattested combinations: *D-TP, *D-CP (in English, at least), *D-vP. Our best shot is probably to follow Reuland's (1983) and Hazout's (1994) accounts, representative of a long line of research. Based on evidence from gerunds in English, Arabic and Hebrew, they essentially claim that there is a gerundive head Ger (Reuland identifying it with *–ing*) that selects verbal projections (or AspPs). Interestingly, they claim this Ger head to be of category N, a noun in other words: in Reuland's (1983:113) words: "a nominal inflection marker […] [i]ts feature matrix will be abbreviated as [N]. The question of why other nouns, lexical nouns more precisely, cannot occupy the position of this head taking AspP complements is not addressed, probably because this is understood to be a matter of lexical selection, hence something to be relegated to the lexicon. Now, the idea that a Ger head is essentially a Noun is interesting, but its being selected by only a possessive D, if at all, and its taking an AspP (not TP) as a complement had better follow from general principles rather than be idiosyncratic specifications of a lexical item. This is desirable for both methodological (we prefer general explanations than ad hoc ones) and typological reasons: gerunds show up in language after language, English, Arabic and Hebrew being at least three of them.

Seeking a general syntactic property governing the behaviour and internal structure of gerundive phrases, I once more turn to Baker's (2003) theory of lexical features.

Baker (2003) takes [N] to be semantically interpreted as 'bearing referential index'. He also notices (Baker 2003: 139) that certain *functional* heads, such as the C *that*, seem to also bear referential indices, without them being nouns, of course:

(16) a. I know [$_{DP}$ Aristotle's definition of tragedy].

b. I know [$_{CP}$ that Aristotle defines tragedy as imitation].

In the face of (16), and recalling (12) and (13), a natural conclusion is that *that*, next to some other functional items belonging to the category Pred and, maybe, D, bear a referential index, hence are specified as [N]. Two questions immediately spring up:

a. Why isn't *that* a noun? Why is it a functional head?

b. Why does *that* take verbal complements?

In a felicitous twist of the affair, the two questions answer each other: *that* bears a referential index, hence is marked as [N]; it is also a verbal functional head, hence is marked as [*V], according to what I claimed in section 5 above. So, *that* bears the following categorial features: [N] [*V]. Its [N] feature makes it referential, its [*V] forces it to have a V head in its complement. It is now easy to see that a gerundive head, conventionally labelled here as Ger, also has the categorial feature specification [N] [*V]. Ger must take a verbal complement, hence is marked for a [*V] feature, and can be argumental and /or be selected by a [*N] category, like D, hence is marked for an [N] feature. Now it is just the fact that not all Ds can take GerP as a complement that must be somehow encoded in the lexicon.[12] What is then the LF contribution of Ger? Exactly that of *that*: making a verbal projection referential by virtue of its interpretable [N] feature. If it is by virtue of [N] and its referential index, rather than the presence of a D head, as Stowell (1991) and Longobardi (1994) claim, that an expression can function as an argument, then we could say that a Ger Phrase can stand as an argument without

[12] Historically gerunds used to take more determiners, most notably *the*; another issues is that gerunds cannot be combined with Num and Q:

(i) * [Four Sam('s) driving her dad's car] were more than enough.

(ii) * [All (Pete's) writing a book] was unsuccessful.

I will return to such examples and provide explanations in future research.

being shelled by a DP. This is borne out if the bracketed phrases in (17) below are bare GerPs, as we would expect on minimal assumptions, on a par with referential determiner-less noun phrase arguments in a great number of languages.[13]

(17) a. [$_{GerP}$ PRO Defining tragedy] is fun.

b. Aristotle enjoyed [$_{GerP}$ PRO defining tragedy].

We now have a way to capture *all* mixed projections and unify them with 'ordinary' items like *that*, see Panagiotidis (in preparation) for details and a more cross-linguistically oriented discussion. All we have to say is that some functional heads, like those of the Ger type, can have interpretable categorial features, [N] in this instance. They still behave as functional though, as they are still marked for an uninterpretable [*V] feature.

The above also explain why there are no nominal-adjectival mixed projections: a hypothetical head taking an adjectival complement but being selected by [*N] heads would be marked as [N]. Nevertheless, such a head would be unable to exclusively select adjectives, as adjectives are not specified for either [N] or [V] (Baker 2003: 190-2). So, our hypothetical head could not mark a [*X] categorial feature, hence being only specified for [N]. An [N] head is of course a lexical or semi-lexical / grammatical noun – see Emonds (1985), Panagiotidis (2002) and contributions in Corver & van Riemsdijk (2002) for discussion. That is why nominal-adjectival (or verbal-adjectival ones, for that matter) mixed projections are impossible and unattested.

Finally, turning to the issue of the 'semantic head' of a projection, we realise that in a mixed projection, there are going to exist more than one interpretable categorial feature. Does this mean, in the case of a gerund for instance, that both V and Ger are semantic

[13] I do not wish to speculate here on the categorial status of bare plural and mass arguments in English as in *Wild ducks love peanut butter*.

heads? Although this is a logical possibility, it seems to be the case that the two actually *combine* in a way so that the gerundive projection receives a Janus-like referential / predicative reading. This is also captured by 'single-head' analyses by Pullum (1991), Lapointe (1993), Malouf (2000), Hudson (2003). Nevertheless, in the next section we will see that no such mechanisms are necessary, in the face of the availability of head movement or a notational variant thereof.

8 Categorial deficiency in Phrase Structure

The central hypothesis of this paper has been that functional items are *categorially deficient*. I have so far presented the consequences of this categorial deficiency functional items are defined by with respect to their LF interpretation as well as the nature and behaviour of mixed categories. I will now proceed to explore the syntactic repercussions of this deficiency by employing it in order to explain the mechanism underlying Head Movement. Having said that, the categorial deficiency hypothesis will paradoxically provide us a means of either motivating it at LF – something quite problematic until this day, see Roberts (2003) and Harley (2003) for discussion, or altogether dispensing with it. Before proceeding, I will supply a, for the purposes of this paper, final version of (10) and (14): *Categorial Deficiency*.

(18) *Categorial Deficiency:* Functional heads are categorially deficient: they are specified for [*X], where X is the categorial feature [N] or [V].

8.1 Head movement, morphological dependency and categorial deficiency

Head (X^0-) movement is usually taken to operate in two instances. The first one typically involves a lexical head climbing up its projection line. Thus, the V head X^0- moves to v in English, to v and then to T in Romance, to v and then to T to end up in C

in main clauses of Verb-Second languages (den Besten, 1983). The N head (probably) X^0-moves to D in Danish. Chomsky (1993: 18; 1995: 269) suggests that the V head must climb at least to T cross-linguistically, whether overtly (as in Romance) or covertly (as in English). Finally, Riemsdijk (1998) shows climbing to be a more general operation than usually thought, attesting inter alia instances of N to D to P movement, in Lezgian for instance. I will informally call this instance of X^0-movement *Climbing*, in order to distinguish it from a second type of X^0-movement. This second kind of X^0-movement is *Incorporation*, a lexical head moving to another one (N to V, ordinarily; Baker, 1988; 1996). *Hale-Keyser Incorporation* (Hale & Keyser 1993) of N to P to V is a subcase of this. I will have not much more to say on Incorporation here.

Turning back to Climbing, we realise it is interesting for three reasons. First, it appears to be unmotivated. Although morphology has been invoked time and again as a possible motivation for it (cf. Rohrbacher, 1999), it has been lately becoming clear that an object of any size and morphological complexity can be in principle inserted under any syntactic node, see Baker (2003: 275-290). At the same time, a semantic motivation for climbing is simply not available.

Second, it is extremely widespread: as Roberts (2003) points out, most languages seem to display some instances of climbing, especially in their verbal domain. It is difficult to give a generalised morphophonological account for a phenomenon that occurs across so large and so morphologically diverse a set of languages.

Finally, Climbing seems to be intimately related to the morphological dependence of functional heads – recall Abney's criteria in section 6.2. Climbing is indispensable as

a. either a mechanism 'assembling' inflectionally complex forms: *lege–* would climb to T 'picking up' *–va–*; then the complex head $[_T$ V T]

lege–va– would end up in 'Agr' / verbal D to pick up *–no* to finally assemble *lege–va–no*.

b. or a mechanism for eliminating uninterpretable features of morphologically complex forms merged in functional heads; recall *legevano* in (1) and suppose it is merged in 'Agr' or verbal D (Ritter 1995; Manzini & Savoia 2002; Panagiotidis, 2002): V would have to climb to T and then, as a complex head [$_T$ V T] to 'Agr' / verbal D to check, inter alia, the uninterpretable tense feature of the form.

It generally seems to be the case that, at least in the verbal domain, inflectionally complex forms tend to appear higher in the clause, one only need remember the classic analysis of Pollock (1989).

Fortunately, (18) provides us with a way of understanding Climbing as an LF-motivated operation, whatever its exact effect on morphology: this type of head movement's function is to eliminate [*X], uninterpretable categorial features, by LF if no other mechanisms are available to this end. Illustrating, let us take a verbal projection consisting of T, Asp, v and V. By hypothesis, V bears a [V] feature, all the other functional heads a [*V] one each. In a language like English, the verb's [V] feature will move overtly to v pied-piping the entire feature matrix with it; the rest of the climbing to Asp and T must be completed before LF. In a language like French, the verb's [V] feature will move overtly to v, to Asp and then to T, pied-piping the entire feature matrix with it.[14] In every 'Extended Projection' only interpretable categorial features will survive, [V] and / or [N].

[14] Exactly this is proposed in Chomsky (1995: 269): "The operation Move F […] raises the categorial feature *v* of the verb V [to T] carrying FF[*v*] along automatically in a derivative chain. If the operation is covert, as in English, then nothing else happens […] if the operation is overt, as in French, then [the moved element] is V itself, pied-piped to allow convergence at PF".

As it is beyond the scope of this paper to discuss the status of head movement in a Bare Phrase Structure model, I will only restrict myself to saying that categorial deficiency does not, of course, propose a *mechanism* of Climbing, or – a fortiori – X^0 movement. Categorial deficiency does not even require that Head Movement exist. What it does is to force only categorially matching heads in a projection line, banning projections such as *D…Asp…Num…X; at the same time it does not preclude the presence of both [V] and [N] in a (typically gerundival) projection line, given that both are interpretable ([N] contributed by Ger and [V] by the verb). In other words, categorial deficiency provides the LF motivation for either Climbing or *an equivalent operation*. The obvious candidate of such an equivalent operation is the relation Agree: in a projection line at LF, heads must be in a relation of Agree with each other as regards their categorial features. Incidentally, both Climbing and categorial Agree capture Extended Projections without the stipulated machinery in (5). Extended Projections are 'secured' by Climbing / Agree, they are not a primitive. Which mechanism is the correct one is ultimately an empirical matter.[15]

8.2 Phrase structure and categorial deficiency

I will now proceed to showing that (18) subsumes Li's Generalisation and a fundamental, but not extensively discussed, aspect of phrase structure: lexical heads can only appear at the *bottom* of projection lines.

[15] In fact, the proposal herein is much closer to van Riemsdijk's (1998) theory of (M-)Projections, although not needing either of his principles that force the construction of 'proper' projections or an F feature for functional and an L one for lexical categories.

Li (1990) observes that it is impossible for a lexical head to X^0-move to a functional head and then for the resulting complex head to move to a *lexical* head. Hence a configuration like the following is impossible:

(19) *$[_N [_D$ N D] N]… $[_D$ ~~N D~~]… ~~N~~

As is almost evident, the above structure, generated by two applications of head movement, one of N to D and the subsequent from D to the higher N is illicit. Li's Generalisation captures this but does not explain it. On the other hand, Categorial Deficiency rules it out and encompasses Li's Generalisation on the grounds of Economy. The first instance of head movement, N to D, is triggered by the need of the [*N] feature of D to be checked by N's interpretable [N]; N head-moves to D and adjoins to it; [N] checks and eliminates [*N]. Subsequently, the resulting complex head $[_D$ N D] bears only one categorial feature, [N], but so does the higher N head. This means that there is no trigger for a second instance of X^0-movement of the complex head $[_D$ N D] to the higher N, as both bear interpretable features, D's [*N] having been eliminated and lexical heads otherwise marking only interpretable features (cf. Chomsky 1995:232, 378).[16] The absence of trigger entails the impossibility of $[_D$ N D] to N movement, on the grounds of Economy.[17]

Suppose now that the higher head were a V, as below:

(20) *$[_V [_D$ N D] V]… $[_D$ ~~N D~~]… ~~N~~

To the problems above, the mismatch between the [N] feature of $[_D$ N D] and the [V] feature of V would be added. In this instance, an Economy violation would be escorted by feature mismatch between [N] and [V]. Hence, Categorial Deficiency, coupled with

[16] I abstract away from uninterpretable φ features on D; these will have already been checked as free riders against those of the adjoined N. I also abstract away from the Num head intervening between D and the lower N.

[17] If V in (20) is replaced by the categorially featureless A, the problem remains.

the simplest version of Economy, derives the impossibility of X^0 movement from a lexical to a functional and on to a lexical head at no extra cost. An added advantage is that such a relation, in terms of *categorial features*, naturally, is banned as an Agree relation, as well. Otherwise phrased, Li's Generalisation also holds for 'covert' head movement.[18]

8.3 Categorial deficiency and the position of the lexical head

It is standardly, and tacitly, understood that in a line of projection, the lexical head always appears at the bottom. Even in the richer system of Riemsdijk (1998), the general hierarchy is functional…semi-lexical…lexical heads. A question at this point is whether this order, which seems to be universal, can be motivated by a universal principle. I will show in this section that Categorial Deficiency has the answer.

Let us start with the attested case, exemplifying on a fragment of verbal projection:

(21) T…Asp…V

T and Asp both bear a [*V] feature each. Climbing or categorial Agree ensures that they will be both checked by the interpretable [V] feature of V. Suppose now a configuration like the one in (22) below were to be created by Merge:

(22) T…V…Asp

First of all, (22) would probably be legible by LF, containing all the LF-interpretable features that (21) does. Although we do expect the semantic component to be concerned with the relative ordering of *functional* heads in a projection – for tense-anchoring purposes, for instance, as in Laka (1990), Enç (1987), Demirdache & Uribe-Etxebarria (2000), Roberts & Roussou (2002) and elsewhere – observe that, in (22), T dominates

[18] P to V movement has been proposed by Neeleman (1997) but, even if P is functional (hence [*N]), this is not a problem for Li's Generalisation (and the analysis here) because the operation is not triggered by categorial features.

Asp exactly as it does in (21). Hence, although caution is necessary when making claims about what the semantic component can or cannot do, knowing little about it, we can probably safely assume here that (22) does not violate any semantic requirements. So, why is it unattested, or, as I will now claim, impossible?

In (22), the [V] feature of V checks via Move or Agree and eliminates [*V] of T. Nevertheless, the [*V] of Asp will remain unchecked. This would lead to (22) crashing at LF, because of the presence of an uninterpretable feature, [*V] on Asp, in the derivation. For general Locality reasons, both attempts of categorial Agree / categorial head movement to save (22) illustrated below are impossible: (23) would involve raising and then lowering, (24) would be an instance of lowering and then raising.

(23) *V to T to Asp

PLEASE INSERT FIGURE 1 HERE

(24) *V to Asp to T

PLEASE INSERT FIGURE 2 HERE

Once more, Categorial Deficiency provides a simple and straightforward account: lexical head projections appear at the bottom of their projection lines for categorial feature-checking purposes. In fact, we now have the theoretical account and the methodological means to fully side with Panagiotidis (2002) and, ultimately, Ouhalla (1991) and Grimshaw (1991): (uninterpretable) categorial features are indeed the ‘glue’ that holds projections together. At the same time, we can uphold and expand Baker’s (2003) insightful and restrictive account of the two interpretable categorial features [V] and [N].

9 Conclusion

I have taken up the idea in Panagiotidis (2002) that functional heads bear uninterpretable categorial features and coupled it with Baker's (2003) account of how many and which categorial features exist. The result is a theory of Categorial Deficiency for functional heads, encapsulated in (18), which can capture the predictions of Grimshaw's (1991) Extended Projection and Riemsdijk's (1998) (M-)Projection, without the need for special mechanisms or ad hoc 'selectional' features. I proceeded to show that in the light of Baker's conception of categorial features, Categorial Deficiency can also capture mixed projections, such as English gerunds, as an [N] feature is what makes a projection 'referential' and there is no stipulated ban on functional items bearing interpretable categorial features as well. Finally, I demonstrated Categorial Deficiency to motivate Climbing (the main instance of X^0-movement) or categorial Agree, as well as simple (but unanswered up until now) facts about phrase structure.

Bibliography

Abney, Steven P. 1987. *The English Noun Phrase in its Sentential Aspect*. Unpublished PhD Thesis. MIT.

Baker, Mark C. 2003. *Lexical Categories: Verbs, Nouns and Adjectives*. Cambridge: CUP.

Barner, David & Bale, Alan 2002. No nouns, no verbs: psycholinguistic arguments in favor of lexical underspecification. *Lingua* 112: 771-791

Borsley, Robert & Kornfilt, Jaklin 2000. Mixed Extended Projections. In: Borsley, Robert (ed.) *The Nature and Function of Syntactic Categories*. [*Syntax and Semantics 32*]. London: Academic Press. 133-166

Chomsky, Noam 1970. Remarks on Nominalization. In: Jacobs, Roderick, and Rosenbaum, Peter. (eds.) *Readings in English Transformational Grammar*. Waltham, Mass.: Ginn & Company. 184-221.

Chomsky, Noam 1993. A Minimalist Program for linguistic theory. In Hale, K. & Keyser, S.-J. *The view from Building 20*. Cambridge, Mass.: MIT Press. 1-52

Chomsky, Noam 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam 2000. Minimalist inquiries: the framework. In Martin, R., Michaels, D. and J. Uriagereka (eds.) *Step by Step. Essays on Minimalist Syntax in honor of Howard Lasnik*. Cambridge, Mass.: MIT Press. 89-155

Chomsky, Noam 2001a. Derivation by Phase. In: Kenstowicz, Michael (ed.) *Ken Hale: A life in language*. Cambridge, Mass.: MIT Press. 1-52

Chomsky, Noam 2001b. Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics* 20.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads.* Oxford: OUP.

Corver, Norbert & Delfitto, Dennis 1999. On the nature of pronoun movement. In: Riemsdijk, Henk van (ed.), *Clitics in the Languages of Europe*. Berlin: De Gruyter. 799–861

Déchaine, Rose-Marie 1993. *Predicates across categories*. Unpublished PhD thesis, University of Massachusetts, Amherst.

Demirdache, Hamida & Uribe-Etxebarria, Myriam 2000. Primitives of Temporal Relations. In: Martin, R., Michaels, D. and J. Uriagereka (eds.) *Step by Step. Essays on Minimalist Syntax in honor of Howard Lasnik*. Cambridge, Mass.: MIT Press. 157-186.

den Besten, Hendrick 1983. On the interaction of root transformations and lexical deletive rules. In Abraham, W. (Ed.) *On the formal syntax of Westgermania*. Amsterdam: John Benjamins

Dubinsky, Stanley and Williams, Kemp 1995. Recategorization of prepositions as complementizers: the case of temporal prepositions in English. *Linguistic Inquiry* 26: 125-137

Embick, David 2000. Features, syntax and categories in the Latin perfect. *Linguistic Inquiry* 31: 185-230

Emonds, Joseph 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.

Enç, Mürvet 1987. Anchoring conditions for tense. *Linguistic Inquiry* 18: 6633-658.

Fukui, Naoki and Speas, Margaret. 1986. Specifiers and projections. *MIT Working Papers in Linguistics* 8: 128–172

Geach, Peter 1962. *Reference and Generality*. Ithaca: Cornell University Press.

Grimshaw, Jane 1991. *Extended Projection*. Unpublished manuscript, Brandeis University.

Grimshaw, Jane 2003. *Extended Projection*. Unpublished manuscript, Rutgers University.

Hale, Kenneth & Keyser, Samuel Jay 1993. On argument structure and the lexical expression of syntactic relations. In: Hale, Kenneth & Keyser, Samuel Jay (Eds.) *The View from Building 20*. Cambridge, MA: MIT Press. 53-109

Halle, Morris & Marantz, Alec 1993. Distributed morphology and the pieces of inflection. In: Hale, Kenneth & Keyser, Samuel Jay (Eds.) *The View from Building 20*. Cambridge, Mass.: MIT Press. 111-176.

Halle, Morris & Marantz, Alec 1994. Some key features of distributed morphology. *MIT Working Papers in Linguistics* 21, 275-288

Harley, Heidi & Noyer, Rolf 1998. Licensing in the non-lexicalist lexicon: nominalisations, vocabulary items and the Encyclopaedia. *MIT Working Papers in Linguistics* 32: 119-137

Harley, Heidi 2003. *Merge, conflation, and head movement: the First Sister Principle revisited.* Talk delivered at NELS 34.

Hazout, Ilan 1994. Nominalizers in theta theory. *The Linguistic Review* 11: 5-48

Hudson, Richard 2003. Gerunds without phrase structure. *Natural Language and Linguistic Theory* 21: 579-615

Jackendoff, Ray 1977. *X' Syntax: A Study of Phrase Structure*. Cambridge, Mass.: MIT Press.

Johnson, Kyle 1988. Clausal gerunds, the ECP and government. *Linguistic Inquiry* 19: 583-609

Kayne, Richard 1998. *A note on prepositions and complementizers*. Ms. New York University

Laka, Itziar 1990. *Negation in Syntax. On the nature of functional categories and projections*. PhD Thesis. MIT

Lapointe, Steve 1993. Dual lexical categories and the syntax of mixed category phrases. In: Kathol, A. and Bernstein, M. (eds.) *ESCOL '93*. Columbus, Ohio: Ohio State University. 199-210

Li, Yafei 1990. X^0-binding and verb incorporation. *Linguistic Inquiry* 21: 399-426.

Longobardi, Giuseppe 1994. *Reference and proper Names: A Theory of N-movement in Syntax and Logical Form*. Linguistic Inquiry 25, 609-665

Malouf, Robert 2000. Verbal Gerunds as Mixed Categories in Head-Driven Phrase Structure Grammar. In: Borsley, Robert (ed.) *The Nature and Function of Syntactic Categories*. [*Syntax and Semantics 32*]. London: Academic Press. 133-166

Manzini, Maria-Rita & Savoia, Leonardo 2002. Parameters of subject inflection in Italian dialects. In Svenonius, Peter (ed.) *Subjects, Expletives, and the EPP*. Oxford: OUP. 157-199

Marantz, Alec 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. *U. Penn Working Papers in Linguistics* 4: 201-225

Marantz, Alec 2000. *Reconstructing the lexical domain with a single generative engine*. Unpublished ms. MIT.

Muysken, Pieter & Van Riemsdijk, Henk (Eds.) 1986 *Features and Projections.* Dordrecht: Foris.

Neeleman, Ad 1997. PP complements. *Natural Language and Linguistic Theory* 15: 89-137

Ouhalla, Jamal 1991. Functional Categories and Parametric Variation. London: Routledge.

Panagiotidis, Phoevos 2002. *Pronouns, clitics and empty nouns: 'pronominality' and licensing in syntax*. Amsterdam: Benjamins.

Panagiotidis, Phoevos (in preparation) *Categorial features and mixed projections*.

Pinker, Steven 1994. *The Language Instinct*. London: Penguin Press.

Pollock, Jean-Yves 1989. Verb movement, Universal grammar and the structure of IP. *Linguistic Inquiry* 20: 365-424

Pullum, Geoffrey 1991. English nominal gerund phrases as noun phrases with verb phrase heads. *Linguistics* 29: 763-799

Reuland, Eric J. 1983. Governing '–ing'. *Linguistic Inquiry* 14:101-136

Ritter, Elizabeth 1995. On the syntactic category of pronouns and agreement. *Natural Language and Linguistic Theory* 13: 405–443.

Roberts, Ian 2003. *Bare Head Movement*. Unpublished ms., University of Cambridge.

Roberts, Ian & Roussou, Anna 2002. The EPP as a condition on the Tense dependency. In Svenonius, Peter (ed.) *Subjects, Expletives, and the EPP*. Oxford: OUP. 125-155

Rohrbacher, Bernhard Wolfgang (1999). *Morphology-driven Syntax*. Amsterdam: John Benjamins

Ross, John 1973. Nouniness. In: Fujimura O. (ed.) *Three Dimensions of Linguistic Theory*. Tokyo: TEC Co. Ltd. 137-258

Starke, Michal 1995. On the format for small clauses. In Cardinaletti, Anna & Guasti, Maria Teresa Eds. Syntax and Semantics 28. San Diego: Academic Press. 237-269

Stowell, Tim 1981. *Origins of Phrase Structure*. Unpublished PhD thesis, MIT.

Stowell, Tim 1991. Determiners in NP and DP. In: Leel, K. & Bouchard, D. (eds.), *Views on phrase structure*. Dordrecht: Kluwer. 37–56

Stowell, Tim 1996. The phrase structure of Tense. In: Rooryck, Johan & Zaring, Laurie (eds.) *Phrase Structure and the Lexicon.* Dordrecht: Kluwer. 277-291.

Riemsdijk, Henk van 1998. Categorial feature magnetism: the endocentricity and distribution of projections. Journal of Comparative Germanic Linguistics 2: 1-48

FIGURE 1

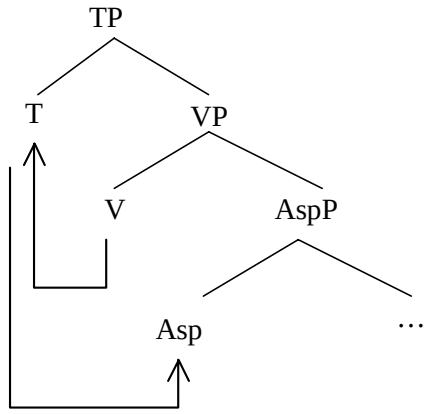

FIGURE 2