Hedde Zeijlstra
University of Tuebingen
Nauklerstrasse 35, #3.11
72074 Tuebingen, Germany
+49.7071.2977156
hedde.zeijlstra@uni-tuebingen.de
www.heddezeijlstra.nl

Functional structure, formal features and parametric variation: consequences of conflicting interface conditions

1. Introduction

In current minimalist reasoning language is assumed to be a 'perfect' solution to the task of relating sound and meaning (Lasnik 2002). The claim that language is 'perfect' in this sense seems to be at odds with several appearances of 'imperfections' that one finds in grammar. In this paper I address three such issues: (i) parametric variation; (ii) (c)overt functional structure; and (iii) formal features. These phenomena have been formed a big puzzle in the study of generative grammar.

First, if language performs a perfect task in relation sound and meaning, why would it opt for so many differences? Since the introduction of the Principles and Parameters program (Chomsky 1981), consensus has arisen on the fact that linguistic variation is not unlimited and the notion of parameter has been introduced.[1] Still, the fact that grammatical variation is limited by a (fixed) number of innately present parameters lacks a principled explanation. It is not clear why languages must vary syntactically. The need for such an explanation becomes more and more urgent since modern research, especially since the success of microparameteric studies, has revealed that the numbers of parameters cannot is no longer easily countable, but is rather big. There is no clear notion at the moment how big this number should be, but given the idea that each human I-grammar results from a different parameter setting, the number must be at least 7846 to serve the number of all human beings current grammar. Since previous and possible I-grammars must be taken account as well, the number may even exuberantly increase.

Second, given the large amount of syntactic variation, the question arises whether languages are structurally identical. Several approaches have taken the strong radical claim that each language has underlying the same functional structure and that grammatical variation is restricted to which positions are overtly realised. Proposal along these line have been formulated by (Cinque 1999; Kayne 2000; Starke 2005) and (to a lesser extent) (Rizzi 1997).[2] If this approach is on the right track, (rich) functional structure can be taken to be part of UG and the language-specific realisation be reduced to parametric variation. However, it is not explained why UG

[1] For a recent debate on the status of parameters in linguistic theory, cf. (Newmeyer 2004; 2005; 2006; Roberts & Holmberg 2006).

[2] The approach that languages share a template for functional structure is known as the cartographic approach. Many papers written within this perspective can be found in (Cinque 2001; Rizzi 2004; Beletti 2005).

should innately be equipped with such a rich structure. Several scholars, e.g. (Ernst 1998; Svenonius 2001; Nilsen 2003) have argued against such a universal UG-based functional sequence (terminology due to (Starke 2001)). In my paper I also provide several empirical arguments against a universal underlying structure. Moreover, I address the question why functional structure is required in the first place. I conclude that functional structure can be best explained as resulting from a phonology-semantics mismatch and should be considered to be flexible amongst grammars.

Third, the notion of formal features requires a principled explanation. Especially the fact that formal features are manifested either as being interpretable [iF] or uninterpretable [uF] seems to be a grammatical 'imperfection.' Why would language use redundant material in the first place? The question why one would adopt redundant material in a grammatical theory is purely grammatical: at first sight language seems to be full of redundancy as suggested by Concord phenomena (such as Negative Concord), overt agreement or multiple markers. The line of reasoning developed in (Chomsky 1995) was to take one ugly thing to license the other. Adopting the principle of Full Interpretation that states that the interfaces should be free of uninterpretable material (i.e. [uF]'s) have to be deleted during the derivation, before reaching LF. If their position in the clause would not provide the proper checking configuration, a syntactic operation such as Move had to be triggered. The necessity of an instance of dislocation has thus been triggered by the presence of redundant material. Note that this may very well explain the presence of dislocation effects against the background that all movement is in principle 'uneconomic,' but it leaves the occurrence of redundant material itself to remain an open question (Chomsky 2006).

In my paper, I address all three 'imperfections' that require principled explanation, and I argue that the existence of parameters, formal features and functional structure follows from the fact that language forms an optimal solution to mapping meaning to form, but that it is not the perfect solution. In its very essence the proposal that I formulate amounts to saying that each grammar is a maximally optimal solution to conditions imposed by the different interfaces. Since interface conditions may be conflicting, language opens up the possibility that different grammars may equally optimal satisfy their interface conditions. This already creates a grammatical space that allows for cross-linguistic variation. Ideally, one would then endorse the most radical hypothesis, namely that the room for grammatical variation that follows from different interface conditions is the exact parametric space.

I adopt the hypothesis that formal features are required to trigger dislocation operations, and that dislocation operations are required in order to spell-out two markers of different semantic operations in one and the same possibility, a desideratum that follows from standard phonological economy conditions. Hence formal features are syntactically flexible. Furthermore, I demonstrate that projection can apply only to formal features, not to semantic features, hence accounting for the flexible nature of functional structure.

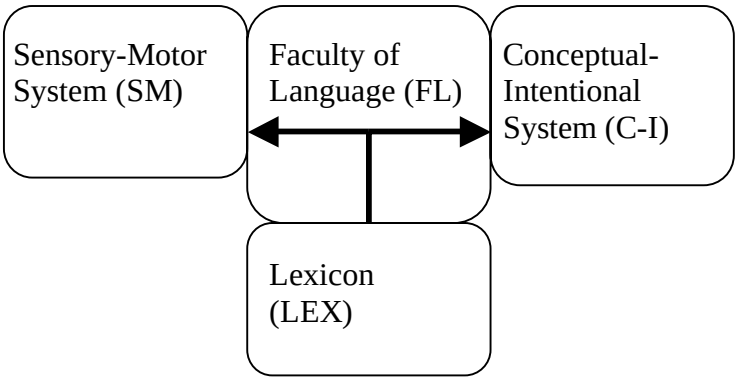
This paper is set up as follows: in section 2 I discuss Chomsky's Strongest Minimalist Thesis (Chomsky 2005) and I present the core of my proposal. In section 3 I discuss the notion of optimality in grammatical architecture and demonstrate why dislocation effects result from phonology-semantics mismatches. In section 4 I discuss the raison-d'etre of formal features, namely that they are required to establish dislocation effects. In section 5 I discuss the way that functional structure consequences of my proposal

w.r.t. L1 acquisition: if formal material is not UG-based, how can it then be acquired and I formulate a hypothesis that says that overt doubling in the language input forms the proper cue to create formal features. In section 5 I address the debate between flexible and rigid syntactic structures and in section 6 I demonstrate that the existence of flexible structures can be proved empirically. I formulate the so-called Flexible Formal Features Hypothesis that states that formal features arise during language acquisition as a result of doubling effects, briefly discuss two phenomena to illustrate the correctness of my hypothesis. In section 7 I re-address the relation between formal features and functional projection and in section 8 I discuss two case studies, Negative Concord and Modal Concord, that seem to confirm the Flexible Formal Feature Hypothesis. In section 9 I motivate the existence of flexible functional structure and in section 10 I argue that the proposals made in this paper allow parameters to be considered not to be innate. Section 11 concludes.

2. The Strongest Minimalist Thesis

Chomsky's SMT says that 'language is an optimal solution to interface conditions that the Faculty of Language (FL) must satisfy' (Chomsky 2005):3. This thesis, tracing back to the philosophical view that language enables human beings to express their thoughts, a view endorsed in the biolinguistic perspective, is implemented in the current perspective on the architecture in the following way: The faculty of language (FL), a mental organ is connected to both the systems that deal with the expression of a sentence and the meaning it conveys, as illustrated in (1) below:

(1)



In the picture in (1) the input for FL exists of a lexical numeration and the output (after separating at Spell-Out) is diverged either to the SM interface or the C-I interface. Consequently, since the output of FL is the input for the SM and C-I systems, the outputs of the derivation should be fully legible for each connected mental component. This amounts to saying that the two interpretational systems impose conditions on the structures that are met at the interface. Such conditions are for instance principles such as Full Interpretation that states that each element of a derivation at the interface may not contain any uninterpretable material. Moreover, since FL is assumed to have no internal levels of interpretation, all other grammatical principles apply at either one of the interfaces. This also holds for Economy conditions. The fact that movement only takes place if strictly necessary follows from Full Interpretation (move in order to delete uninterpretable material) and an Economy condition that also holds at LF (only move if necessary).

Against this background the SMT simply says that language is the simplest possibility to obey all conditions that are imposed on FL by the interfaces. The idea that language is the simplest solution to fulfil these interface conditions leads to the assumption that language is in some way 'perfect.' However, although some conditions are really strong (violating Full Interpretation leads to ungrammaticality), other conditions are less strong. Economy conditions are best understood in the sense that, all other things being equal, comparison between two structures leads to the grammaticality of the more economical structure and ungrammaticality of the less economical one. But, there is nothing that guarantees 'that all other things are being equal.' Crucially, if there are a number of conditions that the SM system imposes on language and a number of conditions that the C-I system imposes, it follows from nothing that these conditions may not be in conflict. On the contrary, it is much more likely to assume that, since the two mental components function autonomously, the two types of conditions are not always compatible. It could in fact very well be the case that if a particular interface condition imposed by the SM-system is maximally satisfied, some other interface condition operating at the C-I interface, cannot be maximally satisfied anymore.

If several interface conditions are indeed in conflict with other conditions, one cannot say anymore that there is only one solution to optimally satisfy interface conditions. In fact it turns out that there are more, equally optimal, solutions to satisfy interface conditions. Hence, I suggest to amend Chomsky's SMT in the following way:

(2) *The Revised Strongest Minimalist Thesis (RSMT):*
Every grammar G forms a maximally optimal solution to legibility conditions at the (different) interfaces.

Note that the main difference between Chomsky's SMT and (2) is that (2), contrary the SMT, invokes grammatical variation. If there are two ways of equally optimally satisfying a pair of conflicting interface conditions, one would expect that some grammars opt for one way and the others for the other. This means that (2) leads to parametric variation. Given that one grammatical 'imperfection' can be explained in this way, the strongest hypothesis is to assume that this holds for each instance of parametric variation:

(3) *The Strongest Parametric Variation Hypothesis (SPVH):*
The Revised Strongest Minimalist Thesis governs the entire range of parametric variation.

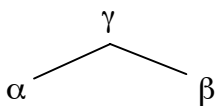
3. Optimal design in the architecture of grammar

In this section I have a closer look at the kind of structures (economy) conditions applying at the two interfaces require different configurations. It will become clear that the semantic conditions (each lexical items should have some semantic content) and phonological economy conditions (keep phonological structure as small as possible) lead to a conflict, and that the only way of repair is the emergence of functional structure.

3.1 Optimal design from the C-I interface perspective

Let us try to enter the mind of a purely semantically biased language engineer. From the semantic perspective the most important requirement on linguistic structure is that it allows for compositional interpretation. This means that if a particular interpretation of an element γ (4), ||γ|| can follow from ||α|| and ||β|| by means of Functional Application (as defined by (Heim and Kratzer 1998) in (5)) there is no reason to assume a different structure, let alone a structure that contains more (functional) material than α and β. Extra structure below γ is only motivated if the interpretation of γ does require so.

(4)



(5) FA: $||\gamma|| = ||\alpha(\beta)|| = ||\alpha||(||\beta||)$

3.2 Optimal design from the SM interface perspective

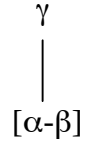
It is a well-known fact about natural language that syntactic structure is not always identical to the easiest structure that compositionality might require. This is due to the fact that language, apart from semantic requirements also needs to satisfy phonological requirements. Take a look at (4) again. Since α and β occupy different structural positions, they should be thought of as different Lexical Items, i.e. different morphological words. The structure is for instance reminiscent of a verb and its object argument. This is not the most economical structure from the perspective of the SM interface however. Let us compare the two structures in (6) and (7). The first corresponds to the structure in (4). The structure in (7) on the other hand is phonologically much more economical. Here there is only one Lexical Items that consists of a root and an affix. The amount of phonological structure is much lower in (7) than in (6).

(6) $[LI_1 + LI_2]$

(7) $[_{LI2}$ Root-$AF_1]$

The problem with the structure in (7) however is that it cannot be interpreted as such. A corresponding structure, as given in (8), cannot be interpreted, since both α and β need to be interpreted in an independent position from which they can take scope. Hence (8) will crash at LF.

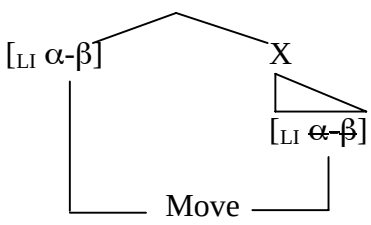
(8)



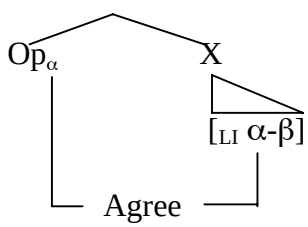
3.3 Dislocation as a result of the mismatch between SM and C-I interface conditions

In order to keep structures like (7) interpretable, the single word [α-β] needs to be matched to two structural positions. This is the moment where dislocation is introduced: either the complex element must be moved to a second position, such that in each position one semantic operator is interpreted, or the element carrying the affix marks the presence of an abstract operator by means of agreement.

(9) a.



b.



The structures in (9) are indeed attested cross-linguistically. (9)a is an instance of movement, where movement itself leads to the expression of two semantic functions, each in a different position. (9)a could for instance be an imperative: in many languages a sentence gets interpreted as having the illocutionary force of an imperative as a result of overt movement from V° to C°. In C° the operator encoding the illocutionary force is interpreted (and not below, since speech act operators are apply on propositions). The verb is interpreted in its base position.

The construction of (9)b is more reminiscent of canonical inflectional marking. Take for instance the expression of past tense by means of the affix *-ed*.

(10) I loved Mary

As is well known, a Past tense operator needs to outscope the entire VP (or even a higher constituent) as operates on events. Since lexical verbs do not raise out of VP in English, it is unlikely to assume that '-ed' is the carrier of past tense itself (see also (Stechow 2002) who provides a series of arguments that for the claim that *-ed* cannot be interpreted as the carrier of the semantics of past tense). This thought is even confirmed by the fact that if a second, overt, past tense is operator, the two elements marking past tense do not yield a cumulative reading. It seems more likely to assume that past tense is induced by an abstract operator, of which the presence is marked by the affix '-ed.' If the past tense operator is overtly present, such as in the case of an adverbial like 'yesterday,' this adverb can be thought of as an overt realisation of the past tense operator. The underlying structure of (10) is then:

(11) $[_{TP}\ Op_{PAST}\ [_{TP}\ I\ [_{VP}\ loved\ Mary]]]$

The cases of dislocation discussed above show that whenever there is one single word, consisting of two elements that correspond to different semantic operators, the modes of interpretation that apply at LF require that such a word is configurationally related to more than one position, triggering a dislocation effect. The cases discussed above are core cases, in which only one position is phonologically visible. In some cases (overt agreement partial movement, etc) both positions are visible. Naturally

these instances of dislocation are accounted for in the same fashion as they are mere varieties of the structures in (9).

4. Formal features

In the previous section I demonstrated how the existence of additional functional structure (creating phonologically invisible positions) follows from the SM desideratum for phonological economy. However, nothing has been said yet about how this extra functional structure arises, it has only been accounted for that it must arise. Two facts play a role. First, it is the case that the structures containing an LI [α-β] without dislocation cannot be interpreted at LF. This means that the sentences in (9) without movement of the LI or an (abstract) operator that is connected to it render ungrammaticality. Second, in order to prevent Look Ahead problems, it must already be lexically encoded that without these instances of dislocation the sentence cannot be interpreted. This is how the existence of formal features comes into being. Shortly, the elements in (9) must carry some information, lexically presented, that is relevant for the syntax only. Such a piece of information must have three properties: it may not be allowed to survive without being connected to (i.e. standing in a particular relation with) a second element; it must indicate with what kind of element it must be combined with; and third it may not contribute to the semantics of the sentence. Given the interface condition of Full Interpretation, i.e. every element present at the interface levels may not contain elements that are not legible, the first and third requirement fall together. If something cannot be interpreted at PF/LF, it is illegible at these levels and thus causes ungrammaticality. Thus the piece of information that the LI must carry, call it X, is something saying that X is uninterpretable and that X must be combined with an element that carries a particular property F. This is the definition of elements standardly known as formal features, in this case [uF].[3]

Now, let us discuss how the existence of uninterpretable features licenses the dislocation structures in (9). Note that nothing new is being introduced here: the main difference between the proposal presented in this paper and standard analyses of dislocation and feature checking is that the reason why feature checking is the trigger of dislocation receives a principal explanation.

[3] A well-known problem regarding formal features is that the syntax, to be more precise a particular state of the derivation, cannot know that that [uF] is not legible at LF. Therefore the notion *unvalued* has been introduced (Chomsky 2001). This amounts however to the same idea: a feature that gets valued during the derivation is gets deleted, a feature that has already been valued within the lexicon is not. In my approach I will stick to the notion uninterpretable. Being uninterpretable here is a purely syntactic property, it encodes information that is only relevant for the course of the derivation. As a result of this it is not legible for the C-I system (and neither for the SM system for that matter), since such information can only be read by the syntactic component. This means that, given that LF is constituted by (one of the) final stages of the derivation, [uF]'s may be deleted at LF. Obviously nothing is wrong with that. If the final syntactic operation is one that causes deletion of the final [uF] in the sentence no problems should arise for the grammaticality of the sentence (there is no material present that is illegible to the C-I and SM systems).

In the canonical case of movement, the result of the syntactic operation Move (Internal Merge), a LI of the form [α-β] is moved to a second position, such that one property is interpreted in the higher copy,, one in the lower copy. Le us take an imperative verb, V_{imp}. V_{imp} contains two pieces of semantic information: its verbal (i.e. predicative) property and its imperative (i.e. speech act property). In order to avoid interpretation at the same time, the sentence must be configured such that without any triggering of dislocation, the sentence would crash. Note that without the imperative morphology the sentence would not crash.[4] Hence there must be a feature [uIMP] at hand, which is uninterpretable, and requires a speech act operator (encoded by the interpretable feature [iIMP]). Now there are two ways to instantiate this: either the verb carries [iIMP] or it carries [uIMP]. In the first case the target of movement must carry [uIMP], in the second case [iIMP]. This leads to the respective configurations in (12).

(12) a. $[[X_{[uIMP]}\ V_{[iIMP]i}\ ]\ [\ldots V_{[iIMP]i}\ \ldots]]$
b. $[[X_{[iIMP]}\ V_{[uIMP]i}\ ]\ [\ldots V_{[uIMP]i}\ \ldots]]$

In (12)a the feature [uIMP] on X needs to be deleted; in (12)b [uIMP] on V gets deleted. In both cases, for some reason, the deletion operation could not act on a distance, and thus the element had to move. After deletion the way is open for the speech act operator to take scope from the highest position, and for the verb from the lowest position. How this actually take place depends on different theoretical preferences (copy theory of movement, trace theory, etc.) and I will not discuss these implementations. It is crucial that in order to create additional functional structure (required for proper interpretation), some [uF] is needed, otherwise the sentence would crash at LF, because if V carried [uIMP] Full Interpretation would render ungrammaticality, and if V carried [iIMP] the two semantic operations would be interpreted within the same position.

Let us now discuss the case of Agree. Agreement comes about in two forms: overt and covert agreement. Overt agreement manifests a connection between two elements in the sentence. Take for instance subject-verb agreement. In (13) the verb agrees with the second person singular subject.

(13) Du komm-st German
You come.2sg
'You come'

The properties of *-st* can be summarised as follows: (i) *–st* may not survive on its own; (ii) *–st* indicates with what kind of element it must be combined (with a second person singular subject); and (iii) it may not contribute to the semantics of the sentence (the meaning of the sentence 'you come' can be derived completely from the contribution of the subject and the verbal stem). The reader will have noticed that these three properties are exactly the properties of uninterpretable formal features as

[4] Unless other requirements require verbal movement as well, as is the case in V2 languages.

discussed above. Hence, the finite verb *kommst* must carry an uninterpretable formal feature [u2SG].[5]
The question why the verb has these inflectional markers carrying these uninterpretable formal features is not clear in this case. Nothing would go wrong in a language if there were not any inflectional subject marking on the verb, as long as the subject is overtly realised. This is however a special peculiarity of German. The example above shows that given the agreement relation between the subject and the verb, an uninterpretable feature must be involved.

There are however cases where the inflectional marking is absolutely necessary. Take for instance the expression of past tense, in which it was argued for that the past tense semantics were induced from an abstract operator. The original example in (11) is repeated below.

(14) $[_{TP}\ Op_{PAST}\ [_{TP}\ I\ [_{VP}\ loved\ Mary]]]$

The question is how the sentence 'I loved Mary' can be assigned a structure as in (14). Why is it impossible that the structure does not require the presence of an abstract operator Op_{PAST} and has a reading 'I love Mary'?
Apparently the expression *loved* must have a requirement that it cannot survive without Op_{PAST}, it indicates that it is a past tense operator that is required (not for instance a future tense operator) and it *loved* itself does not contribute to the temporal semantics, since the abstract operator is already responsible for that. In other words, the properties of *loved* are by definition the same properties of an uninterpretable feature. Apparently *loved* is equipped with an uninterpretable feature [uPAST] that requires the presence of a past tense operator. This explains why (14) must be analysed like this. With no abstract operator the sentence would be ungrammatical. Abstracting away from the past tense operator, one can postulate the following phonological economy rule:[6]

(15) An abstract operator may be assumed iff this prevents the sentence from crashing.

To conclude, what have seen thus far, is that in each instance of dislocation (resulting from the operations Move and Agree) formal features have played a crucial rule. In

[5] I am not making any specific claim whether this should be one feature [u2SG] or two features [u2] and [uSG]. I uses the first feature in the text, but the proposal is compatible with the second perspective as well.

[6] The condition in (15) accounts for the fact that sentences containing an overt element with a [uF] but no overt element with a [uF] do get their correct interpretation. In minimalist terms nothing is wrong with this assumption. If the abstract negative operator is a lexical element, it can simply be part of the numeration and the [uF] can be checked. (15) also says that an abstract operator carrying [iF] can only be assumed to be present in the sentence if it contains an overt element [uF]. In minimalist terms this is strange, since it functions somehow as a filter on derivations. On the other hand, it is known that phonological zero material needs to have some kind of licensing, which may be imposed as a result of a general (interface) condition of licensing abstract material (similar to constraints on e.g. ellipsis).

fact, it has been demonstrated that by definition without formal features dislocation could not have taken place. Dislocation, as shown above, is a result of a semantics-phonology mismatch, and formal features must exist in order to establish the required dislocation effects. This means that a second imperfection in Grammar (formal features), similar to parametric variation, follows from conflicting interface (economy) conditions.
Note that this view on formal features shows that formal features are only needed if a the marking of a particular semantic operator is spelled out on another element.

5. Functional structure

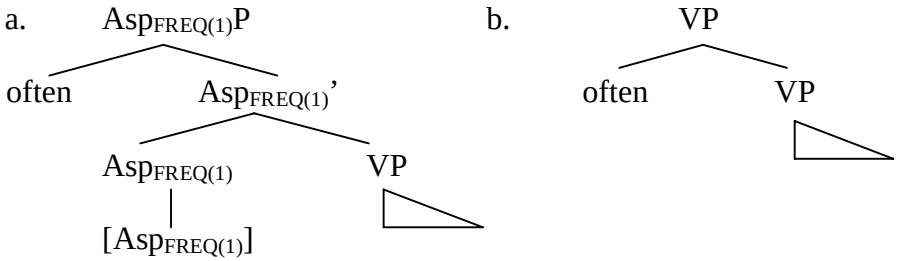
The third imperfection to be discussed in this paper concerns the functional structure. Several scholars have argued that (the amount of) functional structure is a property of UG. One such proposal has been Cinque's adverbial hierarchy (Cinque 1999), who provides a template for the adverbial distribution, as illustrated in (16). Since Cinque's proposal is one of the most radical ones in terms of fixed templates, it is on of the most interesting ones to discuss here.

(16) [*frankly* $\text{Mood}_{\text{speech act}}$ [*fortunately* $\text{Mood}_{\text{evaluative}}$ [*allegedly* $\text{Mood}_{\text{evidential}}$ [*probably* $\text{Mod}_{\text{epistemic}}$ [*once* T(Past) [*then* T(Future) [*perhaps* $\text{Mood}_{\text{irrealis}}$ [*necessarily* $\text{Mod}_{\text{necessity}}$ [*Possibly* $\text{Mod}_{\text{possibility}}$ [*usually* $\text{Asp}_{\text{habitual}}$ [*again* $\text{Asp}_{\text{repetitive(I)}}$ [*often* $\text{Asp}_{\text{frequentative(I)}}$ [*intentionally* $\text{Mod}_{\text{volitional}}$ [*quickly* $\text{Asp}_{\text{celerative(I)}}$ [*already* T(Anterior) [*no longer* $\text{Asp}_{\text{terminative}}$ [*still* $\text{Asp}_{\text{continuative}}$ [*always* $\text{Asp}_{\text{perfect(?)}}$ [*just* $\text{Asp}_{\text{retrospective}}$ [*soon* $\text{Asp}_{\text{proximative}}$ [*briefly* $\text{Asp}_{\text{durative}}$ [C*haracteristically(?)* $\text{Asp}_{\text{generic/progressive}}$ [*almost* $\text{Asp}_{\text{prospective}}$ [*completely* $\text{Asp}_{\text{SgCompletive(I)}}$ [*tutto* $\text{Asp}_{\text{PlCompletive}}$ [*well* Voice [*fast/early* $\text{Asp}_{\text{celerative(II)}}$ [*again* $\text{Asp}_{\text{repetative(II)}}$ [*often* $\text{Asp}_{\text{frequentative(II)}}$ [*completely* Asp]]]]]]]]]]]]]]]]]]]]]]]]]]]]]

Cinque's analysis is based in two different claims: (i) he argues that selectional hierarchies are part of UG; (ii) he argues that each adverb occupies the specifier position of a functional projection that exhibits an X-Bar skeleton. In this paper I will not discuss the syntactic nature of hierarchies. Rather, I want to address the claim that each class of adverbials universally requires a functional projection of its own.

Opposed to Cinque's view that the set of functional projections is uniform across languages is the view that the set of available functional projections is flexible (Ackema, Neeleman et al. 1993; Weerman and Neeleman 1997; Koeneman 2000; Zeijlstra 2006). In a flexible system the existence of a particular functional projection in one language does not imply the existence of such a projection in another language. One can compare the two perspectives w.r.t. the English adverb *often*. Under Cinque's analysis *often* should be located in the specifier position of $\text{Asp}_{\text{frequentative(I)}}\text{P}$; under a flexible approach it can be taken to be an adjunct of a vP. The two options are illustrated below:

(17) a. $Asp_{FREQ(1)}P$ b. VP



The question to be asked now is how these two structures can be distinguished empirically. Cinque suggests that since each two adverbs in principle allow a verbal participle between them. This indicates that in between every two adverbs an empty head position should be available.

(18) Da allora, non hanno <rimesso> di solito <rimesso> mica <rimesso> più <rimesso> sempre <rimesso> completamente <rimesso> totto bene in ordine
Since then NEG have.3PL <put> usually <put> NEG <put> any longer <put> completely <put> always <completely> everything in order
'Since then, they haven't usually not any longer always put everything well in order' (Cinque 1999): 45

However, this argument does not show that there are as many head positions available as there are adverbial projections. If one would allow multiple specifiers, only two verbal heads can exhibit the entire distribution in (18). One position is the head position that the verb occupies and of which the preverbal adverbials are specifiers; the other position may be left empty and hosts all postverbal adverbials as specifiers.

(19) a. $[_X$ ADV1 X° $[_Y$ ADV2 Y° $[_Z$ ADV3 Z° $[_U$ ADV4 Y°]]]]
b. $[_X$ ADV1 ADV2 X° $[_Y$ ADV3 ADV4 Y°]]

Hence, the only way to distinguish between the two possible structures in (17) lies in the fact that in (17)a a particular formal feature $[Asp_{FREQ(1)}]$ must be available to project $Asp_{FREQ(1)}P$. In (17)b such a feature however does not have to be present. If it can be shown that English lacks particular formal features, this forms evidence for the existence of structures as in (17)b. If however it cannot be proven that such features are absent, no empirical distinction can be made between (17)a and (17)b. Hence the question which structure in (17) is correct for a language like English depends completely on the question whether the set of formal features is universal (UG-based) or based on L1 acquisition (and thus empty at the initial stage). Since the absence of formal features is decisive, the most plausible track to follow is to hypothesize that that the set of formal features is non-universal, and therefore empty in UG. The reason for this is purely methodological: a hypothesis that takes all formal features to be part of UG, predicts the availability of both (17)a and (17)b, whereas the non-universal approach only allows (17)b. If the predictions that the flexible hypothesis makes are correct, the correct structure for English must be the on in (17)b.

The reader may already have noticed that the discussion above implicitly assumed that functional projection is reserved to formal features, i.e. only formal features are allowed to project. This is a standard assumption being made in the literature, tracing back to (Giorgi and Pianesi 1997) in their Feature Scattering Principle:

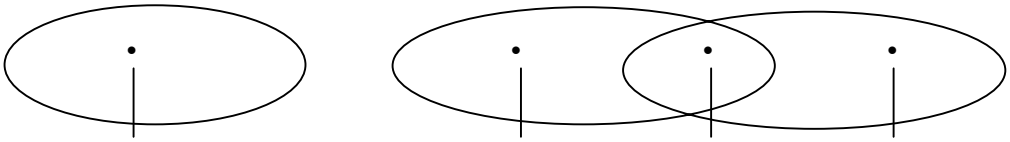
(20) *Feature Scattering Principle*
Every feature can head a projection

Although it is highly likely that syntactic operations are only accessible to syntactic material (and thus only to formal features and not to semantic features), the fact that only formal features may project needs to be explained. I will do so after the discussion on flexible features.

6. The Flexible Formal Features Hypothesis

Grammatical features are thought to constitute three categories: phonological features, formal features and semantic features (Chomsky 1995). Phonological features are legible at the SM interface and semantic features are legible at the C-I interface. Formal features come about in two kinds: interpretable and uninterpretable formal features. Interpretable formal features are also legible at the C-I interface and are therefore also members of the set of semantic features. The set of formal features and semantic features thus intersect. Uninterpretable formal features are neither legible at one of the interfaces and thus need to be deleted during the derivation, i.e. they need to stand in a proper agreement relation with a formal feature in order to prevent the derivation from crashing at the interfaces.

(21) Phonological features Formal features Semantic features



Suppose that the set of formal features is empty in the initial stage. L1 learners must then acquire which features are formal(ised) and which features are not. In other words, an L1 learner needs to have positive evidence that certain lexical elements carry formal features. Let us look again at the properties of formal features. Since they come about in two kinds, we need to discuss the properties of both interpretable and uninterpretable features in order to see what property can act as a proper cue (in the sense of (Lightfoot 1999)) during language acquisition.

Formal features have two different properties: they are interpretable at LF and they can check uninterpretable features.[7] However, they are not recognisable as such for a language learner. Their first property is not decisive, since semantically speaking, formal features are undistinguishable from non-formal semantic features ([F]'s), as shown in (22).

(22) $||X_{[F]}|| = ||X_{[iF]}||$

[7] Feature checking here is used since it is the common term for the process that iss going on here. Checking means that a licensing requirement of an uninterpretable feature is fulfilled, thus leading no longer to ungrammaticality.

Secondly, the fact that interpretable formal features are required to check uninterpretable features cannot trigger the acquisition of formal features either. A formal feature [iF] can occur without any problems in a sentence without any [uF]'s. Only a [uF] cannot survive without the presence of a proper [iF]. However, this can only be acquired on the basis of negative evidence, which is virtually absent during L1 acquisition. Hence the properties of interpretable formal features can never lead to the acquisition of formal features as such.

Uninterpretable features on the other hand do give rise to cues. Let us look at the properties of uninterpretable features again, using the insights of section 4. Uninterpretable formal features are semantically vacuous. Moreover they require the presence of an interpretable formal feature and they give rise to dislocation effects, thus triggering syntactic operations such as Move and Agree. All these properties are recognisable for a language learner. In fact, they all reduce do so called doubling effects. Let us define doubling in the following way.

(23) F exhibits a doubling effect iff the presence of a semantic operator Op_F is manifested overtly by more than one element in the morphosyntax

Now, let us look at possible instances of doubling effects. Doubling is the case in overt agreement. Take the past tense example:

(24) John loved Mary

Since *-ed* cannot be the realisation of Op_{PAST}, since the semantics of the sentence must correspond to the situation in which (24) is true. Hence, there must be a second (abstract) Op_{PAST} available in the sentence. Consequently, since the sentence soes not contain two Op_{PAST}'s, -ed must be semantically vacuous and thus carry a feature [uPAST] that triggers the presence of the abstract Op_{PAST}. As a result the feature corresponding to past tense is no longer a semantic feature [PAST], but a formal feature [i/uPAST]. Past tense operators, such as the abstract Op_{PAST} are then assigned a formal feature [iPAST].

Movement is another instance of doubling. Take for instance Wh elements. In languages such as English they require movement to Spec,CP. That means that the sentence contains two copies of a Wh term. Hence Wh exhibits a doubling effect. Moreover, as I discussed befor, movement is always triggered by an uninterpretable feature, so if Wh triggers movement, at least at some position the structure the presence of an uninterpretable [uWh] feature is required. Then it follows that the elements that directly contribute the semantics of the interrogative contain a feature [iWh]

Note that this does not only hold for covert movement. LF movement (such as Quantifier Raising) also exhibits a doubling effect, since the language learner knows that the position that a particular quantifier is interpreted (e.g. a quantifier) does not correspond to its position at the surface. Again this is a doubling effect. The language learner knows again that in such cases the presence of an uninterpretable feature (say [uQ]) is needed in order to enable the movement operation. The quantifiers themselves then carry a feature [iQ] which enables them to be 'read' as quantifiers ar LF.

Hence the presence of formal features can be acquired by L1 learners, since the presence of uninterpretable features can be acquired. The presence of formal features then immediately follows. Hence the following hypothesis can be formulated.

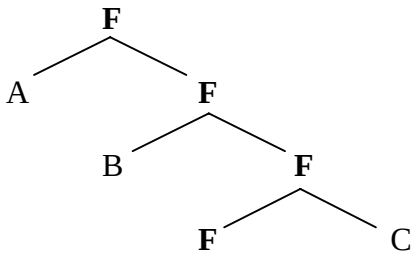
(25) The Flexible Formal Feature (FFF) Hypothesis:[8]
A language has a formal feature [i/uF] iff it exhibits doubling effects with respect to F.

This means that a language only has a formal feature [i/uPAST] if it exhibits doubling effects with respect to past tense; a language exhibits only a formal feature [i/uNEG] if it exhibits doubling effects w.r.t. negation, etc.

7. Formal features and projection

As promised, the fact that only formal features are allowed to project needs to be addressed. This follows from the observations made above: formal features may give rise to doubling effects; semantic features do not (if they would they had to be reanalysed as formal features). Let us see now what the consequences are for formal features and projection. Let us take the following abstract functional projection:

(26)



The reader may already have observed that projecting F exhibits a doubling effect w.r.t. F: a functional projection consists of multiple layers, each assigned a label that correspond to the head. But obviously not all instances of F are semantic operators: a NegP for instance does not contain three or four negative operators; it contains only one. And (17)a only contains one adverbial, *often*, and not four elements to be interpreted as $Op_{AspFREQ(1)}$ (candidates being the specifier *often*, $Asp_{FREQ(1)}P$, $Asp_{FREQ(1)}$' and $Asp_{FREQ(1)}°$). The latter scenarios however would have been the case if semantic features were allowed to project. This explains why projection is restricted to formal features and not to semantic features.

The argument that functional projection is only allowed addresses some problems that have been mentioned in the literature regarding the necessity of functional projection, especially the apparent redundancy of feature doubling that is the case with spec-head agreement (cf. (Brody 2000; Starke 2005)). Given the fact each instance of merge needs a label and that features provide such label, redundancy occurs immediately. However, this redundancy is not problematic if it is motivated on independent grounds. Clearly, that is the case here. Formal features are needed in order to enable dislocation so that conflicting interface conditions can be fulfilled in multiple optimal

[8] This hypothesis has first been introduced in (Zeijlstra 2006) in a slightly different form

ways. Now if these formal features create functional projections as a by-product, this instance of redundancy is not any longer problematic.

Thus far the FFF hypothesis has not been proven. It has only been demonstrated that formal features can be acquired, since uninterpretable features can be acquired (and therefore their interpretable counterparts) and that functional projection is subject to formal features. But this makes the FFF hypothesis empirically testable. If it can be shown that a functional F° is present cross-linguistically if apart from projection F exhibits doubling effects, the FFF hypothesis is confirmed. If F°'s may be attested without giving rise to any doubling effect, except for projection, than the FFF hypothesis must be rejected, and the set of formal features is then likely to be part of UG instead of resulting from L1 acquisition.

A few words need to be said about the distinction between phrases and heads. In current minimalist reasoning, lexical items are not marked for head or phrasal status. This does however not imply that there is no difference between heads and phrases. The only difference is that X°'s, X's and XP's should no longer be thought of as syntactic primitives, but as derived notions. As is well known heads can be rephrased as having a property [-Max, +Min] and phrases as [+Max, - Min] (see (Grohmann, Hornstein et al. 2005) for discussion). Consequently, relativized minimality effects (Rizzi 1990) can still be attested empirically (see (Starke 2001)). Hence standard diagnostics for the distinction between heads and non-heads still apply. Heads, for instance, do not allow movement of other heads across them (following Travis' Head Movement Constraint (Travis 1984)), and adjunction may only take place between two elements with identical syntactic status (heads adjoin to heads, phrases to phrases), as has been shown by many scholars (take (Merchant 2001) as an example).

The fact that heads are empirically detectible, and the prediction that the FFF hypothesis makes, namely that formal features [i/uF] are acquired as result of doubling effects and that only formal features are allowed to project, gives rise to the following templates for grammatical universals:

(27) F° → Overt doubling effects w.r.t. F

This template for typological universals can easily be explained. The FFF hypothesis accounted for the fact that it takes doubling effects w.r.t. F for the L1 learner to trigger the acquisition of formal features [i/uF]. Only formal features [i/uF] are allowed to project (given that projection is an instance of doubling). Hence, if a particular F° os overtly realised, it must have been analysed as carrying a formal feature [i/uF] and thus there have the have been doubling effects w.r.t. F in the language input.

Note that this template is uni-directional. It does not say that whenever there are formal features [i/uF] there must be an overt head F°. It only says that if such a head is there, there must be doubling effects too.

In the next section I will highlight one functional category, negation, and show that the prediction that (27) makes is correct. After that I will elaborate a bit more on other functional categories to see whether they fit in the template as well. Needless to say, one must evaluate (27) for a series of phenomena thus giving rise to a research

program rather than a few analyses. However, the results thus far seem promising and indicate that the flexible approach is on the right track.

8. Concord

In order to evaluate the FFF Hypothesis one must find cases where a particular semantic operator Op_F is marked at a particular head position F° and evaluate whether or not doubling effects w.r.t. F arise. If not, then this is a counterexample against the FFF Gypothesis. The FFF Hypothesis thus makes some very clear predictions about the relation between the presence of a functional head F° and so-called Concord. Concord is a phenomenon where not every morpho-syntactic instance of as particular category corresponds to a semantic operator, a clear case of a doubling effect. Therefore it is useful to study the relation between functional heads and concord in order test the FFF Hypothesis. In this section I first discuss the case of Negative Concord, and I demonstrate the FFF hypothesis makes correct predictions about its cross-linguistic distribution. After that I address a less known instance of concord, namely Modal Concord.

8.1 Negative Concord

Languages differ w.r.t. the interpretation of sentences consisting of multiple negative elements. In Double Negation (DN) languages, such as Dutch, each morpho-syntactically negative element corresponds to a semantic negation. In Negative Concord (NC) languages, like Czech, multiple negative elements correspond to one semantic negation.[9] Some examples are given in (28)-(29).

(28) a. Jan ziet *niemand* — Dutch
Jan sees n-body
'Jan doesn't see anybody'
b. *Niemand* zegt *niets*
N-body says n-thing
'Nobody says nothing'

(29) a. Milan *(*ne*)vidi *nikoho* — Czech
Milan NEG.saw n-body
'Milan didn't see anybody'
b. Dnes *(*ne*)volá *nikdo*
Today NEG.calls n-body
'Today nobody calls'

[9] Some languages do not exhibit the possibility to mark negation on more than one element. In those languages (many languages are South-East Asian, such as Canton Chinese) only one negative marker can be used. Indefinites are marked for negation by the fact that they are Negative Polarity Items (NPIs), such as English *anything* or *anybody*. In those languages negation is part of the verbal inflectional morphology, and thus supporting the FFF hypothesis (if morphemes are assumed to have head status). The behaviour of these languages does not provide any counterexample for the FFF Hypothesis. In this section, I leave out the discussion on these languages and focus on languages that either exhibit NC or DN.

c. Dnes *nikdo* *(*ne*)volá
Today n-body NEG.calls
'Today nobody calls'

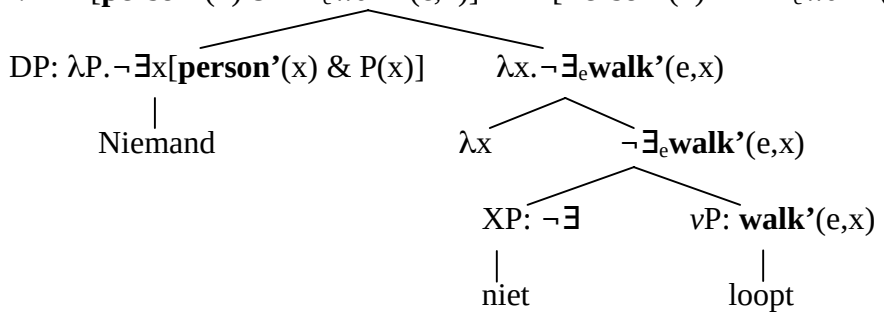
d. Nikdo *(*ne*)vidi *nikoho*
N-body NEG.saw n-body
'Nobody saw anybody'

The sentences in (28) show that in languages such as Dutch each morpho-syntactic instance of negation corresponds to a semantic negation (that is interpreted in its surface position). According to the FFF Hypothesis no cue in the language input is available for the language learner to assume that negative elements should be analysed as having a negative feature [i/uNEG]. Negative elements can only carry a semantic feature [NEG]. The semantics follow straightforwardly from the lexical semantics of the negative elements, and syntactic operations, such as Move/Agree, are not involved, as demonstrated below.

(30) *Niemand* loopt *niet* Dutch
Nobody walks neg
'Nobody doesn't walk' = 'Everybody walks'

(31) $[_{TP}$ *Niemand*$_i$ $[_{vP}$ *niet* $[_{vP}$ t_i loopt]]]

(32) TP: $\neg\exists x$[**person'**(x) & $\neg\exists_e$**walk'**(e,x)] = $\forall x$[**Person'**(x) → $\neg\exists_e$**walk'**(e,x)]



Things are different in a language like Czech however. The examples in (29) show that the presence of a negative operator (which is semantically negative) is marked by two elements. Thus, in each example not every negative element can correspond to a semantic negation. This forms a clear cue that (according to the FFF Hypothesis) negation constitutes a formal feature in Czech. The question is now which kind of feature ([iNEG]/[uNEG]) should be assigned to which element. Since multiple n-words may establish NC relations, n-words cannot carry a feature [iNEG], since otherwise the (29)d should have two negations in its semantics. However, the negative marker cannot carry an [iNEG] feature either, since it does not always c-command all other negative material. If it were the case that *ne* carried an [iNEG] feature, (29)c would have a reading in which the indefinite outscopes the negation.[10]

[10] See (Giannakidou 2000) for an analysis that takes n-words to be non-negative universal quantifiers that raise across negation. Under such approach Czech *ne* can be said to carry an [iNEG] feature. Although I think that the arguments given in (Zeijlstra 2004) demonstrate some problems with this analysis, it is not incompatible with the proposal formulated in this paper. Crucial is that the language learner decides

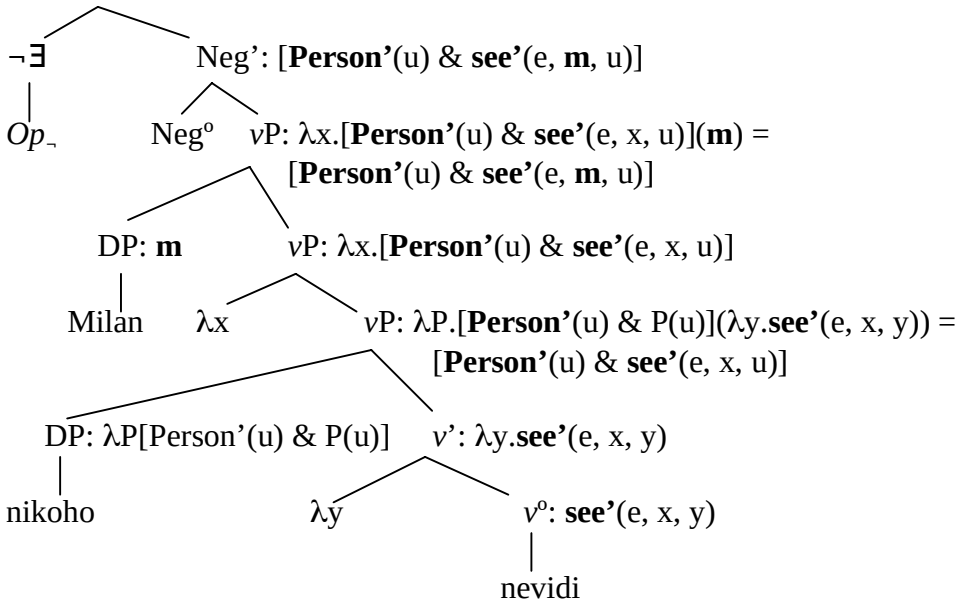
Hence I argue, following (Zeijlstra 2004), that both negative markers and n-words in Czech carry [uNEG] and that an abstract operator $Op_{\neg}$ is responsible for the negative semantics. Assuming that n-words are indefinites, with a semantics as in (33) and the assuming, as is standardly done, that operators may bind free variables, the syntax and semantics of Czech negation are than analysed as in (35)-(36).

(33) ||n-Q|| = λP.[Q(x) & P(x)], where Q ∈ {**Person'**, **Thing'**, **Place'**, …}

(34) Milan *ne*vidi *nikoho* Czech
Milan NEG.sees n-body
'Milan doesn't see anybody'

(35) $[_{NegP}\ Op_{\neg[iNEG]}\ [_{Neg^o}\ ne\text{vidi}_{[\sout{uNEG}]i}\ [_{vP}$ Milan $nikoho_{[\sout{uNEG}]}\ t_i]]]$

(36) NegP: $\neg\exists_{u,e}$[**Person'**(u) & **see'**(e, **m**, u)]



The assumption that NC is an instance of syntactic agreement follows straightforwardly from the FFF Hypothesis, but is not uncontroversial. It has been assumed prior by (Zeijlstra 2004; Zeijlstra 2006), following (Ladusaw 1992; Brown 1999). However, other analyses have been formulated as well. Unfortunately space limits prevent me here from doing justice to other accounts of NC. The reader is referred to (Zeijlstra 2004; Zeijlstra 2006) for a critical evaluation of other theories (Haegeman 1995; Haegeman and Zanuttini 1996; de Swart and Sag 2002) who propose an analysis in terms of polyadic quantification, (Ladusaw 1992; Giannakidou 2000) who account for NC in terms of NPI licensing or (Herburger 2001) for an analysis in terms of lexical ambiguity) and a discussion of problems that these theories have been facing.

The syntactic agreement analysis of NC is on the other hand supported by the fact that NC obeys all conditions that grammatical/syntactic dependencies require (cf. (Koster 1987; Neeleman and Van de Koot 2002)):

---

that not every instance of morpho-syntactic negation corresponds to a semantic negation.

(37) a. Each dependent must take an antecedent
b. The antecedent must be in a c-commanding position
c. The antecedent must be sufficiently close to the dependent
d. Each dependent must take a unique antecedent
e. An antecedent may be linked to multiple dependents

Ad (a): each element carrying [uNEG] must take an element carrying [iNEG], otherwise the sentence would lead to ungrammaticality. Ad (b): the element carrying [iNEG] must be in a c-commanding position: a negative operator must be the highest element in negative chain. Note that this requires agreement between a higher [iF] and a lower [uF], as has been suggested by (Adger 2003). Ad (c): the element carrying [iNEG] must stand in a proper locality relation with an element carrying [uNEG]. NC obeys standard locality constraints, such as island effects and clause boundedness (Giannakidou 2000; Zeijlstra 2004). Ad (d): each element carrying [uNEG] must take a unique element carrying [iNEG]. It may not establish an Agree relation with multiple (abstract) operators. Ad (e): An element carrying [iNEG] may be linked to multiple elements carrying [uNEG]. This is also characteristic for NC, since multiple negative elements may participate in an NC relation. Note that this requires multiple Agree in the sense of (Haraiwa 2000; Haraiwa 2001; Ura 1996).

Given this support that NC should indeed be thought of as an instance of syntactic agreement, the prediction that the FFF Hypothesis makes is that DN languages lack formal negative features, whereas NC languages do not. Restricting myself to these two classes of languages, the FFF Hypothesis predicts that only in NC languages a negative feature may project. In DN languages there is no formal negative feature, and only formal features may project. Then it must also follow that whenever one attests an overt negative head, a negative marker Neg°, this language must exhibit NC as well. The prediction is formalised below.

(38) Overt Neg° → NC

This prediction is born out. In (Zeijlstra 2004) it is shown on the basis of 25 diachronic and 267 dialectal varieties of Dutch and a sample of 50 other languages that every language/variety that exhibits an overt negative head is also an NC language/variety.[11] The strong uni-directional correspondence between the syntactic status of the negative marker and the occurrence of NC forms strong evidence for the FFF hypothesis.

[11] The only possible exception known against this observation is standard English. Standard English, being a non-NC language, allows for the negative marker *n't*, which behaves like a negative head. Assumably, this is related to the fact English is on its way of transforming itself into an NC language cf. (Zeijlstra 2004; Anderwalt 2002). Then, English is actually an NC language in disguise. DN effects may then be analysed as the result of language conservatism during the transit period, which are known to play a delaying role during language change (Weiss 2002).

8.2 Modal Concord

Obviously one phenomenon is too few to form firm evidence in favour of the FFF Hypothesis, although the data about negation are very strong. Hence, I address a one second case: Modal Concord.

It is well known that modal verbs in English form a separate class ((Pollock 1989) amongst many others). This follows from facts as in (39)

(39) a. Modal verbs may not be infinitival:
*To can swim is useful
b. Modal verbs cannot iterate
*He shall must do it
c. Modal verbs do not take DP complements:
*I shall you a penny
d. Modal verbs may precede *not*:
If I {*gave/shall} not …
e. Modal verbs move to C° under inversion:
How many languages {*<speak>/can} you <speak>

A consequence of the FFF Hypothesis would be that multiple modals (may) give rise to concord readings. It has been shown for English that they do indeed (Halliday 1970; Lyons 1977; Geurts and Huitink 2006). The readings in (40) do not have an iteration of two modal operators in their semantics. Note that it must hold that the modal types (epistemic, deontic, …) and the quantificational force (existential, universal) must be identical.

(40) a. You may possibly have read my little monograph upon the subject[12]
'The speaker things that it is possible that you read his little monograph '
b. Power carts must mandatorily be used on cart paths where provided
'It is obligatory that power cats are used on cart paths where provided'

Adverbials are most likely carriers of modal force. This follows for instance from the fact that they narrow down the meaning of the modal auxiliary. Then modal auxiliaries themselves must carry uninterpretable modal features of a particular kind. May would then be analysed as carrying a feature [uMOD-EPIS-∃]. If a modal adverb is present, it can check the auxiliary's feature; if it is not present, an abstract modal operator will license it.

(41) [$_{IP}$ You [$_{ModP}$ possibly$_{[iMOD\text{-}EPIS\text{-}\exists]}$ [$_{Mod°}$ may$_{[uMOD\text{-}EPIS\text{-}\exists]}$] [$_{VP}$ …]]]

(42) [$_{IP}$ You [$_{ModP}$ *Op*$_{MOD\text{-}DEON\text{-}\exists}$ may$_{[uMOD\text{-}DEON\text{-}\exists]}$ [$_{VP}$ read]]]

[12] Data taken from (Geurts & Huitink 2006)

The idea that modal auxiliaries are semantically vacuous also sheds some light on an old problem in the study of modality. Why is the scopal order in (43) the reverse of the surface order? If the modal auxiliary is not the carrier of the modal semantics, the abstract operator can be said to occupy a position lower than NegP, and the auxiliary raises to Neg°, as is shown in (44).

(43) You cannot read
'It is not possible for you to read'
¬ > CAN

(44) [You [$_{NegP}$ can$_i$-not [$_{ModP}$ Op$_{MOD}$ t$_i$ [$_{VP}$ read]]]]

If modal auxiliaries constitute functional projections of their own, then Modal Concord is expected to be available. So far this prediction is born out. Languages, such as Dutch, German, English and Hungarian exhibit modal Concord; Russian, that does not have any modal auxiliaries that are a designated head, lacks Modal Concord either.

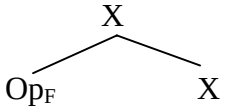
8.3 Conclusion

The strong case of Negative Concord in combination with the Modal Concord data show that the FFF Hypothesis seems to make correct predictions. Of course, the evidence is not complete as of yet. Much research needs to be done in order to make definite claims. However, at this moment the evidence speaks much more in favour of the FFF Hypothesis than against it. I therefore tacitly conclude that the FFF Hypothesis is correct and that the set of formal features is not part of UG. This provides support for the flexible syntax approach that takes functional structure to be syntactically flexible.

9. Functional structure revisited

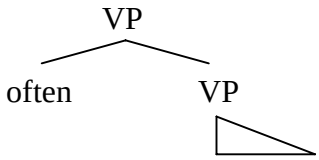
Now that the balance has turned in the direction of flexible functional structure, the question immediately arises why must there be functional structure in the first place. In other words, why do some languages have NegP's, ModP's AspP's, etc.? The answer to this question lies within the fact that each language needs to have some way to express a particular semantic operator. As has been discussed before, the fact that interface conditions may be conflicting leads to equally optimal strategies to express a semantic operator. From the C-I perspective, a structure as is required.

(45)



In this case no extra functional structure is required. Functional Application can be applied and the interpretation of the highest X follows directly from Op$_F$ and the lower X. Since F does not give rise to doubling effects that language will not contain any formal features [i/uF] and F can thus not project. This structure is reminiscent of (17)b, repeated below for convenience.

(46)



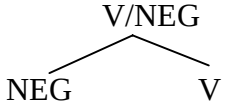
However, as explained before, SM preferences prefer spell-out of multiple elements on the same node. This requires dislocation effects and these dislocation effects require the presence of formal features. Given the (assumed) correctness of the FFF Hypothesis, these formal features are only acquired as a result of doubling effects, and doubling effects cover all dislocation effects. Hence if the SM-biased strategy for expressing a semantic operator requires additional structure, during L1 acquisition the relevant features will be formalised (i.e. analysed as a formal feature). Given that formal features may project, the dislocation effect required for the expression of semantic operators may lead to additional functional structure that is hosted by the required formal feature. Let me illustrate this with an example. The expression of negation can either be realised without formal features (DN languages) or with formal features (NC languages). Both expression strategies are equally economical, which explains why both are attested in natural language. If a negative marker is used to express the (presence of a) negative operator, this requires extra structure, resulting from merger with the negative marker. This is both the case in Yiddish and in Italian.

(47) Ikh hob *nit* gezen *keyn* moyz — Yiddish
I have beg seen n- mice
'I haven't seen any mice'

(48) *Non* ha telefonato a *nessuno* — Italian
Neg has.3sg called to n-body
'He didn't call anybody'

Let us assume for the sake of the argument that both negative markers carry [iNEG] and both n-words [uNEG]. Then the presence of [uNEG] requires merger of VP with the operator. VP is projected by V. Now nothing a priori determines what the label of merger should be. Both V and [iNEG] are proper candidates:

(49)



Hence, languages vary w.r.t. the element that projects in these cases. Italian *non* projects, and therefore exhibits X° behaviour, yielding a NegP. Bavarian on the other hand does not have its negative marker project and therefore it is a specifier or adjunct of VP. The fact that different functional projections are available, although this is not strictly necessary, simply follows because nothing prevents it. What grammar does is requiring additional structure, but does not impose any restriction on the label of the new structures. Hence, languages will vary w.r.t these labels.

10. Consequences for parameters

Thus far, I have addressed three 'imperfections' in natural language: cross-linguistic variation, redundancy and flexible functional structure. They all have received an explanation. I have argued that cross-linguistic variation arises as a result of conflicting interface conditions. Furthermore I have argued that formal features are required to enable dislocation effects that an SM-based expressing strategy for semantic operators. Finally I have argued that flexible functional structure is a result of the fact that only formal features may project and that those formal features are only acquired if necessary. By means of reasoning along these lines all grammatical differences between languages seem to follow from the way that a semantic operator can be expressed in different equally optimal ways. This has been addressed by the *Strongest Parametric Variation Hypothesis (SPVH)*, repeated below

(50) *The Strongest Parametric Variation Hypothesis (SPVH):*
The Revised Strongest Minimalist Thesis governs the entire range of parametric variation.

(51) *The Revised Strongest Minimalist Thesis (RSMT):*
Every grammar G forms a maximally optimal solution to legibility conditions at the (different) interfaces.

The question is now what the consequence of SPVH is for the status of parameters. As has already been addressed in the introduction the innate status of parameters is not unproblematic. Under this approach it becomes unnecessary either. The fact that languages can express semantic operators in different ways, accounts for the existence of parametric choices. The first choice that an L1 learner must make in order to determine how a particular Op_F is expressed is by determining whether a formal feature [i/uF] does exist. This follows from the FFF Hypothesis. If not, the language learner does not have to acquire more syntactic information in order to express Op_F in his/her grammar. If F on the other hand is formalised, new choices emerge. Which elements have [iF] and which have [uF]. Moreover, questions arise such as when does [i/uF] project (i.e. on which item).

Crucially is however that such questions impose themselves to language learners as a result of the previous choices. Therefore those questions, i.e. those parameters, do not have to be assumed to be innate. They follow automatically from previous choices. Note that this view safes quite a lot of ballast in the sense that it does not meet all the problems that innate parameters suffer from. On the other hand, it still limits the entire space of grammatical variation. Moreover, in its essence it is still very close to two dominant perspectives on parameters: Bakers Parameter Hierarchy (Baker 2001) and the Chomsky-Borer conjecture (Borer 1984; Chomsky 1995).

The first perspective states that parametric variety is hierarchical. This means that a second parametric choice is only possible after setting the first parameter(s) in a particular way. Note that under the approach formulated above, this also follows. The only difference is that the hierarchy is not innate, it creates itself. Some choices require further choices, whereas some other choices do not require these further choices. The idea that parameters are not innate does not exclude them from being hierarchical w.r.t. each other

The Chomsky-Borer conjecture states that parameters are reduced to properties of functional heads. Under the approach that I propose parameters cannot

be properties of functional heads in the first place, since functional heads are not part of UG. The first parameters, so to say, reduce to properties of semantic operators. Only if they are formalised elements that carry a formal feature can be analysed such that this feature will project. Those elements are functional heads. After this procedure, things are similar, since these heads can serve as locus for more specified parameters. The rationale behind the Chomsky-Borer conjecture is that parametric variation reduces to lexical variation. This also follows from my proposal. Semantic operators are lexical items in the first place and thus parametric variation still reduces to lexical variation.

Hence, the main advantages of the above mentioned perspectives remain. Parametric hierarchies are well motivated empirically, and are also predicted by the SPVH. The idea that parametric variation is lexical variation is also kept.

11. Conclusions

In this paper I have concluded that Chomsky's SMT must be revised by the thesis that every grammar G forms a maximally optimal solution to legibility conditions at the (different) interfaces. This thesis leads to several consequences.

First, it has been shown how different economy conditions, applying at the SM and C-I interfaces, lead to different expressing strategies of semantic operators. A more C-I biased strategy uses different lexical items to express a particular semantic operator; a more SM biased strategy spells out markers of different semantic operations on one and the same lexical item. As a result dislocation is needed in order to make those structures interpretable at LF.

Second, I have demonstrated that in order to license dislocation effects, (uninterpretable) formal features are needed, thus accounting for the existence of redundant material in grammar, a puzzle that has remained unsolved so far.

Third, I have shown that it is possible to describe functional structure in a flexible way. I have presented an empirical testable hypothesis, the FFF Hypothesis that argues that formal features are syntactically flexible. According to this hypothesis, the set of formal features in empty in UG, and formal features are acquired as a result of doubling phenomena in the language input.

Fourth, I have demonstrated that the behaviour of concord phenomena, such as Negative Concord and Modal Concord, provide firm evidence for this hypothesis.

Fifth, the idea that formal features are acquired as a result of doubling effects explains why only formal features are allowed to project. The fact that only formal features may project, in combination with the syntactic flexibility of formal features, accounts for cross-linguistic variety w.r.t functional structure.

Finally, the hypothesis that all grammatical variation follows from the Revised Strongest Minimalist Thesis provides a new perspective on parameters that maintains all the benefits of parameters, namely that parametric variation is limited, hierarchically ordered and lexically encoded, but that does not presuppose that parameters are innately present.

Of course the programmatic nature of this paper leads to many open questions, and I am fully aware of the fact that many problems need to be solved as well. On the other hand, I think the proposals formulated above solve many questions that have been problematic thus far. Moreover, the proposals formulated analyse many aspects of

grammar in terms of interface conditions rather than pointing in the direction of UG, a desideratum in current minimalist reasoning.

References:

Ackema, P., A. Neeleman, et al. (1993). Deriving functional projections. NELS 23.

Adger, D. (2003). Core Syntax: a minimalist approach. Oxford, Oxford University Press.

Anderwalt, L. (2002). Negation in Non-Standard British English. London, Routledge.

Baker, M. (2001). The Atoms of Language: the Mind's Hidden Rules of Grammar. New York, Basic Books.

Borer, H. (1984). Parametric Syntax: Case Studies in Semitic and Romance Languages. Dordrecht, Foris.

Brody, M. (2000). "Mirror Theory. Syntactic Representation in Perfect Syntax." Linguistic Inquiry **31**: 29-57.

Brown, S. (1999). The Syntax of Negation in Russian. Stanford, CSLI Publications.

Chomsky, N. (1981). Lectures on Government and Binding. Dordrecht, Foris.

Chomsky, N. (1995). The Minimalist Program. Cambridge, MA, MIT Press.

Chomsky, N. (2001). Derivation by Phase. Ken Hale: A Life in Language. M. Kenstowicz. Cambridge, Massachusetts, MIT.

Chomsky, N. (2005). OnPhases, MIT.

Chomsky, N. (2006). Turing's Thesis. Lecture, presented at the InterPhases workshop, held in Nikosia, Cyprus, May 18-20.

Cinque, G. (1999). Adverbs and Functional Heads – A Cross-Linguistic Perspective. New York, Oxford University Press.

de Swart, H. and I. A. Sag (2002). " Negation and Negative Concord in Romance." Linguistics and Philosophy **25**: 373-417.

Ernst, T. (1998). The Syntax of Adjuncts. New Brunswick, Rutgers University.

Geurts, B. and J. Huitink (2006). Modal Concord, University of Nijmegen.

Giannakidou, A. (2000). "Negative ... concord?" Natural Language and Linguistic Theory **18**: 457-523.

Giorgi, A. and F. Pianesi (1997). Tense and Aspect. From Semantics to Morphosyntax. New York/ Oxford, Oxford University Press.

Grohmann, K., N. Hornstein, et al. (2005). Understanding Minimalism. Cambridge, Cambridge University Press.

Haegeman, L. (1995). The Syntax of Negation. Cambridge, Cambridge University Press.

Haegeman, L. and R. Zanuttini (1996). Negative Concord in West Flemish. Parameters and Functional Heads. Essays in Comparative Syntax. A. Belletti and L. Rizzi. Oxford, Oxford University Press.

Halliday, M. A. K. (1970). "Functional diversity in language as seen from a consideration of mood and modality in English." Foundations of Language **6**: 322-361.

Haraiwa, K. (2000). Agree and closeness in multiple specifiers, MIT.

Haraiwa, K. (2001). Multiple Agreement and the Defective Intervention Effect, MIT.

Heim, I. and A. Kratzer (1998). Semantics in Generative Grammar. Oxford, Blackwell.

Herburger, E. (2001). "The negative concord puzzle revisted." Natural Language Semantics **9**: 289-333.

Kayne, R. (2000). Paramteres and Universals. Oxford, Oxford University Press.

Koeneman, O. (2000). The flexible nature of verb movement, Utrecht University.

Koster, J. (1987). Domains and Dynasties. Dordrecht, Foris.

Ladusaw, W. (1992). Expressing Negation. SALTII, Cornell University.

Lasnik, H. (2002). "The Minimalist Program in syntax." Trends in Cognitive Sciences **6**: 432-437.

Lightfoot, D. (1999). The development of language: Acquisition, change and evolution. Oxford, Blackwell.

Lyons, J. (1977). Semantics. Cambridge, University Press.

Merchant, J. (2001). Why no(t). University of Chicago.

Neeleman, A. and H. Van de Koot (2002). "The Configurational Matrix." Linguistic Inquiry **33**: 529-574.

Newmeyer, F. (2004). "Against a parameter-setting approach to language variation " Linguistic Variation Yearbook **4**: 181-234.

Newmeyer, F. (2005). Possible and probable languages. A generative perspective on linguistic typology. Oxford, Oxford University Press.

Newmeyer, F. (2006). Newmeyer's rejoinder to Roberts and Holmberg on parameters, University of Washington.

Nilsen, O. (2003). Eliminating Positions, PhD Disseration, Universiteit Utrecht.

Pollock, J.-Y. (1989). Verb Movement, Universal Grammar, and the Structure of IP. Linguistic Inquiry. Mass., MIT. **20, no 3:** 365-424.

Rizzi (1997). The fine structure of the left periphery. Elements of Grammar: Handbook in Generative Syntax. L. Haegeman. Dordrecht, Kluwer.

Rizzi, L. (1990). Relativized Minimality. Cambridge MA, MIT Press.

Roberts, A. and A. Holmberg (2005). On the role of parameters in universal grammar. A reply to Newmeyer. Organizing Grammar. Linguistic Stusies in honor of Henk van Riemsdijk. H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz and J. Koster. Berlin, Mouton de Gruyter.

Starke, M. (2001). Move Dissolves into Merge: a Theory of Locality, University of Geneva.

Starke, M. (2005). On the inexistence of specifiers and the nature of heads. **Structures and Beyond. The Cartography of Syntactic Structures**. A. Belletti. Oxford, Oxford University Press.

Stechow, A. v. (2002). Temporal phrases with quantifiers University of Tuebingen.

Svenonius, P. (2001). Subject positions and the placement of adverbials. Subjects, expletives, and the EPP. P. Svenonius. Oxford/New York, Oxford University Press**:** 199-240.

Travis, L. (1984). Parameters and Effects of Word Order Variation, MIT.

Ura, H. (1996). Multiple Feature Checking, MIT.

Weerman, F. and A. Neeleman (1997). Flexible Syntax. Dordrecht, Kluwer.

Weiss, H. (2002). Three types of negation: a case study in Bavarian. Syntactic microvariation. S. Barbiers, L. Cornips and S. Van der Kleij, Meertens Institute Electronic Publications in Linguistics**:** 305-332.

Zeijlstra, H. (2004). Sentential Negation and Negative Concord. Utrecht, LOT.

Zeijlstra, H. (2006). On the syntactic flexibility of formal features. The Limits of Syntax. A. Holmberg and I. Roberts. Amsterdam, Benjamins**:** Submitted.