DISLOCATION EFFECTS, UNINTERPRTABLE FEATURES, FUNCTIONAL HEADS 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[1] and meaning (Chomsky 2000; Chomsky 2001; Lasnik 2002; Chomsky 2005a, b). This perspective takes language to be an optimal solution to conditions that are imposed to the Faculty of Language by its neighbouring mental modules, the Sensory-Motor system and the Conceptual-Intentional systems.

However, the idea that language is 'perfect' in this sense seems to be at odds with several 'imperfections' found in grammar, such as agreement phenomena or the dislocation property. Implemented in linguistic theory, at least four properties of language appear to be 'imperfections' rather than 'perfections': (i) the existence of uninterpretable formal features; (ii) dislocation; (iii) the cross-linguistic flexibility of morpho-syntactic categories; and (iv) the existence of cross-linguistic variation.

In this paper I argue that all four properties addressed above are not linguistic imperfections, but are actually predicted by the perfectness

[1] As is well known, sound is interpreted in a broad sense, including signs in sign languages, gestures, etc.

hypothesis. In a nutshell, I argue that the different conditions imposed to the Faculty of Language by are not always compatible to each other, and that therefore the Faculty of Language can offer multiple, equally optimal solutions to these conflicting interface conditions. In this sense, fulfilling interface conditions is some kind of a trade-off. Being a perfect solution to one interface condition may imply that another interface condition cannot be maximally solved, and vice versa. Hence perfect solutions to interface conditions, which are in conflict with other interface conditions, can only exist by virtue of less perfect solutions to these other interface conditions. Consequently, some linguistic imperfections are consequences of conflicting interface conditions and thus epiphenomenal in nature.

The central claim of this paper is that the existence of uninterpretable formal features, dislocation effects, the cross-linguistic flexibility of morpho-syntactic categories and the existence of cross-linguistic variation are all epiphenomena of the perfectness hypothesis.

This paper is organised as follows: in section 2, I discuss Chomsky's Strongest Minimalist Thesis (Chomsky 2005a) which takes the language to be an optimal solution to interface conditions imposed to that are imposed to the faculty of language and I discuss the nature of two such interface conditions (one sound-based, one meaning-based). In section 3, I introduce the four linguistic 'imperfections' and argue why they appear to be problematic for the perfectness hypothesis. In section 4 I discuss the notion of optimality in grammatical architecture and demonstrate why dislocation effects and

uninterpretable formal features result from conflicting interface conditions. In section 5 I discuss the notion of functional heads and projections and I argue that the set of formal features is not universal but triggered during L1 acquisition as a consequence of the existence of uninterpretable features. In Section 6 I argue that the different marking strategies that are pregiven by a 'perfect' grammatical architecture constitute the range of parametric variation that is attested in natural language. Section 7, finally, concludes.

2. The Perfectness Hypothesis

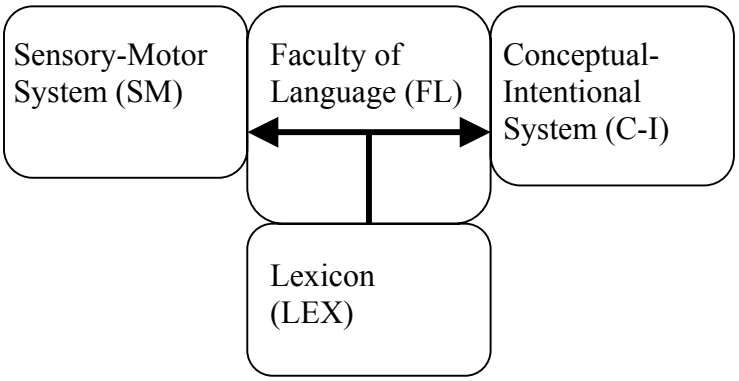
In current minimalist reasoning language is assumed to be a 'perfect' solution to the task of relating sound and meaning (Chomsky 2000; Chomsky 2001; Lasnik 2002; Chomsky 2005a, b). This perspective takes language to be an optimal solution to conditions that are imposed to the Faculty of Language by its neighbouring mental modules, the Sensory-Motor system and the Conceptual-Intentional systems.

2.1 The Strongest Minimalist Thesis

Chomsky's Strongest Minimalist Thesis (SMT) states that 'language is an optimal solution to interface conditions that the Faculty of Language (FL) must satisfy' (Chomsky 2005a: 3). This thesis, tracing back to the philosophical view that language enables human beings to express their

thoughts (a view endorsed in the biolinguistic program) 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 systems that deal with the expression of a sentence and the meaning it conveys, the Sensory-Motor (SM) system and the Conceptual-Intentional (C-I) system respectively, as well as with an instance of memory, the lexicon. This is illustrated in (1) below:

(1)



In the diagram in (1) the input for FL consists of a lexical numeration and the output (after separating at Spell-Out) passes on either to the SM interface or to 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 to each connected mental component. This amounts to saying that the two interpretational systems impose conditions on the structures that have to be met at the interface.[2]

[2] I do not take into account the possibility that the lexicon should be considered a mental module as well. Note that nothing in this paper is incompatible with that idea either.

It is important to distinguish two different types of conditions: hard en soft conditions. Hard conditions are conditions that must always be satisfied. The Principle of Compositionality, which states that the meaning of a sentence follows from the meaning of its parts and the way in which these parts are ordered, for instance, is a hard condition that the C-I system imposes on FL. If Compositionality cannot be applied, the derivation will crash at the level of Logical Form (LF), the interface between FL and C-I.

Soft conditions, on the other hand, are conditions that express preference. Economy conditions are well known examples, such as the last resort constraint with respects to movement. This constraint does not rule out movement, but only states that movement must be as late as possible. In other words, it favours late movement over early movement, but does not exclude early movement a priori. Such a constraint does not rule out any kind of early movement a priori, but compares a number of possible derivations and assigns grammaticality to only one candidate.

Hence, in principle, a grammatical architecture like (1) allows in total four kinds of conditions: both neighbouring mental modules may impose both hard and soft conditions. Whereas hard conditions must be completely satisfied, soft conditions must be optimally satisfied.

However, it follows from nothing that the different soft conditions cannot be in conflict to each other. On the contrary, it is much more likely that, if mental components function autonomously, soft conditions from different mental modules are not always compatible. It could in fact very well be the case that by satisfying a particular interface condition imposed by the

SM-system, some other interface condition, for instance applying at the C-I interface, cannot be maximally satisfied anymore, and vice versa. This implies that if two soft conditions are in conflict with each other multiple, equally optimal, solutions may fulfil these conditions.

Hence, natural language cannot be seen as a single optimal solution to different interface conditions, but its various instantiations should be thought of as multiple, equally complex, solutions to different interface conditions. The central aim of this paper is to reduce existing cross-linguistic or more specifically, parametric variation, to the variety that natural languages constitutes in order to the variation that natural language allows a result from SMT. This amounts to adopting the following hypothesis:

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

Under (2) parametric variation is no longer an imperfection that natural language seems to exhibit, but an epiphenomenon of the supposed perfection of natural language.

Thus, the SPVH leads to a view on individual grammars that takes the SMT not only to be a hypothesis on FL and the nature of natural language. The SMT is now also on individual natural languages and possible grammars. The variety of natural languages, or to be more precise, the different

characteristics they exhibit, follow from different conflicting soft conditions imposed at the interfaces FL shares with other mental components.

Before discussing the range of variation that the SPVH constitutes, let us first discuss two important interface conditions that are imposed to FL.

2.2. Optimal design in the architecture of grammar

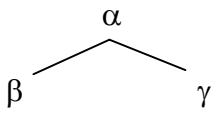
In this subsection I take a closer look at the kind of hard and soft conditions that apply at the two interfaces. I first argue that the hard C-I condition that interpretation follows from Functional Application and Predicate Modification results in a simplicity metric that favours structures that lack uninterpretable features over structures that have them. Then I argue that, from the SM perspective, light prosodic structure is preferred over rich prosodic structure.

2.2.1 Optimal design from the C-I interface perspective

Let us try to enter the mind of a purely semantically (C-I) biased language engineer in order to investigate what kind of conditions the C-I systems impose on FL. From the semantic perspective the most important requirement on linguistic structure is that it allows for compositional interpretation (Frege 1892; Janssen 1983; Partee 1984; Hendriks 1993; Szabó 2000). This means that a particular interpretation of a non-terminal element α (as in (3)), $||\alpha||$

follows from ||β|| and ||γ|| through Functional Application (FA) or Predicate Modification (PM), as defined by (Heim and Kratzer 1998) and illustrated for extensional semantics in (4) and (5) respectively.[3]

(3)



(4) *FA*: If f is the set of daughters of branching node α, and [[β]] is in $D_{<a,b>}$ and [[γ]] is in D_a, then [[α]] = [[β]]([[γ]])

(5) *PM*: If {β, γ} is the set of daughters of branching node α, and [[β]] and [[γ]] are both in $D_{<e,t>}$, then [[α]] = λx.[[[β]](x) & [[γ]](x)]

Consequently, from the semantic/C-I perspective there is no reason to assume more structure to be present than the compositional interpretation of the top node requires. In other words, there is no reason to assume any abstract structure projected by semantically vacuous elements. If a particular lexical item does not contribute to the meaning of the sentence there is no need to assume its presence at LF. This assumption leads to the following two conclusions: (i) the C-I systems prefer no semantically vacuous elements at the level of LF as they cannot be motivated by any C-I condition and (ii) nothing a priori rules out the presence of semantically vacuous material at LF.

[3] For illustratory purposes I only included the extensional definitions.

These two conditions allow us to formulate the following semantic simplicity metric:

(6) *Semantic Simplicity Metric*: A structural representation R for a substring of input text S is simpler than an alternative representation R' iff R contains fewer uninterpretable features than R'.

Note that (6) is a weaker version of Full Interpretation (Chomsky 1995) as its does not forbid the occurrence of semantically vacuous material at LF.[4] If for an independent reason uninterpretable material can be motivated at LF, nothing would rule out the sentence. In the following section I demonstrate that SM-based soft interface conditions may in fact require the presence semantically vacuous material at LF.

2.2.2 Optimal design from the SM interface perspective

[4] The idea that Full Interpretation requires that the semantic content of structures at LF must be interpretable and that there is no ban against uninterpretable material at LF, as long as its syntactic licensing requirements have been met, could be expected to rule in vacuous quantification. However, the constraint on vacuous quantification has been argued not to be a necessary constraint on syntax in the first place. See (Potts 2002) for a series of arguments in favour of this view.

It is a well-known fact about natural language that syntactic structure is not always identical to the simplest structure that meets all compositionality requirements. This must be due to the fact that language, apart from semantic requirements, also needs to satisfy conditions, which are imposed by the SM system. Otherwise, following the SMT, syntactic structure should only reflect the simplest possible configuration that would allow for a compositional interpretation. Once again, we should try to enter the mind of a language engineer, but this time the mind of purely phonologically (SM) biased one.

Work by (McCarthy 1986) and (Hopper and Traugott 1993) on phonological simplicity and grammaticalisation claims that the SM system prefers phonologically weak items over strong ones: affixes over clitics; particles over lexical words, etc. (see (7))

(7) a. *Word > Foot > Syllable > Mora* (McCarthy 1986)

b. *Content word > Particle > Clitic > Affix* (Hopper and Traugott 1993)

Although generalisations like the ones in (7) have been formulated from the 19th century onwards, these preferences are puzzling from the perspective of the SM system. In the previous subsection on semantics I argued that elements lacking meaning, i.e. elements that are uninterpretable for the C-I system, are dispreferred. But there is no preference in favour of 'small meanings' over 'large meanings'. Hence, the fact that small words like affixes are preferred over big words such as content words may seem intuitive, as has

been argued for in studies on grammaticalisation, but the SM system does not disfavour large words in any way.

However, the crucial distinctions in (7) are not about word length, but about what separates different words or word parts. Adopting a proposal by (Neeleman and Koot 2006) that prosodic representations are string-based, in the line of (Chomsky and Halle 1968; McCawley 1968), prosodic categories (utterances (U), intonational phrases (I), prosodic phrases (Φ) prosodic words (ω), feet (F) and syllables (σ)), are considered to be prosodic boundaries. This means that phonological elements are interrupted by phonologically uninterpretable, unpronounceable, material, as their example (8) shows.

(8) *U* John's *ω* father *Φ* suggested *ω* a two-seater *I* but *ω* John's *ω* mother *Φ* preferred *ω* a fur *ω* coat *U*.

Thus the preferences in (7) can be replaced by the single preference in (9).

(9) $\sigma > F > \omega > \Phi > I > U$

Similar to the semantic case, nothing a priori bans prosodic boundaries, but they are not motivated by the SM-system, for it is a system that merely interprets phonological material. The fact that prosodic boundaries are not motivated but neither forbidden by the SM system, induces a phonological simplicity metric (10).

(10) *Phonological Simplicity Metric*: A structural representation R for a substring of input text S is simpler than an alternative representation R' iff R contains less prosodic boundaries than R'.

What (10) reflects is that prosodic phrase boundaries are dispreferred over prosodic word boundaries, which in their turn are dispreferred over foot boundaries, etc. Note that (10) can be regarded as a SM variant of the Principle of Full Interpretation. In other words, the expression of two particular elements carrying semantic content preferably constitutes a single word (e.g. a root affix combination (11) rather than two different prosodic words (12).

(11) $[_{PW}\ Root_1\text{-}AF_2]$

(12) $[[PW_1] \ldots [PW_2]]$[5]

Of course, it is not the case that structures that are prosodically richer than what would be desired from an SM based perspective are excluded from natural language, as there is no principled motivation to rule out rich prosodic structures. Again, the preference takes the shape of a soft condition that prefers to assign poor rather than rich prosodic structure to a particular string of phonological elements.

[5] In this paper I concentrate only on the difference between affixes and prosodic words.

3. Four linguistic 'imperfections'

In the previous section I have presented two different simplicity metrices that are both soft conditions imposed on FL. However, these two principles only induce the SM-based preference for poor prosodic structure and the CI-based preference against uninterpretable material. However, natural language features many more characteristics, which do not seem to be the result of the simplicity principle behind the SPVH (and therefore behind the SMT). For instance uninterpretable features are dispreferred from the Semantic Simplicity Metric, but still widely attested in natural language. In this chapter I introduce, next to the notion of uninterpretable features, three other instances of linguistic imperfections that do not seem to result from the SMMT: dislocation effects, the cross-linguistic flexibility of functional heads and the existence of parametric variation. After this section, I demonstrate that each of these four linguistic imperfections are actually the result of the SPVH and are thus correct predictions of the perfectness hypothesis.

3.1 Dislocation effects

First, dislocation effects, the fact that a Lexical Item (LI) may occupy a different position in the structure (its position in the phonological representation of the sentence) to that in which it is interpreted semantically

(the position of the semantic representation), seems to be one of the core properties of natural language. As is well known, many LI's contribute to dislocation effects, such as fronting of Wh elements, topicalisation, scrambling, verb movement, Quantifier Raising (QR), etc. Against the background of the perfectness hypothesis this immediately leads to the question as to why the semantic position of a particular LI does not simply coincide with its phonological position. Note that this question is a different one to the question as to why dislocation is possible in the first place. Arguing that there is no principled ban on re-merging elements has adequately solved this question (Starke 2001; Chomsky 2005a-b), although this leaves unexplained how semantic compositionality requirements remain unviolated after remerging a linguistic object. However the fact that remerge (or internal merge) is not blocked as a matter of principle does not answer the question why natural language exhibits dislocation. In other words, the existence of unrestricted Merge, accounts for the possibility of dislocation, since there is no principled ban on remerging. However, that does not mean that dislocation effects are immediately expected to occur. As is well-known, remerging, and therefore dislocation effects, are heavily restricted. The question why dislocation, despite being freely available is so much restricted, must be due to the fact that although it is possible, it is not necessary, and since it is not necessary it is ruled out. The notion of merge suffices to explain the possibility of dislocation, but not its necessity.

In (Chomsky 2005a) it has been argued that the duality of semantics (i.e. the distinction between the expression of argument structure and

discourse properties) calls two different modes of syntactic structures into being: arguments structure is realised by external merge, and discourse properties are expressed by means of internal merge. This idea has, however, met a fair amount of critique in the literature (cf. (Moro 2000; Moro 2004; Hinzen 2006)), who apart from presenting some arguments against a semantic motivation for internal merge on evolutionary grounds, argue that there is no independent evidence for the duality of semantics. Moreover it is not clear why discourse properties cannot be expressed by means of external merge. In fact many markers of discourse properties are externally merged, such as West Germanic discourse particles or Classical Arabic focus particles. Hence the question why natural language exhibits dislocation effects is still in need of a principled explanation, as it does not seem to be inevitable in order to perfectly connect sound to meaning.

3.2 Uninterpretable features

Second, the notion of uninterpretable formal features requires a principled explanation. Uninterpretable features are those features that cannot be interpreted, neither by the SM system, nor by the C-I systems. This immediately opens the question as to why natural language would allow for redundant material in the first place? At first sight language seems to be full of redundancy as suggested by concord phenomena (such as Negative Concord) or overt agreement (subject-verb agreement). The line of reasoning developed in (Chomsky 1995) was to take one imperfection to license the

other. Following the principle of Full Interpretation (Chomsky 1986) that states that the interfaces should be free of uninterpretable material, uninterpretable features ([uF]'s) must be deleted during the derivation, before reaching LF or PF. Deletion takes place by establishing a feature checking relationship with a local element carrying an appropriate interpretable formal feature. However, if the structural distance between a particular LI carrying some uninterpretable feature and its possible feature checker (i.e. an LI that carries a matching interpretable feature) is too big to allow for feature checking, a syntactic operation such as Move may be triggered, thus motivating the triggering of internal merge. Not moving this element would lead to a violation of Full Interpretation at LF. 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, but that it leaves the existence of redundant material itself an open question (Chomsky 2006).

3.3 Functional heads

Third, natural language exhibits a flexible distribution of functional heads. Since the introduction of multiple clausal functional heads (most notably by (Pollock 1989)) it is observed that languages differ with respect to which functional heads are (overtly) realised. Why would a particular language exhibit some F° if another language can do without it? Roughly speaking, two different approaches have been proposed to account for this flexible

distribution of functional heads. The first approach, the so-called cartographic approach[6], has taken the strong, radical claim that each language underlyingly has the same functional structure that reflects the many hierarchies that have been observed in natural language (Larson 1988; Rizzi 1997; Cinque 1999; Belletti 2004b; Ramchand 2007) and that grammatical variation is restricted to which positions are overtly realised. Proposal along these lines have been formulated by (Cinque 1999; Kayne 2000; Starke 2005) and (to a lesser extent) (Rizzi 1997).[7] If this approach is on the right track, (rich) functional structure can be taken to be part of UG and the language-specific realisation can be reduced to parametric variation. Under such an approach, the abstractness of functional structure (i.e. why are several functional heads allowed be covertly realised whereas others are realised overtly) remains unexplained and therefore lacks principled explanation. As there is no explanation why UG should innately be equipped with such a rich structure, several scholars, e.g. (Ernst 2002; Svenonius 2001; Nilsen 2003) have argued against such a universal UG-based functional sequence (terminology due to (Starke 2001)). These scholars have pointed out several problems with respect to the cartographic approach by arguing that the clausal hierarchy that the

[6] Not every analysis that is cartographic makes this strong assumption. Several analyses actually allow for cross-linguistic differences with respect to clausal structure. Cf. (Iatridou 1990; Giorgo and Piranesi 1997) among many others.

[7] Many papers written within this cartographic approach can be found in (Cinque 2001; Rizzi 2004a; Beletti 2005).

cartographic approach imposes turns out to be problematic, as many functional orders can in fact be reversed. Moreover, it has been argued that many hierarchical effects, as well as the observed transitivity failures can be explained by adopting a semantic motivation for the orderings observed within clausal structure. Under such an approach the universal functional sequence is rejected and functional heads themselves become subject to parametric variation. However, also under this perspective, it remains unclear what determines the existence/availability of functional heads of they are not innately provided by UG.

3.4 Parametric variation

Finally, if language performs a perfect task in relation sound and meaning, why would different instantiations of it (i.e. different languages) opt for so many differences? Why would morphosyntax not be identical across different languages? The introduction of the Principles and Parameters program (Chomsky 1981) has provided an initial answer to this question by arguing that linguistic variation is not unlimited and the notion of parameter has been introduced.[8] 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

[8] For a recent debate on the status of parameters in linguistic theory, cf. (Newmeyer 2004; 2005; 2006; Roberts & Holmberg 2006, Baker 2007).

due to the success of microparameteric studies, has revealed that the number of parameters is no longer easily countable. Even the most optimistic analyses take the number of parameters to be larger than 100-150 (cf. (Newmeyer 2004)). At the moment there is no clear notion what the exact number should be, but given the idea that each human I-grammar results from a different parameter setting, the number of parameters must be accordingly large. Since previous and future I-grammars must be taken into account as well, the number may even exuberantly increase. If the number of parameters is indeed as high as it is esteemed these days, their innate status is getting less and less likely, not only because such a large amount of innate acquisition instructions is hard, if not impossible, to explain in terms of language evolution and genetic encoding, but also since an explanation for such a distribution of innately present parameters is lacking. If linguistic variation is indeed constrained by parameters, what constitutes parameters themselves? Why do parameters exist in the first place?

A hypothesis put forward by (Baker 2001; Baker 2007) is to separate different types of parameters, thus distinguishing microparameters from macroparameters, and possibly intermediate types, such as mesoparameters. Whereas microparameters can be reduced to particular properties of functional heads (following the Borer-Chomsky conjecture (Borer 1984) and (Chomsky 1995)), macroparameters should distinguish different types of language families. However, this distinction between possible types of parameters does not solve the problems that have been addressed. First, even under a perspective that only macroparameters are innately present (not a

conclusion that Baker is necessarily committed to), the question as to why a perfect system like natural language would allow for parameters, remains unanswered. The major question is not how many parameters there are, or what form they can have, but why they are there in the first place. Second, even if many parameters are reduced to properties of functional heads, it still needs to be accounted for what properties enable functional heads to constitute the particular amount of cross-linguistic variation that has been attested.[9] Moreover, if parametric variation reduces to properties of functional heads, the previously addressed question again emerges: what determines the set of functional heads. Why would natural language allow a series of functional heads if, at least from a superficial point of view, not every language exhibits all possible functional heads.

3.5 Line of argumentation

In this paper, I address all four 'imperfections' that require principled explanation, and I argue that the existence of dislocation effects, uninterpretable formal features, functional heads and parameters follows from the fact that language constitutes an optimal solution to the task of 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 natural is an optimal solution to conditions imposed by the different interfaces. However, the fact that these interface conditions can be conflicting opens up the possibility that different

[9] See (Gianollo et. al.) for an analysis if these properties.

grammars may equally optimally satisfy their interface conditions. This already creates a grammatical space that allows for cross-linguistic variation. I propose the radical hypothesis, that this room for grammatical variation, which follows from conflicting interface conditions, forms the entire parametric space.

I demonstrate that dislocation operations are required in order to spell out two markers of different semantic operations in one and the same position. If two semantic functions cannot be interpreted in one and the same position, since for instance their semantic types form a mismatch, remerge creates a second syntactic position, so that each semantic function can be interpreted in a unique position. This argument is close to the argument following from the duality of semantics (since it derives movement from the fact that LI's may induce multiple semantic functions), but crucially differs from it since it takes SM-based soft conditions to be responsible for the fact that multiple semantic functions are spelled out in one syntactic position.

Remerge is, however, not the only available mode of repairing the fact that mismatching semantic functions are realised in one syntactic position. A second, equally economical alternative is by assuming that if two semantic functions are marked on one LI, that one of these two markers can be semantically vacuous and licenses the presence of a phonologically abstract element that is interpreted as the second semantic function. Note, however, that this mode of repair calls uninterpretablity into being as such elements need to carry uninterpretable formal features in order to license the higher abstract semantic function. Hence the same mismatch between phonological

and semantic economy conditions that causes movement is also the cause of the existence of redundant material in natural language.

Finally, I discuss the relation between functional heads and formal features, demonstrating that feature projection is only allowed for formal features, and that the set of functional heads in a particular language depends on which formal features are available in that language. This alludes to the cartography-flexibility debate: is the set of formal features identical cross-linguistically or not? In this paper I demonstrate that formal features must be acquired through positive evidence. In short, I demonstrate that features can only be taken to be formal if there is positive evidence that at least one instance of this formal feature is uninterpretable. This explains why the set of functional heads is flexible. Only if [F] is some formal feature in a particular language, may F project in this language. If a language lacks a particular formal feature [F], it must also lack a syntactic head F°. I discuss some phenomena (Negative Concord, Modal Concord) that support this prediction.

The most far-reaching consequence of this idea is that the notion of parameters, as well as the motivation for other syntactic operations, such as Move (Remerge) and the feature checking system underlying Agree, are pregiven by the language system as a result from the fact that it is a perfect system. Thus, simply arguing that natural language is an optimal system connecting the thought and speech systems already accounts for the available linguistic tools (dislocation, uninterpretable formal features, the flexible distribution of functional projections and parameters) without having to allude to biologically innate knowledge or to argue against the consequences

of poverty of the stimulus arguments. In this sense these characteristics can be seen as 'Factor III' properties in the sense of (Chomsky 2005a).

4 Uninterpretablity and dislocation as a result of the mismatch between SM and C-I interface conditions

In this section I argue that in many cases the two simplicity metrices, (6) and (10), cannot always be satisfied simultaneously. To be more precise, in every case where two semantic operators are not of a matching type, it is impossible to have these semantic expressions spelled out in a prosodically poor construction, and at the same time have them take scope from that position. To illustrate this I demonstrate that grammatical tense cannot be interpreted from the same position where it is spelled out, namely on the finite verb. In a nutshell, this opens up different marking strategies for natural language to express tense. Either by an adverbial operator, which occupies a different clausal position than the verb (as is the case in for instance in Greenlandic), or by merging the finite verb in two different position while spelling out only one copy (e.g. French) or by taking the tense markers on the verb to be semantically vacuous, i.e. pure scope markers realising a higher covert operator (e.g. Dutch).

In order to see this, take the following sentence:

(13) John loved Mary

The sentence contains two arguments (*John* and *Mary*) as well as a finite verb marked for 3rd person singular and past tense. Focussing on the latter, the question rises where past tense is actually interpreted in the sentence. As (von Stechow 2002) has demonstrated the first suggestion that may come to mind, namely that past tense is interpreted in situ (i.e. on the finite verb), cannot be correct for semantic reasons. This is illustrated in (14).

(14) Wolfgang played tennis on every Sunday (von Stechow 2002)

= 'For every Sunday in Pastc there is a time t at which Wolfgang plays tennis'

≠ 'There is past time on every Sunday at which Wolfgang plays tennis'

≠ 'For every Sunday, there is time before it such that Wolfgang plays tennis at that time'

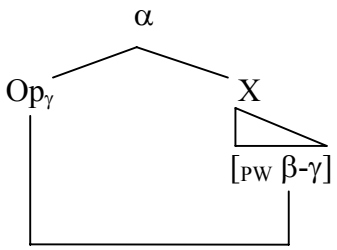
As can be seen from the correct interpretation in (14), past tense cannot be interpreted in the same position where the verbal content ('play') is interpreted since the past tense outscopes the quantifying PP *on every Sunday*, whereas the predicate 'play' is outscoped by this PP. In more technical terms,

the logical types of 'play' and the past tense operator (Op_{PAST}) do not match.[10]

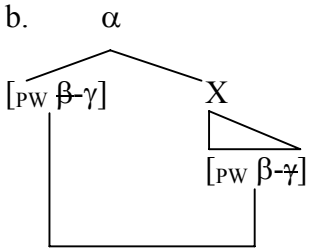
The prosodic word *played* thus cannot induce the semantic contents of the predicate 'play' and Op_{PAST} at the same time. In other words, what seems to be the case here is that the phonological preference to express Op_{PAST} by means of single affix (*-ed*) yields a semantic problem: how is past tense interpreted in sentences like (14)?

As it is obvious that *–ed* is the only marker of past tense in this sentence, two logical possibilities arise. The first possibility is that the finite verb does not move to a higher position, but that it somehow licenses the presence of a phonologically abstract operator that has the semantics of Op_{PAST}. The second possibility is that the finite verb (*loved*/*played*) has remerged (Chomsky 2005a) in such a way that past tense is interpreted in the higher copy and the predicate in the lower copy (i.e. partial reconstruction after movement). The abstractions behind both options are illustrated in (15).

(15) a.



b.



[10] Adopting Von Stechow's representation for Op_{PAST}: [[PASTc]] c = $\lambda w.\lambda P_{it}.\exists t[t < t_c \ \& \ P(t)]$

In the next subsections I argue that the first possibility is an instance of the syntactic operation known as Agree and the second possibility reflects Move. Moreover I argue that both strategies cannot be instantiated without the linguistic notion of uninterpretable features. When I refer to both strategies in (15), I use the term doubling, as in both strategies, γ is manifested more than once in the syntactic structure.

4.1 Uninterpretable features

Let us first concentrate on the first strategy to express multiple markers of semantic functions on one and the same element. In the previous subsection I explained how the existence of additional structure hosting an abstract operator follows from the SM desideratum for phonological economy. However, nothing yet has been said about how this extra functional structure arises; it has only been explained why it must arise.

In the case of (15)a it is a property of the prosodic word [β-γ] that γ has no semantic value itself, but that the presence of γ implies that it is c-commanded by an abstract operator Op_{γ} that is responsible for the semantics that correspond to the affix γ. Hence the prosodic word [β-γ] has an additional particular property such that it meets the following three criteria:

(16) a. $[Op_{\gamma}$ […[β-γ] …]]

b. *[… [β-γ]…]]

c *$[Op_{\gamma}$ [… [β] …]]

The criteria in (16) state that [β-γ] is grammatical if c-commanded by Op_{γ}, but that ungrammaticality is yielded if either Op_{γ} is merged in the clause without the occurrence of the affix γ or if the affix γ occurs without being c-commanded by Op_{γ}. In other words, γ is a morpho-syntactically visible element that is semantically empty, but simultaneously requires the presence of an operator that dominates it in order to survive at LF. Readers who are familiar with the minimalist program will immediately recognize γ as the phonological realisation of a so-called uninterpretable formal feature [uF] that must establish a syntactic relationship with an interpretable formal feature [iF] in order to prevent crashing at LF (in the sense of (Chomsky 1995)).

It should be noted, however, that the perspective on uninterpretable features in this sense is not exactly similar to the perspective in (Chomsky 1995). Uninterpretable features as described above are not illegible to the C-I system. They are thought to be semantically vacuous, which basically means that the C-I system is blind to them, contrary to the view in (Chomsky 1995) where the presence of undeleted uninterpretable features at LF makes the derivation crash. Consequently, this means that uninterpretable formal features do not have to be deleted before or at the level of LF. They only need to be properly licensed in syntax. Note also that such a definition for uninterpretable features prevents look-ahead problems. As no uninterpretable feature needs to be deleted, its semantic status does not play any role during the syntactic derivation. The only information that uninterpretable features carry, and which is lexically encoded, is purely syntactic in nature.

Now let us see how exactly the three criteria in (16) are met, given the notion of uninterpretable feature introduced above. Clearly the case of (16)a follows immediately, since the [uF] feature that γ carries is properly licensed by Op_{γ}. Note that, contrary to more traditional analyses of feature checking the hierarchical structure here is the reverse: it is the element carrying [iF] that c-commands [uF]. The reversal of this hierarchical structure has been proposed on different grounds by, amongst others, (Pesetsky and Torrego 2001; Adger 2003; Wiklund 2005; Pesetsky and Torrego 2006), and follows the essentials behind Rizzi's notion of criteria (where semantic operators occupy specifier positions that must share their features with their heads) (Rizzi 1997).

The second condition in (16) is also immediately met as, by definition, any [uF] requires checking by an element carrying [iF].

The third criterion, however, is not directly met. Given the nature of Merge, nothing in principle forbids merger of an abstract operator carrying [iF] with a syntactic object that does not include [uF]. This is not a problem particular to this theory, but is more general one concerning the nature of abstract operators, or even more generally, the nature of abstract material. In order to restrict the inclusion of abstract material to those cases in which it is required, i.e. cases like (16)c, let us adopt the following condition:

(17) Only postulate a covert element if a particular sentence is grammatical and none of its overt elements is responsible for its grammaticality of the sentence

Note that (17) is a truism, if the feature checking is the mechanism behind grammaticality. It only states that if no overt element can be responsible for the fact that some uninterpretable feature has been checked, it must be a covert element. Note that (17) is not a condition that allows inclusion of abstract material in order to save sentences from ungrammaticality. It only states that if a sentence is grammatical, it may be the case that abstract material is responsible for it.[11] Condition (17) is a soft condition in the sense that abstract material may occur in those cases where its presence is somehow unavoidable. It is exactly this economy condition that has also been applied to license *pro* (Rizzi 1986) and it permits inclusion of abstract material only in those cases where the derivation would not have been convergent otherwise. Adopting (17) derives the ban in (16)c: the fact that there is no γ-affix carrying an uninterpretable feature [uF] renders the sentence without Op_{γ} grammatical (all other things being equal) and thus (17) can never be applied.

[11] One may wonder why languages can realise an overt operator if a covert operator is also available. An example would be the realisation of a pronominal subject in a pro-drop language. Note however, that such an overt realisation always comes about with a particular semantic effect, such as topicalisation or emphasis. As such, overt realisation not only denotes the semantic properties of the subject, but also additional semantic/pragmatic properties.

The status of (17) depends on the perspective on grammar one adopts. In a representational system it operates as a filter on representations, which excludes configurations like (16)c. In a derivational system, condition (17) cannot be properly implemented, as nothing can forbid the creation of (16)c and filters do not apply at the interfaces. However, (17) does not necessarily have to function as a syntactic filter. Following (Ackema and Neeleman 2002), who take rightward movement to be banned on parsing grounds and do not take it to be ruled out in syntax proper (as opposed to (Kayne 1994)), (17) can be thought of as a parsing constraint as well. Hence, although grammatical background assumption may alter the status of (17) varies, they do not block its application.

Thus far we have reached the following situation: on the basis of the two simplicity metrices defined in section 3, it follows that there are already two, equally optimal solutions to express two semantic functions β and γ that are not each other's semantic complement. Either β and γ are introduced in the clausal position from which they take scope, or they are expressed on one and the same prosodic word. In the first case the semantic simplicity metric is fully satisfied, but the phonological simplicity metric is not, as the two elements must both be prosodic words. In the second case β and γ are expressed on one word only, but as both semantic functions cannot be interpreted in the position where the prosodic word has been base-generated, one of the two markers actually carries an uninterpretable feature that marks the presence of an abstract operator. In this case the phonological simplicity

metric is fully satisfied, but the semantic simplicity metric is not, as the structure now contains an uninterpretable feature [uF], carried by γ and checked by Op_{γ}.

4.2 Dislocation

However, marking abstract operators is not the only way to optimally fulfil the phonological simplicity metric. Another way, represented in (15)b, is to remerge a particular word that contains two semantically mismatching elements, to a higher position in the tree and have one of the two take scope from the higher position, and the other one from the lower position.

Following (Truckenbrodt 2006; Zeijlstra 2006), who argue that V-to-C movement in the Germanic languages is always semantically motivated (contra (Chomsky 1995; Chomsky 2000), who postulates that this movement in fact takes place at PF)[12], let us focus on verbal movement triggered by imperative morphology. 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). Although both semantic functions (the predicate and speech act operator) are encoded on the verb, they cannot be interpreted in the same position:

[12] See also (Lechner 2007) for a number of arguments in favour of the idea that head movement causes interpretable effects.

(18) Slaap niet!

Dutch

Sleep not

'Don't sleep!'

Sentence (18) means 'it is imperative that it is not the case that you sleep'. The imperative operator (i.e. the operator that has the illocutionary force of an order) takes scope above the negation, whereas negation outscopes the predicate. The scopal relationships immediately follow if movement is assumed to be a markings strategy as sketched above. In order to see this, let us go through the derivation step by step.

First the entire prosodic word V-IMP is copied (or remerged) in the higher position:

(19) [V-IMP] → [[V-IMP] … [V-IMP]]

Now the formal, phonological and semantic content of the prosodic word V-IMP has been copied. However, although the possibility of remerging is given by any system that allows unrestricted application of an operation like Merge, from the semantic and phonological perspective, it faces serious problems. Semantically speaking, the operation applied in (19), is a blatant violation of compositionality. The only way to avoid this violation is to delete all semantic features that have been copied, once. In principle, it does not matter on which copy which semantic feature is deleted, as long as compositionality

is maintained. This means that deletion could target all semantic features in one copy or some in one copy and some in the other copy. In this case, given the semantics of both predicate and speech act operators, the only division that would not yield any uninterpretable construction at LF is one in which the imperative feature is deleted below, and the verbal feature above, as shown in (20). Semantically speaking, (20) has escaped its compositionality violation.

(20) $[\text{V-IMP}_{[V][IMP]}] \rightarrow [\text{ V-IMP}_{[\sout{V}][IMP]} \ldots \text{V-IMP}_{[V][\sout{IMP}]}]$

From the phonological perspective, doubling all phonological features would introduce an uneconomical effect, as there is no reason to spell out lexical material twice, when spelling it out once suffices. In fact, the general idea behind movement is that it fulfils the phonological simplicity metric. However, just as deletion may target semantic features, it may also target phonological features. Following standard minimalist assumptions, the phonological features of the lower copy are deleted and will thus not be realised. It should be addressed however, that contrary to semantic deletion, phonological deletion does not have to take place, as no SM hard condition is violated if phonological material is spelled out twice. This also explains, why in some languages in cases of e.g. Wh movement traces of movement are phonologically realised (McDaniel 1989; Cheng 2000).

Hence, remerging does not suffer from any phonological or semantic problems. However, (20) is still invalid from a syntactic point of view. If, as

in (20), all verbal features are deleted in the higher copy, the moved element could no longer be analysed as a (finite) verb. Yet, clearly the moved element's category remains unchanged. A moved verb remains a verb and behaves like a verb in every respect. But even more crucial: if the finite verb lacked any formal feature in the first place, no feature could ever have projected it. Although the higher copy no longer carries any semantically verbal properties, it must thus still have purely formal verbal properties. In other words, the higher copy must have such a feature that is syntactically recognizable as a verb, but semantically not. The most likely hypothesis now is that the highest copy must carry an uninterpretable verbal feature, whereas the lower copy carries an interpretable one. Then, (21) denotes correct representation after movement has taken place.

(21) $[[\text{V-IMP}_{[\text{IMP}][\text{uV}]\text{i}}\ ] \ldots [\text{V-IMP}_{[\sout{\text{IMP}}][\text{iV}]\text{i}}]]$

All features, except the verbal features, are present at only one copy. The verbal feature must be a formal feature that remains present on both copies, albeit with a different value. Again, [uV] exhibits the diagnostics of an uninterpretable feature: it does not contribute to the semantics of the sentence, and it must stand in a syntactic relationship with a particular element carrying [iV].

The movement solution to optimally fulfil the requirements of the phonological simplicity metric mirrors exactly the agreement approach described in the previous section. There it was the affix that had to be realised

by a formal feature; here it is the root itself. Note that this account of movement comes about with two major benefits. First, it accounts for the fact why uninterpretable formal features are involved in enabling movement. If there was not any formal verbal feature, movement could never have taken place. In a sense the formal feature in the case of movement functions as the vehicle. Note that, the verb's higher [uV] and lower [iV] feature exactly represents the earlier probe-goal configuration underlying movement, but this configuration has now received an explanation in terms of linguistic simplicity. Second, it enables marking by means of uninterpretable features, the strategy discussed in the previous section. This can be explained as follows: suppose movement did not involve uninterpretable features. Then movement could in principle take place without violating the semantic simplicity metric. Hence movement would be a more economic strategy than the Agree strategy, thus ruling out the latter. The fact that movement cannot be realised without uninterpretable features, motivates the notion of semantically redundant features in natural language.

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 crucial roles. It has been demonstrated that without formal features movement and Agree cannot take 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 (dislocations), similar to uninterpretable formal features, follows from conflicting soft interface conditions.

5. Functional structure

5.1 Functional projections

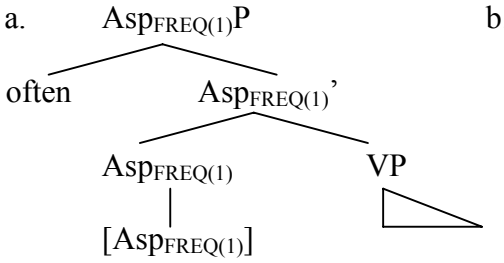
The third imperfection to be discussed in this paper concerns the functional structure. Several scholars have argued that (the amount of) functional structure is a property of UG. One such proposal has been Cinque's adverbial hierarchy, cf. (Cinque 1999), who provides a template for the adverbial distribution, as illustrated in (22). Since Cinque's proposal is one of the most radical ones in terms of fixed templates, it is one of the most interesting ones to discuss here.

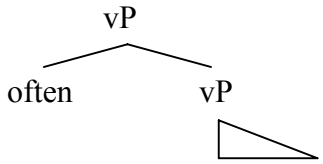
(22) [*frankly* $\text{Mood}_{\text{speech act}}$ [*fortunately* $\text{Mood}_{\text{evaluative}}$ [*allegedly* $\text{Mood}_{\text{evidential}}$ [*probably* $\text{Mod}_{\text{epistemic}}$ [*once* T(Past) [*then* T(Future) [*perhaps* $\text{Mood}_{\text{irrealis}}$ [*necessarily* $\text{Mod}_{\text{necessity}}$ [*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* Asp$_{\text{SgCompletive(I)}}$ [*tutto* Asp$_{\text{PlCompletive}}$ [*well* Voice [*fast/early* Asp$_{\text{celerative(II)}}$ [*again* Asp$_{\text{repetative(II)}}$ [*often* 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 2007). 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 with respect to the English adverb *often*. Under Cinque's analysis, *often* should be located in the specifier position of Asp$_{\text{frequentative(I)}}$P; under a flexible approach it can be taken to be an adjunct to vP. The two options are illustrated below:

(23) a. 

b. 

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

(24) 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 allows multiple specifiers, only two verbal heads can account for the entire distribution in (24). 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.

(25) a. $[_{XP}$ ADV1 X° $[_{YY}$ ADV2 Y° $[_{ZP}$ ADV3 Z° $[_{UP}$ ADV4 U°]]]]

b. $[_{XP}$ ADV1 ADV2 X° $[_{YP}$ ADV3 ADV4 Y°]]

Hence, the only way to distinguish between the two possible structures in (23) lies in the fact that in (23)a a particular formal feature [$Asp_{FREQ(1)}$] must be available to project $Asp_{FREQ(1)}$P. In (23)b, by contrast, this feature 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 such as (23)b. If, however, it cannot be proven that such features are absent, no empirical distinction can be made between (23)a and (23)b. Hence the question which structure in (23) 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 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 (23)a and (23)b, whereas the non-universal approach only allows (23)b. If the predictions that the flexible hypothesis makes are correct, the correct structure for English must be the one in (23)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

in the literature, tracing back to (Giorgi and Pianesi 1997) and their Feature Scattering Principle:

(26) *Feature Scattering Principle*

Every feature can head a projection

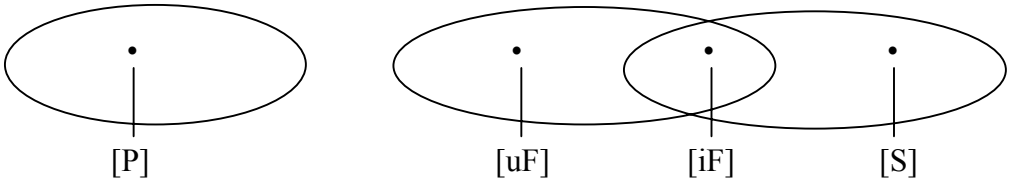
Although it is highly likely that syntactic operations can only access syntactic material (and thus only formal features and not semantic features), the fact that only formal features may project needs to be explained. I will do so after the discussion on flexible features in section 5.3.

5.2 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 interpretable at the SM interface and semantic features are interpretable at the C-I interface. Formal features come in two kinds: interpretable and uninterpretable formal features. Interpretable formal features are also interpretable at the C-I interface, i.e. they carry semantic content, and are therefore also members of the set of semantic features. The sets of formal features and semantic features thus intersect. Uninterpretable formal features need to stand in a proper agreement relation with an

interpretable formal feature in order to prevent the derivation from crashing at the interfaces.

(27) Phonological features Formal features Semantic features



Suppose that the set of formal features is empty at 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 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.

Interpretable formal features have two different properties: they are interpretable at LF and they can check uninterpretable features.[13] However, they are not recognisable as such for a language learner. Their first property is not decisive, since semantically speaking formal features are

[13] Feature checking here is used since it is the common term for the process that is going on here, described in section 4. Checking thus means that a licensing requirement of an uninterpretable feature is fulfilled, thus leading no longer to ungrammaticality.

undistinguishable from non-formal semantic features ([F]'s), as shown in (28).

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

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 doubling effects, thus triggering syntactic operations such as Move and Agree. All these properties can be identified by a language learner. In fact, they all reduce to so called doubling effects. Let us define doubling in the following way.

(29) 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

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 in natural languages then immediately follows. Hence the following hypothesis can be formulated.

(30) The Flexible Formal Feature (FFF) Hypothesis:[14]
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 past tense exhibits doubling effects with respect to past tense; a language only has a formal feature [i/uNEG] if negation exhibits doubling effects with respect to negation, and so on.

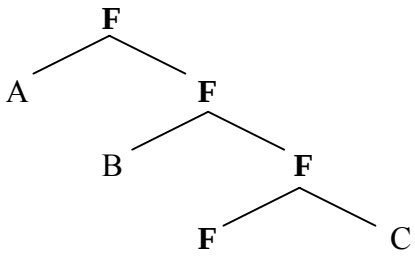
5.3 Formal features and projection

The proposal that only formal features are allowed to project still needs to be addressed. This follows from the observations made above: formal features

[14] This hypothesis was first introduced in (Zeijlstra 2007) in a slightly different form.

may give rise to doubling effects; semantic features do not (if they would they had to be reanalysed as formal features). Let us now see what the consequences are for formal features and projection. Let us take the following abstract functional projection:

(31)



The reader may already have observed that projecting F exhibits a doubling effect with respect to F: a functional projection consists of multiple layers, each assigned a label that corresponds 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 (23)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, which, arguably, is the case here. Formal features are needed in order to enable doubling effects so that conflicting interface conditions can be fulfilled in multiple optimal ways. Now, if these formal features create functional projections as a by-product, this instance of redundancy is not any longer problematic. In fact redundancy in natural language straightforwardly follows from the perfectness hypothesis.

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 can too) and that functional projection is subject to the presence/availability of formal features. But this makes the FFF hypothesis empirically testable. If it can be shown that if a functional F° is present cross-linguistically, 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 not, however, imply that there is no difference between heads and phrases. The only difference is that X°'s, X-Bar'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 (Rizzi 2001; Starke 2001)). Hence standard diagnostics for the distinction between heads and non-heads can 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 2006) as an example).

The fact that heads are empirically detectable, 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:

(32) F° → Overt doubling effects with respect to F

This template for typological universals can easily be explained. The FFF hypothesis accounted for the fact that it takes doubling effects with respect to 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° is overtly realised, it must have been analysed as carrying a formal feature [i/uF] and thus there must have been doubling effects with respect to 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.

The FFFH has not been tested for a wide range of domains, as of yet. However the results that have appeared until now, point in the direction of the FFFH. Let me briefly give two examples.

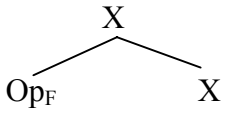
Negation, has been long known to be a functional category that exhibits doubling effects. In many languages two morpho-syntactically negative elements may give rise to a single semantic negation, a phenomenon known as Negative Concord (see (Laka 1990; Haegeman and Zanuttini 1991; Ladusaw 1992; Haegeman 1995; Haegeman and Zanuttini 1996; Zanuttini 1997; Giannakidou 2000; De Swart and Sag 2002; Zeijlstra 2004) amongst many others). In (Zeijlstra 2004) it has been concluded on the basis of a large domain of languages that the FFFH prediction is correct for negation. Every language with a negative marker that is a syntactic head also exhibits Negative Concord.

Another domain is modality. Although many have argued that in languages modal auxiliary occupy a syntactic projection of their own, as is the case in English. According to the FFFH these languages are expected to exhibit 'Modal Concord' effects. In (Zeijlstra 2007) it is shown that this prediction is correct. Modal Concord is indeed present in languages that have a particular modal head.

5.4 Functional structure revisited

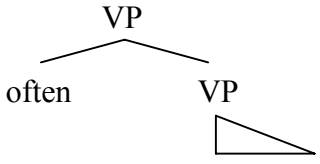
Now that the balance has turned in the direction of flexible functional structure, the question immediately arises why must there be functional structure in the first place. In other words, why do some languages have NegP's, ModP's AspP's, etc.? The answer to this question lies in the fact that each language needs to have some way to express a particular semantic operator. As has been discussed before, the facts that interface conditions may be conflicting leads to equally optimal strategies to express a semantic operator. From the C-I perspective, a structure as (33) would be required.

(33)



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 (23)b, repeated below for convenience.

(34)



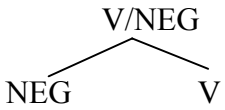
However, as explained before, SM preferences result in a preference for spell-out of multiple elements on the same node. This requires doubling effects and these doubling effects require the presence of formal features. Given the FFF Hypothesis, these formal features are only acquired as a result of doubling effects. Hence if the SM-biased strategies for expressing a semantic operator require 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 doubling effects 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 the case in both Yiddish and in Italian.

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

(36) *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:

(37)



Hence, languages vary with respect to the element that projects in these cases. Italian *non* projects, and therefore exhibits X° behaviour, yielding a NegP. Yiddish 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 require additional structure, but it does not impose any restriction on the label of the new structures. Hence, languages will vary with respect to these labels.

6. Consequences for parameters

Thus far, I have addressed three 'imperfections' in natural language: uninterpretable formal features, dislocation effects, and flexible functional structure. They have all 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 would prefer. 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

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

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 postulating innate parameters in fact becomes unnecessary. 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] exists or not. 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 to when [i/uF] projects (i.e. on which item).

What is crucial, however, is that such questions impose themselves on language learners as a result of previous choices. Therefore those questions, i.e. those parameters, do not have to be assumed to be innate. Note that this view saves quite a lot of ballast in the sense that it does not encounter 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 the two dominant perspectives on parameters: Baker's Parameter Hierarchy (Baker 2001) and the Borer-Chomsky 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 with respect to each other.

The Borer-Chomsky 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 instance since functional

heads are not part of UG. The "first" parameters reduce to properties of a semantic operator (Op_F). Only if these semantic operators have to be analysed as carrying an interpretable formal feature ([iF]), elements carrying a formal feature [i/uF] may project. And only if they do, the language has a functional head F°. After this procedure, things are similar since these heads can serve as locus for more specified parameters, once again established in the input-driven manner outlined above. The rationale behind the Borer-Chomsky 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.

7. Conclusions

In this paper I have concluded that Chomsky's SMT governs the entire range of parametric variation.

First, it has been shown how different economy conditions, applying at the SM and C-I interfaces, lead to different strategies for expressing semantic operators. A C-I biased strategy uses different lexical items to express a particular semantic operator; an SM biased strategy spells out

markers of different semantic operations on one and the same lexical item. As a result doubling 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 empirically testable hypothesis, the FFF Hypothesis, which argues that formal features are syntactically flexible. According to this hypothesis, the set of formal features is empty in UG, and formal features are acquired as a result of doubling phenomena in the language input.

Fourth, the idea that formal features are acquired as a result of doubling effects explains why only formal features are allowed to project: projection is an instance of doubling. The fact that only formal features may project, in combination with the syntactic flexibility of formal features, accounts for cross-linguistic variety with respect to 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 traditional 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 still need to be solved. 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., and Neeleman, A. (2002). 'Effects of Short-Term Storage in Processing Rightward Movement.' in S. Nooteboom et al. (ed), *Storage and Computation in the Language Faculty.* Dordrecht, Kluwer**:** 219-256.

Ackema, P., Neeleman, A. and Weerman, F. (1993). 'Deriving functional projections.' in A. Schafer (ed), Proceedings of NELS 23. Amherst, MA, GSLA: 17-31

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

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

---- (2007). 'The Macroparameter in a Microparametric World.' to appear in T. Biberauer, Holmberg, A. and Roberts, I. (eds), *Limits of parametric variation*. Amsterdam: John Benjamins.

Belletti, A. (2004a). *Structures and Beyond. The cartography of Syntactic Structures*. Oxford: Oxford University Press.

---- (2004b). 'Aspects of the low IP area.' in L. Rizzi (ed), *The Structure of IP and CP. The Cartography of Syntactic Structures*. Oxford: Oxford University Press.

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.

Cheng, L. (2000). 'Moving just the feature.' In Lutz, U, Mueller, G., and Von Stechow, A. (eds), *Wh scope marking*. Amsterdam: John Benjamins**:** 77-99.

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

---- (1986). *Barriers*. Cambridge MA: MIT Press.

---- (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

---- (2000). 'Minimalist Inquiries. The Framework.' in Martin, R., Michaels, D., and Uriagereka, J. (eds), *Step by Step*. Cambridge, MA: MIT Press.

---- (2001). 'Derivation by Phase.' in M. Kenstowicz (ed), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press.

---- (2005a). ‘Beyond Explanatory Adequacy.’ in Belletti (ed), *Structures and Beyond. The Cartography of Syntactic Structures*. Oxford: Oxford University Press.

---- (2005a). ‘Three factors in language design’ *Linguistic Inquiry* 36: 1-22.

---- (2005b). *On Phases*. Ms MIT.

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

---- and Halle, M. (1968). *The Sound Pattern of English*. Cambridge: MIT Press.

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

Cinque, G. (2001). *Functional structure in DP and IP. The cartography of syntactic structures*. Oxford: Oxford University Press.

De Swart, H., and Sag, I. (2002). ‘Negation and Negative Concord in Romance.’ *Linguistics and Philosophy* 25: 373-417.

Ernst, T. (2002). *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.

Frege, G. (1892). ‘Sinn und Bedeutung.’ *Zeitschrift für Philosophie und philosophische Kritik* 100: 25-50.

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

Gianollo, C., C. Guardiano, and LonmgobardiG. (2007). ‘Three fundamental issues in parametric linguistics.’ to appear in T.

Biberauer, Holmberg, A. and Roberts, I. (eds), *Limits of parametric variation*. Amsterdam: John Benjamins.

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

Grohmann, K., Hornstein, N., and Nunes, J. (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.

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

---- and Zanuttini, R. (1991). 'Negative Heads and the NEG-Criterion.' *The Linguistic Review* 8: 233-51.

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

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

Hendriks, H. (1993). *Studied Flexibility*. PhD Dissertation, University of Amsterdam.

Hinzen, W. (2006). *Mind design and minimal syntax*. Oxford: Oxford University Press.

Hopper, P., and Traugott, E. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Iatridou, S. (1990). 'The Past, the Possible and the Evident.' *Linguistic Inquiry* 21: 123-9.

Janssen, T. (1983). *Foundations and Applications of Montague Grammar*. Amsterdam: Mathematisch Centrum.

Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. (2000). *Parameters and Universals*. Oxford: Oxford University Press.

Koeneman, O. (2000). *The flexible nature of verb movement*. Utrecht: LOT Publications.

Ladusaw, W. (1992). 'Expressing negation.' in Barker, D and Dowty, D (eds), *Proceedings of SALT II*. Cornell, NY, Cornell Linguistic Circle: 237-259.

Laka, I. (1990). *Negation in Syntax: On the Nature of Functional Categories and Projections*. Cambridge, MA: MIT Press.

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

Lechner, W. (2007). 'Interpretable effects of head movement.' *Lingbuzz* 000178.

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

McCarthy, J. (1986). OCP Effects: Gemination and antigemination. *Linguistic Inquiry*, 17: 207– 263.

McCawley, J. (1968). *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.

McDaniel, D. (1989). 'Partial and Multipe Wh-Movement.' *Natural Language and Linguistic Theory* 7: 565-604.

Merchant, J. (2006). 'Why no(t).' *Style* 20: 20-23.

Moro, A. (2000). *Dynamic antisymmetry*. Cambridge, MA: MIT Press.

Moro, A. (2004). 'Linear compression as a trigger for movement.' in Van Riemsdijk, H., and Breitbarth, A. (eds), *Triggers*. Berlin, Mouton de Gruyter.

Neeleman, A., and Van der Koot, H. (2006). 'On syntactic and phonological representations.' *Lingua* 116: 1524-1552.

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.' **Lingbuzz** 000248.

Nilsen, O. (2003). *Eliminating Positions*. Utrecht: LOT Publications.

Partee, B. (1984). 'Compositionality.' in Landman, F., and Veltman, F. *Varieties of Formal Semantics*. Dordrecht, Foris**:** 281-312.

Pesetsky, D., and Torrego, E. (2001). 'T-to-C movement: causes and consequences.' in M. Kenstowicz (ed), *Ken Hale: a Life in Language*. Cambridge, MA: MIT Press.

Pesetsky, D. and Torrego, E. (2006). 'Probes, goals and syntactic categories.' in Otsu, Y, (ed), *Proceedings of the 7th Annual Tokyo Conference on Psycholinguistics*.

Pollock, J.-Y. (1989). 'Verb Movement, Universal Grammar, and the Structure of IP.' *Linguistic Inquiry*. 20: 365-424.

Potts, C. (2002). 'No vacuous quantification constraints in syntax.' in, Hirotani, M. (ed), *Proceedings of the North East Linguistic Society 32*. Amherst, MA, GLSA**:** 451-470.

Ramchand, G. (2007). *First Phase Syntax*. Ms. Univeristy of Tromsoe.

Rizzi, L. (1986). 'Null Objects in Italian and the Theory of pro.' *Linguistic Inquiry* 17: 501-557.

---- (1997). 'The fine structure of the left periphery.' In Haegeman, L. (ed), *Elements of Grammar: Handbook in Generative Syntax*. Dordrecht: Kluwer.

---- (1990) *Relativized Minimality*. Cambridge MA: MIT Press.

---- 'Relativized Minimality Effects.' in Baltin, M., and Collins, C. (eds), *The handbook of comtemprorary syntactic theory*. Oxford: Backwell**:** 89-110.

---- (2004). *The structure of CP and IP, The cartography of Syntactic Structures*. Oxford: Oxford University Press.

Roberts, A., and Holmberg, A. (2005). 'On the role of parameters in universal grammar. A reply to Newmeyer.' in Broekhuis, H. Corver, N., Huybregts, R, Kleinhenz, R., and Koster, J.

Organizing Grammar. Linguistic Stusies in honor of Henk van Riemsdijk. Berlin: Mouton de Gruyter.

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

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

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

Szabó, Z. (2000). 'Compositionality as Supervenience.' Linguistics and Philosophy 23: 475-505.

Travis, L. (1984). *Parameters and Effects of Word Order Variation*. Cambridge, MA: MIT Press.

Truckenbrodt, H. (2006). 'On the semantic motivation of syntactic verb movement to C in German.' *Theoretical Linguistics* 32: 257-306.

Von Stechow, A. (2002). 'Temporal Prepositional Phrases with Quantifiers: Some Additions to Pratt and Francez (2001).' Linguistics and Philosophy **25**: 755-800.

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

Wiklund, A.-L. (2005). *The syntax of tenselessness: on copying constructions in Swedish*.PhD Dissertation University of Lund.

Zanuttini, R. (1997). Negation and Clausal Structure. A comparative study of Romance languages. New York/Oxford: Oxford University Press.

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

Zeijlstra, H. (2006). 'The Ban on True Negative Imperatives.' *Empirical Issues in Formal Syntax and Semantics* 6.

Zeijlstra, H. (2007). 'On the syntactic flexibility of formal features.' to appear in T. Biberauer, Holmberg, A. and Roberts, I. (eds), *Limits of parametric variation*. Amsterdam: John Benjamins.

Zeijlstra, H. (2007). 'Modal Concord.' in Gibson, M., and Friedman, T. (eds), *Proceedings of Semantics and Linguistic Theory 17 (SALT 17)*. Ithaca, NY: CLC Publications.