Hard and soft conditions on the Faculty of Language: constituting parametric variation

Zeijlstra, H.

Publication date
2008

Published in
Linguistics in Potsdam

Hard and soft conditions on the Faculty of Language: constituting parametric variation [*]

Hedde Zeijlstra

University of Amsterdam

In this paper I argue that both parametric variation and the alleged differences between languages in terms of their internal complexity straightforwardly follow from the Strongest Minimalist Thesis that takes the Faculty of Language (FL) to be an optimal solution to conditions that neighboring mental modules impose on it. In this paper I argue that hard conditions like legibility at the linguistic interfaces invoke simplicity metrices that, given that they stem from different mental modules, are not harmonious. I argue that widely attested expression strategies, such as agreement or movement, are a direct result of conflicting simplicity metrices, and that UG, perceived as a toolbox that shapes natural language, can be taken to consist of a limited number of markings strategies, all resulting from conflicting simplicity metrices. As such, the contents of UG follow from simplicity requirements, and therefore no longer necessitate linguistic principles, valued or unvalued, to be innately present. Finally, I show that the SMT does not require that languages themselves have to be optimal in connecting sound to meaning.

Keywords: Parameters, Simplicity, Complexity, Uninterpretablity, Agreement, Movement

1   Introduction

Following current minimalist reasoning, language is thought to be a perfect system connecting sound and meaning (cf. Chomsky 2000, 2001, 2005a,b,

Lasnik 2003). The strongest formulation of this idea is The Strong Minimalist Thesis (SMT): Language is an optimal solution to interface conditions that the Faculty of Language (FL) must satisfy (Chomsky 2005b: 3).

However, the idea that language is some kind of a perfect solution seems to be at odds with the huge amount of cross-linguistic variation that can be attested. This leads to the following question: if language is an optimal solution to interface conditions that the Faculty of Language (FL) must satisfy why would not all languages be morpho-syntactically uniform?

Implementing this question within the Principle and Parameters model, initiated by Chomsky 1981, that takes cross-linguistic variation to be the result of a relatively small amount of parameters to be set during the process of language acquisition, the question rises as to why parameters should exist in the first place?

In this paper I argue that parametric variation is not incompatible with the SMT. Instead, I argue that the SMT, given that it takes language to be an optimal solution to conditions that are imposed on FL by *different* mental modules, allows for multiple solutions as long as these are all optimal. If the SMT allows for multiple solutions it would even require additional explanation why natural language would not exhibit cross-linguistic variation.

However, a question that then immediately rises is whether all languages are actually equally simple? If two languages both form an optimal solution in the task of relating sound to meaning, one language is expected not to be more complex than the other, since otherwise the simplest solution would the only optimal one.

Although the idea that languages are equally complex has been proposed by a number of scholars (see for an overview and discussion Degraff 2001), this view is far from being uncontroversial. In a number of recent proposals (e.g.

Kusters 2003, Gil 2001, Ramchand and Svenonius 2006) it has been argued that languages actually differ with respect to their internal complexity.

In this paper, I argue that that the interplay between principles governing FL and principles governing the process of language acquisition actually allows one language to be more complex than the other.

In a nutshell, I propose that UG should be regarded as a toolbox (to use Jackendoff's 2002 metaphor) that contains different strategies for expressing semantic functions (those strategies are the tools, so to speak). The existence of these tools (why exactly these and why not any more ore any less) follows directly from the SMT. The process of language acquisition then is considered as a process where language learners detect on the basis of their language input which tool(s) are used to express each semantic function. If the target language happens to use multiple tools for the expression of a single semantic function, then the language acquirer is forced to adopt both expression strategies.

This paper is set up as follows: first, in section 2, I discuss the implications of the SMT and I argue that hard conditions applying to FL invoke soft conditions as well that take the shape of simplicity metrices. In section 3, I zoom in on one particular hard condition that applies at both the interface between FL and the Sensory-Motor system (SM) and the interface between FL and the Conceptual-Intentional systems (C-I), namely the requirement that the derivational outputs of FL are legible for both SM and C-I (dating back to Chomsky's 1986 formulation of Full Interpretation and I argue that Chomsky's later version of Full Interpretation, which bans uninterpretable features at LF (cf. Chomsky 1995), actually follows from simplicity metrices that are invoked by this legibility condition. In section 4, I argue that the simplicity metrices that are invoked by the hard condition that the derivational output must be legible both at the level of Logical Form (LF) and Phonological Form (PF) are in conflict and I demonstrate that this conflict calls syntactic operations such as Move and

Agree into being. A side effect of these assumptions is that the existence of uninterpretable features, albeit conceived slightly different from the original notion of uninterpretable features in Chomsky 1995, receives a principled explanation. In section 5, I come back to the alleged problem that different languages may exhibit different levels of complexity and I demonstrate that, contrary to what is generally assumed, combinatorial usage of different expressing strategies of semantic functions is not banned by the SMT, but actually falls out of it, given that principles that shape UG also shape the language acquisition process. Section 6 finally concludes.

2   Hard and soft conditions imposed on the Faculty of Language

I adopt, following Chomsky 1995 and subsequent work, the model in (1) that takes FL to be an autonomous mental module that is connected with other (autonomous) mental modules. In this model, FL interacts with three other mental modules: SM, C-I and the lexicon (LEX), an instance of memory.

(1)   Faculty of Language (FL) and its neighboring mental modules

The Sensory-Motor System (SM)	FL	The Conceptual-Intentional Systems (C-I)
	The Lexicon (LEX)	
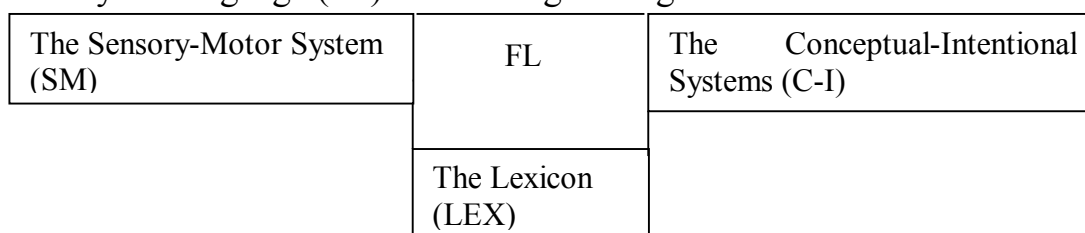
Following, Chomsky who claims that if the SMT holds 'UG would be restricted by properties imposed by interface conditions' (Chomsky 2005b: 3), FL, in the model in (1), must then be restricted by conditions that the SM system, the C-I system or LEX induce.

Both at the level of LF, i.e. the interface between the C-I system and FL, and at the level of PF, i.e. the interface between the SM system and FL, hard

conditions, such as legibility conditions, apply. Such conditions restrict the possible grammatical outcomes of the derivational process.

Although hard conditions applying to FL already severely restrict UG, the SMT not only requires that hard conditions be fulfilled, but also that they are fulfilled in an optimal way. This claim implies that different solutions to hard conditions are evaluated against simplicity metrices that evaluate possible solutions and rule out non-optimal solutions Putting this formally:

(2)   If some hard condition **C** comes along with a simplicity metric **S** and solution S_1 to fulfill **C** is a simpler solution w.r.t **S** than solution S_2, then the possible application of S_1 rules out S_2.

Hence the fact that **C** invokes **S** imposes another restriction on FL: *S_2. But it should be noted that *S_2 is not a hard condition by itself. On the contrary, S_2 is only ruled out by virtue of S_1 being a possible solution. If for some reason application of S_1 is ruled out on independent grounds, S_2 is no longer banned. In fact, S_2 in that case is even the preferred solution.

But then the question immediately rises as to what could rule out S_1, given that it optimally satisfies **C**. Two logical possibilities arise. First S_1 could violate another hard condition. In that case S_1 may never apply. But a second possibility arises as well. Suppose that not one but two hard conditions apply: **C₁** and **C₂**, both with corresponding simplicity metrices **S₁** and **S₂** respectively. Now suppose that **S₁** and **S₂** have the following forms:

(3)   **S₁**: $S_1 > S_2 > \ldots$
         S₂: $S_2 > S_1 > \ldots$

If the simplicity metrices in (3) both apply, they cannot be both optimally satisfied. Satisfying C_1 in an optimal way entails that S_1 is preferred over S_2. But preferring S_1 over S_2 entails that C_2 is not satisfied optimally. Alternatively, if S_2 is favored over S_1, C_2 is optimally fulfilled at the expense of C_1.

Note that in cases like (3) nothing requires that one simplicity metric is stronger than the other. Hence, if there is no external ground that forces optimal satisfaction of C_1 over C_2, both simplicity metrices must be equally strong. As a result, the SMT invokes two different strategies that enable FL to optimally satisfy interface conditions imposed on it.

The question then arises as to whether situations like (3), where two optimal solutions cancel each other out, are natural or expected on conceptual grounds. The answer to this question is univocally yes. Especially since mental modules neighboring FL are (semi-)autonomous, it would in fact require independent motivation if all simplicity metrices induced by hard interface conditions were in harmony. Nothing guarantees that different cognitive systems like the C-I and the SM systems work in such a way that the conditions they impose on FL are identical with respect to the way that they be optimally fulfilled. Of course nothing rules out conditions that do not face any contradictory simplicity metric and such conditions will always be optimally satisfied, but since not all conditions are in harmony, variation is already called into being.

The general and most radical hypothesis following from this line of reasoning is that the entire range of cross-linguistic variation results from conflicting simplicity metrices induced by different hard interface conditions imposed on FL. In this paper I argue that cross-linguistic variation with respect to two expressing strategies of semantic functions, morphological marking and (head) movement, are a direct result of the fact that the SMT constitutes multiple strategies for FL to fulfill hard interface conditions in an optimal way.

3   Full Legibility and Full Interpretation

As discussed above, one hard condition that is imposed on FL is that derivational outputs must be legible to the respective interpretational systems at the levels of the interfaces. At LF the derivation must be legible for the C-I system, and at PF it must be legible for the SM system. In this section I argue that this hard condition induces a weaker version of Chomsky's 1995 formulation of the Principle of Full Interpretation and that the current stipulative formulation of this principle is too strong.

3.1   Full legibility and the C-I interface

Let me formalize the hard condition that derivational outputs must be legible at the level of interfaces by introducing the Principle of Full Legibility (PFL):

(4)   *Principle of Full Legibility (PFL)*: the derivational output of FL must be fully legible for any interpretational system for which such an output forms an input.

PFL is of course reminiscent of Chomsky's Principle of Full Interpretation (PFI), but it is a weaker notion. It only requires LF representations to be legible, nothing more. In this sense it crucially differs from Chomsky's original formulation of PFI (Chomsky 1986), which is also meant to rule out vacuous quantification. But, as Potts 2002 has demonstrated, the ban on vacuous quantification is not a necessary constraint on syntactic structures and therefore does not have to follow from any hard conditions applying to FL.

   PFL is also weaker than Chomsky's 1995 version of PFI, which states that every element at LF must receive interpretation:

(5)   *Principle of Full Interpretation (PFI)*: every element of an output
        representation should provide a meaningful input to the relevant other
        parts of the cognitive system (after Chomsky 1995).

The main difference between PFL and this version of PFI is that PFL states that
every (part of a) syntactic object must be legible to the C-I systems, whereas PFI
requires that every (part of a) syntactic object must have semantic content.
However, legibility does not presuppose semantic content.

    To illustrate this, take for instance the tree in (6):

(6)

```
              E
            /   \
          C       D
                /   \
              A       B
```
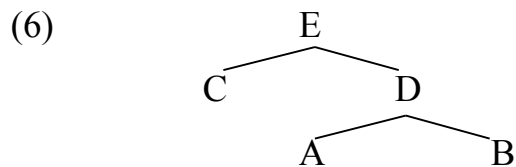
Now suppose that A is semantically empty, i.e. it contains only formal features
at LF. In that case the denotation of D is identical to the denotation of B. If no
other grammatical ban is violated and D can be a semantic complement of C (or
vice versa), nothing renders (6) illegible at LF. Following PFL (6) is ruled in.
Hence, the PFI condition that rules out semantically empty elements in syntactic
representations at the level of LF does not follow from any legibility condition
and therefore counts as a stipulation.

    Note that in a way such a stipulation is even counterintuitive. Saying that the
presence of some element blocks the interpretation of a structure that would
otherwise receive a proper interpretation at LF presupposes that the presence of
such an element has interpretational effects and as such it cannot be said to be
fully uninterpretable.

    The question now rises what kind of elements have such properties that they
can appear at LF without adding anything to the semantic interpretation. Note

that traces, if they are perceived as copies, do have semantic content. In standard semantic theory they are considered to be variables (cf. Heim & Kratzer 1998, Sportiche 2005), whereas unininterpretable features in the sense of Chomsky 1995, 2000, 2001 are said to be free from semantic content.

Note that in their very essence uninterpretable formal features ([uF]'s henceforward) only drive syntactic operations and are strictly speaking only formal in nature: it is a formal requirement that at some point in the derivation they must stand in some particular configurational relation with an interpretable counter feature ([iF]). [iF]'s are both formal and semantic in nature (formal in the sense that they can establish so-called agree relations with [uF]'s, semantic because they are non-vacuously interpreted at LF), but given the fact that [uF]'s are only formal and therefore purely blind to the semantic properties of [iF]'s, it is not a semantic property of [iF]'s, but a formal property that allows it to license [uF]'s.

A major advantage on this more formal perspective on (un)interpretable features is that no look ahead problems arise, as recognition whether some element carries an [iF] or not is now taken to be part of the derivational syntactic process. Suppose that the semantically empty element A in (6) carries a feature [uF] and suppose that C carries [iF]. Than after merger of C with D all formal requirements of A have been met, even before the structure is transferred to LF.

Still, it remains an open question as to why uninterpretable features would occur at LF against the economical background of the SMT. Initially, this was the ground on which their occurrence was banned at LF, inducing the still unsolved question as to why uninterpretable features exist in the first place. The ban on uninterpretable features at LF does not follow from PFL, the hard condition that requires that derivational outputs be legible.

But if the SMT holds, it also follows that PFL is optimally satisfied. Although legibility is not affected by the presence of uninterpretable features,

their presence does not facilitate legibility either. Hence PFL induces the following soft condition:

(7)   C-I Simplicity Metric (Zeijlstra 2007):

A structural representation R for a substring of input text S is simpler than an alternative representation R' iff R contains less uninterpretable features than R'.

Following the line of reasoning sketched in section 2, the simplest solution to satisfy (7) is by banning all uninterpretable features. Then Chomsky's assumption that uninterpretable features be ultimately removed at the level of LF, now follows from the simplicity metric in (7), modulo one major difference: the C-I simplicity metric is a soft condition. If for some reason the null-option, i.e. absence of [uF]'s at LF, blocks another, equally strong, simplicity requirement, their presence may be motivated again.

In the following subsection I argue that the application of PFL at the interface between FL and SM induces a simplicity metric that prefers derivational outputs at LF that contain semantically uninterpretable features over outputs that lack them.

3.2   Full legibility and the SM interface Document Setup

PFL does not only apply at the level of LF, but also at the level of PF. Derivational outputs must be legible for the SM system and this requirement must be met in an optimal way. Equivalent to the application of PFL at LF, this means that PFL induces a second simplicity metric that, being a soft condition, bans the presence of what could be metaphorically called 'phonologically uninterpretable features', i.e. formal features without any phonological content.

Such features must have the property that they are purely formal in nature but lack phonological content, i.e. phonologically null elements.

Neeleman and Van der Koot 2006 who base themselves on Chomsky and Halle 1968 and McCawley 1968, take phonological outputs to be linear with prosodic categories thought of as phonological boundaries. The prosodic/phonological structure of sentence like (8) is thus represented as (9) where U stands for Utterance, I for intonational phrase, Φ for prosodic phrase, ω for prosodic word, F for foot and σ for syllable. U at the beginning and at the end of (9) means that the sentence be preceded and followed by an intonational break.

(8)   John's father suggested a two-seater but John's mother preferred a fur coat

(9) 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.

Neeleman and Van der Koot argue that prosodic categories are hierarchically ordered (from weak to strong) as in (10).

(10) $\sigma < F < \omega < \Phi < I < U$

Prosodic categories, perceived as prosodic boundaries, are thus not banned from phonological representations, but their occurrence should be as limited as possible, since their appearance cannot be motivated in terms of phonological legibility either. Hence, PFL must invoke the following simplicity metric.

(11)   SM Simplicity Metric:

A formal representation R for a substring of input text S is simpler than an alternative representation R' iff R contains less prosodic boundaries than R'.

Hence, both uninterpretable formal features and prosodic boundaries are dispreferred by PFL, given the C-I and SM simplicity metrices. In the next section, I demonstrate that these metrices are in conflict and that for that reason the C-I and SM simplicity metrices can never be optimally satisfied at the same time.

4   Conflicting simplicity metrices

In this section I demonstrate that the C-I and SM simplicity metrices cannot be optimally fulfilled at the same time. That is to say, expressing a semantic function without using uninterpretable features must lead to the introduction of prosodic boundaries, and alternatively, expressing such a semantic function without such prosodic boundaries will inevitably lead to an introduction of a semantically uninterpretable feature. This immediately leads to the question as to why expression strategies for semantic functions that lack both uninterpretable features and strong prosodic boundaries are forbidden.

In its very essence, the answer to this question is the following: the main effect of the SM Simplicity Metric is to spell out as much possible on one and the same lexical node, but semantic functions cannot occupy any arbitrary position in the syntactic structure and require uninterpretable features to ensure possible interpretation at LF.

Let me illustrate this by discussing past tense, which is subject to cross-linguistic variation with respect to the way it is expressed. One way to express past tense is by using a single word for it, something like *past*, as is the case in

expression strategy (12). Under this strategy, there is a 1:1 correspondence between the word for past tense and the semantic past tense operator. Although such a strategy is not very frequently attested, several languages, e.g. West-Greenlandic (cf. Bittner 2005), express past tense in such a way.

However, in many languages, and English is no exception to this observation, past tense is expressed by means of a temporal affix rather than a temporal adverb that, being a prosodic word, would stand on its own. English thus prefers a different expressing strategy (the one in (13)) to the West-Greenlandic type of strategy in (12).

(12)   Wolfgang *past* play tennis

(13)   Wolfgang play-*ed* tennis

But under this strategy semantic past tense no longer follows directly from the temporal morphological marker *-ed*, since the semantics of past tense does not allow for a direct interpretation in the position where the lexical verb *play-ed* is base-generated. Past tense is a semantic operator that must outscope the entire vP, i.e. the fully saturated argument structure of the predicate (see e.g. Klein 1994, Ogihara 1996, Abush 1997, Kratzer 1998, Von Stechow 2002). This is illustrated in (14).

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

   = 'For every Sunday in Past$_c$ 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'

The only available reading is the one where past tense outscopes the distributive quantifier *every Sunday*, which in its turn outscopes the lexical verb *play*. Consequently, past tense affix *-ed* therefore cannot be assigned the semantics of the past tense operator in the position that it occupies at surface structure. But what then is the contribution of *–ed*, if it cannot be interpreted at surface structure? How can *–ed* induce the semantics of past tense, if at the same time it cannot be interpreted at its base position in the sentence?

Two logical possibilities arise: either (i) *-ed* is not the semantic past tense operator itself, but merely a true marker of an abstract operator that is responsible for the semantics of past tense; or (ii) *-ed* is the semantic past tense operator itself, but a structural transformation takes place such that both *–ed* and *play* can be interpreted in the proper position. The two strategies can be tentatively called *marking* and *displacement* and they way they function is sketched in (15).

(15)   a.   (Marking)

Op_{PAST}

… play-*ed* …

b.   (Displacement)

-ed

… play- …
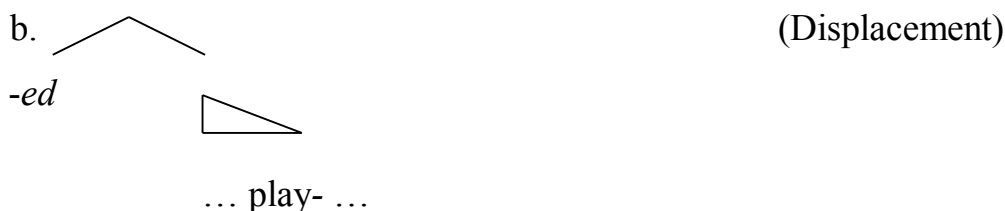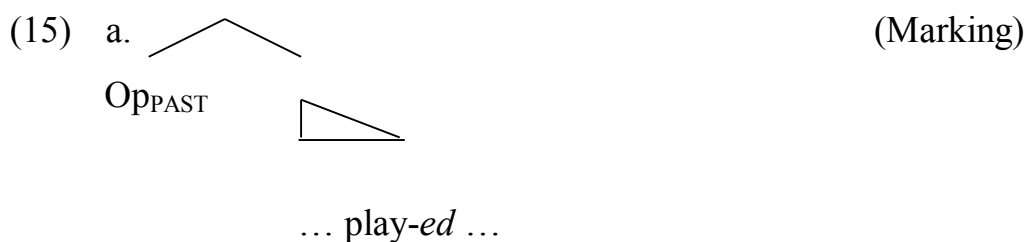
In (15a) *-ed* is a marker that signals the presence of the past tense operator in its proper position, i.e. above VP; in (15b), due to the transformational operation that has been applied, *-ed*, itself being the carrier of the semantic contents of the

past tense operator, is now in a position where it can be properly interpreted and as a result of the same transformational operation, be disconnected from the position where the lexical content of the verb *play* is interpreted.

In the following two subsections section I demonstrate that in both cases the presence of an uninterpretable feature is required. In a nutshell, I argue that in (15a) *–ed* must carry an uninterpretable past tense feature, marking the presence of an abstract operator carrying an interpretable past tense feature. And I argue that in (15b) we find two copies of the finite verb, whereas only one gets interpreted. For that reason, at least one of the two verbs may be interpreted as carrying verbal contents; the other copy must be analyzed as carrying an uninterpretable verbal feature.

If those analyses are correct, the SM Simplicity Metric favoring expression of as much material as possible on the same lexical node, can then only be maximally satisfied at the expense of the C-I Simplicity metric which bans uninterpretable features. Vice versa, a strategy as in (12) which can be directly interpreted at LF without any rescue strategy that requires uninterpretable features, violates the SM Simplicity metric, as it introduces new prosodic boundaries. Thus, the interplay between the SM and the C-I Simplicity Metric already gives rise to two different types of strategies: one type that prefers prosodic boundaries over uninterpretable features and one type of strategies where uninterpretable features are preferred over prosodic boundaries.

4.1 Marking, Uninterpretablity and Agree

Now let's zoom in at the marking strategy exemplified in (16a), where a marker only indicates the presence of an abstract operator in the appropriate position. More abstractly, this means that some root X is equipped with an additional marker. Such a marker can be an affix, but it does not necessarily have to be one: vowel alternation or other instances of marking (e.g. syncretisms of

multiple markers) are equally well possible. Let us call the marker F. In the case of affixation, a root plus marker is thus of the form X-F.

As discussed before, F is not the carrier of the semantic contents of the operator. The structure of a sentence containing X-F is rather like (16), where a covert operator (Op_F) is responsible for the semantic contribution, which is manifested by F.

(16)   (Marking)

Op_F

… X-F …
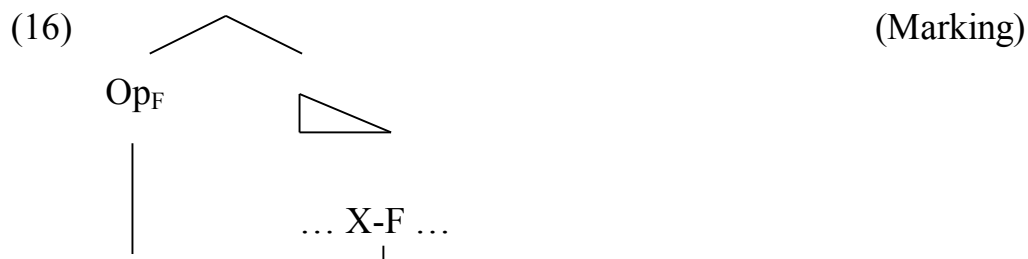
This structural relation in (16) are governed by the following three conditions:

(17)   a.   [Op_F [… X-F …]]

      b.   *[… X-F …]]

      c   *[Op_F [… X …]]

The conditions in (17) state that F demands the presence of an operator Op_F and that abstract Op_F may only be included if F is present. The conditions in (17) are an implementation of what Ladusaw refers to as a mechanism of self-licensing (Ladusaw 1992). The abstract operator is licensed by the presence of Op_F and Op_F fulfils the licensing requirements of F. Since Op_F is abstract, marking strategy (16) serves the SM Simplicity metric in the sense that only an affix is sufficient to express past tense and that no new morphological word needs to be

included. At the same time, this marking strategy needs the formal properties that F exhibits. Hence, the question rises as to what properties does F actually exhibit, such that the conditions in (17) follow?

To recapitulate, F must be morpho-syntactically visible, F may only occur in a grammatical sentence while standing in a syntactic relation with Op_F; and F must be se semantically empty. The reader will recall from the previous section that these are exactly the properties that define uninterpretable formal features. Thus, F must be an uninterpretable formal feature [uF]. In other words, it is only possible to mark a semantic function by means of an (affixal) marker, that itself does not contain any semantic contribution, if that (affixal) marker itself is the carrier of an uninterpretable feature in the Chomskyan sense.

The idea that F carries an uninterpretable formal feature directly entails the conditions in (17). Conditions (17a) and (17b) follow directly, but also (17c) is a consequence of this implementation: if a sentence is grammatical and its grammaticality is not due to any of its overt elements, then a covert element must be responsible for its grammaticality; if the grammaticality of a sentence follows directly from its overt elements on the other hand there is no ground for adopting abstract material. Note that this is in its very essence a truism.

To sum up, the conditions under which an element F may mark the presence of an abstract matching operator, without contributing to the semantics of the sentence in which F occurs itself, follow immediately once it is assumed that F carries an uninterpretable feature [uF] that matches with an interpretable feature [iF] on the operator.

4.2   Displacement, Remerge and Move

Marking strategies (where elements carrying uninterpretable formal features signal the presence of matching abstract operators) are not the only ways to

enable spell-out of semantically mismatching elements on one and the same morphological word, as favored by the SM Simplicity Metric. Another way would be to induce a displacement effect, such that F is semantically non-empty, and therefore does not contain any uninterpretable features, but takes scope from a different position than where X has been base-generated.

(18)   (Displacement)

~~X~~-/F/

… X-/~~F~~/ …
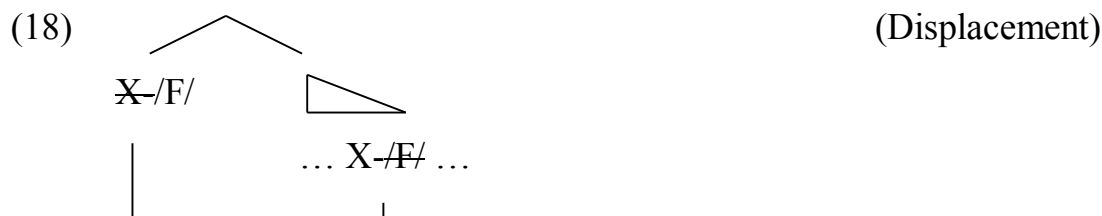
This is of course reminiscent of Chomsky's copy theory of movement or chain formation is the sense of Brody 1995. Before continuing the argument that movement is motivated by the semantic content on a particular lexical item, which can only be interpreted in a higher position, let me avoid one possible misunderstanding. The existence of movement itself does not have to be motivated. As Chomsky 2005 has argued for numerous times, a generative operation like Merge can apply internally and therefore the operation Remerge is pregiven by Merge. However, movement is derivationally complex, and is therefore ruled out if it is unmotivated. Hence what needs to be motivated is the trigger for movement, not movement itself.

Suppose a root again takes the form X-G, where X-G now means that it is the realization of two elements, X and G, that carry both semantic content, but only X can be interpreted in situ, G cannot. One solution then would be to remerge X-G and create a higher position where G is interpreted.

(19) PF: $[\ldots X\text{-}G\ldots] \rightarrow [\ldots\ X\text{-}G\ \ldots\ [\ldots\ \cancel{X\text{-}G}\ \ldots]]$

LF: $[\ldots X\text{-}G\ldots] \rightarrow [\ldots\ \cancel{X}\text{-}G\ \ldots\ [\ldots\ X\text{-}\cancel{G}\ \ldots]]$

The representation in (19) is a direct result of the copy theory of movement, where first a lexical item has been copied, and then all doubly manifested material is deleted once, either in the highest or in the lowest position. In this case the highest copy is phonologically interpreted in the highest position (and deleted in the lowest position), but at LF X is deleted in the highest position and G in the lowest position. Note that this is the only structural representation that is legible at LF. All other combinations (X-G interpreted in the same position, or G interpreted below and X above) would be illegible at LF. This is exactly what has been observed in the case of past tense. A past tense operator must be interpreted outside the vP; the lexical content of a verb must be interpreted vP in situ.

However, (19) cannot be the correct derivational outcome yet, since G lacks a syntactic category. In the example G is nothing but a purely semantic past tense operator that does not carry any formal feature at all. Therefore, if X gets deleted in the highest copy, no formal feature is left over and G would not even be a syntactic object then. Syntax would be completely blind to it. Even if X is interpreted in the lowest position, it must still be syntactically visible in the highest position. A moved noun keeps the syntactic status of a noun; a moved verb the syntactic status of a verb, etc.

Hence the picture in (19) can not be complete. Although semantically X is only present downstairs, X must be formally present upstairs, without receiving any interpretation. In other words, X must be an uninterpretable formal feature in the highest position, and in interpretable formal feature downstairs. The

representation of semantically driven movement of X-G, due to G's semantic requirement to be interpreted in a higher position, must be as in (20).

(20)   […X-G…] $\rightarrow$ [… [uX]-G … [… [iX]-~~G~~ …]]

This (simplified) view on movement takes movement to be semantically movement. As it has standardly been assumed that head movement does not affect semantic interpretation, the burden of evidence is actually to demonstrate that head movement is indeed semantically motivated.

Applying these ideas to the expression of semantic tense, V-to-T movement can then be the consequence of the fact that the past tense morpheme is actually the carrier of past tense. Movement of the verb then results in interpretation of the past tense operator in the highest position and of the verbal contents in the lowest position.

(21)   […V-PAST…] $\rightarrow$ LF: [… [uV]-PAST … […[iV]-~~PAST~~ …]]

The idea that movement is essentially triggered by semantic properties rather than by morpho-syntactic requirements is reminiscent of foot-driven movement analyses (though these analyses have never been based on semantic motivations), such as Platzack 1996, Koeneman 2000 and Van Craenenbroeck 2006.

This view on movement is also supported by a recent analysis by Truckenbrot 2006, who argues that V-to-C movement activates speech act operators and is thus semantically driven. Implemented in the proposal above, V_{fin} carries initially a purely semantic feature speech act feature and a formal feature V. As the feature that has the illocutionary force of a speech act cannot

be interpreted on V°, it must move to a higher position. Given that all operators encoding illocutionary force have to precede all elements carrying propositional contents, speech act formation is easily (but not necessarily) executed by verbal fronting. V_{fin} then copies itself and the speech act is interpreted in the highest head position and the verbal contents are interpreted below. This is shown in (23) below where the phonological, syntactic and semantic representations are given for the imperative sentence in (22).

(22)   Kill Mary!

(23)   SEM   IMP(Kill(Mary))

         SYN   [V-fin$_{[uV]IMP}$] … [$_V$ ~~V-fin~~$_{[iV][IMP]}$ D]]

         PHON   /Kill Mary … /

Move is then, similarly to Agree, is a marking strategy that is imposed on FL by the SM interface condition to express as much material as possible on one and the same lexical node. This condition can only be fulfilled if natural language exhibits uninterpretable material.

     Note that this view on head movement unifies head movement with other types of movement (A movement, A' movement) in the sense that head movement is now also an instance of pied-piping. The formal features of V constitute the vehicle that allows G to move.

4.3   Concluding remarks

In this section I hope to have shown that both agreement and movement strategies are can only be realized if uninterpretable features are involved and that uninterpretable features are motivated because expressing strategies involving uninterpretable features can be equally optimal as expressing

strategies (such as (12)) that lack them for the very reason that reduction of prosodic boundaries can only be established by means of inclusion of uninterpretable features. Purely external merge based strategies, internal merge-based strategies and agreement strategies are thus all tools that are directly motivated by the SMT and thus constitute the UG toolbox.

Note that by no means these three strategies are exhaustive. I only demonstrated that as a result of the SMT, which invokes both the C-I and the SM simplicity metric, these strategies are called into being. Other simplicity metrices, e.g. LEX simplicity metrices could induce additional possible expressing strategies.

In any case, this line of reasoning has, I think, two major benefits. It gives a principles explanation for the existence of uninterpretable features, up till now an unsolved problem in minimalist theory; and it also gives a motivated answer to the question as to why movement is triggered, which does not rely on stipulated notions such as EPP-features.

5   Grammatical simplicity and the parametric space

So far, the proposal explains why different types of expressions for a particular Op_F exist and are cross-linguistically attested. For instance, it provides an answer to the question why uninterpretable features exist in the first place.

However, it cannot be taken to say that all languages are simplest solutions, i.e. that all languages are maximally simple, and thus select exactly one expression strategy for each semantic function. Let me illustrate that with the following examples from Afrikaans, Italian and German.

(24)   Ek sing   Afrikaans
         I sing

'I sing'

(25) Canto   Italian

Pro$_{[i1SG]}$ sing$_{[u1SG]}$

'I sing'

(26) Ich singe   German

I$_{[i1SG]}$ sing$_{[u1SG]}$

'I sing'

Afrikaans has a C-I-based strategy to express pronominal subjecthood, Italian an SM-based strategy (Agree). Both seem to be equally simple in that respect. However, German exhibits both. It has both an Agree strategy and a C-I strategy.

As has often been observed, language seems to be suboptimal rather than optimal, often used as an argument against simplicity as an underlying force in grammar. So the question rises as to how examples like (26) can be accounted for against the background of the SMT?

As discussed before, FL being an optimal solution to conditions imposed on it by neighboring mental modules, constitutes UG, i.e. the linguistic toolset. This toolset consist of a number of optimal solutions to connect form and meaning.

At the same time, FL drives language acquisition. In fact, it is even the backbone of generative theory that principles that shape UG also govern language acquisition. Simplicity metrices therefore do not only constitute possible grammars, but also guide language learnability.

Simplicity metrices applying to the language acquisition process ensure that the simplest grammar is selected during the language learning process. However, this only explains why languages are not maximally simple, but just

as simple as their target language. If for some reason the target language is not maximally simple, but takes for instance two marking strategies to express a single semantic function, than the language learner can do nothing but assign those two marking strategies to his or her own grammar.

The only question is why target languages should be non-optimal given the SMT. Note that the existence of such languages is not expected on the ground of the SMT. However, nothing excludes that external factors may play a role as well. Effects that are due to L2 acquisition in situations of language contact often cause language change effects. Thus two maximally optimal languages may interact and yield a language that ultimately is less optimal. This has been the case for instance with German.

Weerman 2006 argues that proto-Germanic was a pro-drop language, just like current Italian, and that processes of language contact led to a process of deflection in the verbal paradigm: some forms eroded and the 1:1 relation between person and verbal agreement disappeared. As a result (cf. Rizzi 1986 and Neeleman and Szendroi 2006) pro-drop was no longer licensed. On the other hand, the language contact situation did not go so far that all distinctions in the paradigm were gone, as is for instance the case with Afrikaans. German is somewhere in between a full pro-drop paradigm and the Afrikaans zero-paradigm. Consequently, it exhibits both C-I-biased and SM-biased marking strategies for pronominal subjecthood.

German language learners will adopt both strategies, since this is the simplest way to satisfy the simplicity metrices. Only if the language input will undergo total deflection, the language learner will adopt the single external-merge strategy.

On the basis of this line of reasoning, a new view of cross-linguistic variation can be presented, which includes both the notion of possible and probable languages (see Newmeyer 2005 for discussion). The grammatical

space is dynamic and governed by simplicity metrices that are both upward entailing (a set of expression strategies allows application of multiple strategies) and downward entailing (select the smallest number of strategies possible for each semantic function). At the same time, every grammar that exhibits expressing strategies that are not part of UG is impossible. This view on the grammatical space, or as it is structured, on the parametric space can be modeled in (27).

(27)   The parametric space:

Set of simplest grammars

Set of not maximally simple grammars

Simplicity metrices applying
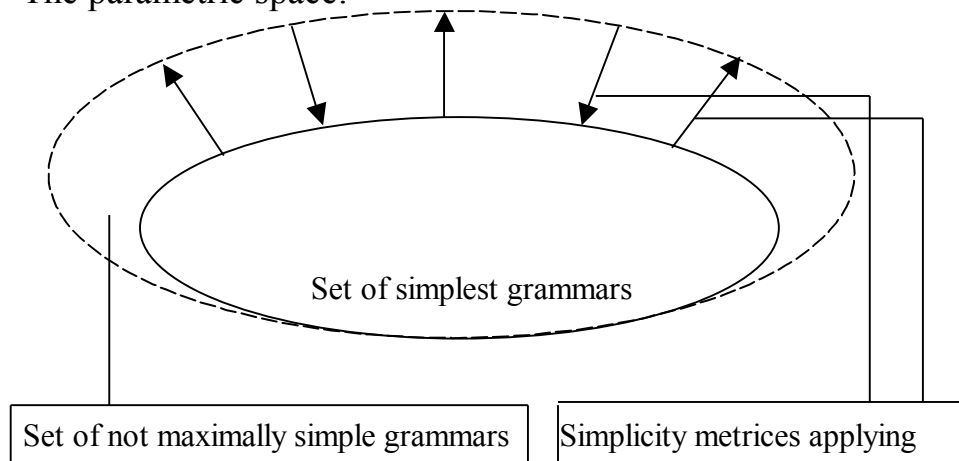
Note that this view on language allows for parametric variation following directly from the SMT. Hence parameters do not need to be thought of as innately present, but can be taken to be derived from the idea that FL is maximally simple. Moreover, this view ensures that even though FL is taken to be maximally simple, this does not entail that every possible grammar has to be simple as well.

6   Conclusions

In this paper I argue that both parametric variation and the attested differences between languages in terms of their internal complexity straightforwardly follow from the Strongest Minimalist Thesis that takes FL to be an optimal solution to conditions that neighboring mental modules impose on it.

In this paper I argue that hard conditions like legibility at the linguistic interfaces invokes simplicity metrices that, given that they stem from different mental modules, do not necessarily have to be harmonious. In fact, I demonstrate that legibility at the interface between FL and the SM system and between FL and the C-I systems respectively already invokes two simplicity metrices that cannot be maximally satisfied at the same time.

I demonstrate that maximal satisfaction of the SM simplicity metric cannot take place without alluding to the notion of uninterpretable features, and that maximal satisfaction of the C-I simplicity metric bans spelling out multiple semantic functions on one and the same word.

I argue that expression strategies, such as agreement or movement, are a direct result of these conflicting simplicity metrices, and that UG, perceived as a toolbox that shapes natural language, can be taken to consist of a limited number of markings strategies, all resulting from conflicting simplicity metrices. As such, the contents of UG follow from simplicity requirements, and therefore no longer necessitate linguistic principles, valued or unvalued, to be innately present.

Finally, I show that the SMT, contrary to what has often been thought, does not require that languages themselves have to be optimal in connecting sound to meaning. Since UG drives the process of language acquisition, language acquisition can be modeled as a selectional procedure where it is detected for

each semantic function how it is to be expressed, which does no a priori require each semantic function to be expressed by one single marking strategy only.

References

Abusch, D. (1997). Sequence of Tense and Temporal de re. *Linguistics and Philosophy* 20: 1-50.

Bittner, M. (2005). Future discourse in a tenseless language. *Journal of Semantics* 22: 339–87.

Brody, M. (1995). *Lexico-Logical Form: A radically minimalist theory.* Cambridge, MA: MIT Press.

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

Chomsky, N. (1986). *Knowledge of Language.* New York: Praeger

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

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

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

Chomsky, N. (2005a). 'Beyond Explanatory Adequacy'. In Belletti (ed), *Structures and Beyond. The Cartography of Syntactic Structures.* Oxford: Oxford University Press.

Chomsky, N. (2005b). 'Three factors in language design' *Linguistic Inquiry* 36: 1-22.

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

DeGraff, M. (2001). On the origin of Creoles: A Cartesian critique of Neo-Darwinian linguistics. *Typology* 5: 213-310.

Gil, D. (2001). Creoles, complexity, and Riau Indonesian. *Linguistic typology* 5: 325-371.

Heim, I. and Kratzer, A. (1998). *Semantics in generative grammar*. Malden, MA/Oxford: Blackwell

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Klein, W. (1994). *Time in language*. London: Routledge.

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

Kratzer, A. (1998). 'More Structural Analogies Between Pronouns and Tenses'. In *proceedings of SALT VIII*. Cornell: CLC Publications.

Kusters, W. (2003). *Linguistic complexity. The influence of social change on verbal inflection*. PhD Thesis Utrecht University.

Ladusaw, W. A. (1992). 'Expressing negation'. In C. Barker and D. Dowty (eds), *Proceedings of SALT II*. Cornell, NY: Cornell Linguistic Circle.

Lasnik, H. (2003). *Minimalist Investigations in Linguistic Theory*. London: Routledge Ltd.

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

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

Neeleman,A. and Szendroi, K. (2007). 'Radical pro drop and the morphology of pronouns'. *Linguistic Inquiry* 38: 671-714.

Ogihara, T. (1996). *Tense, Attitudes, and Scope*. Dordrecht: Kluwer.

Platzack, C. (1996). 'Germanic Verb Second - Attract versus Repel: on optionality, a-bar movement and the symmetrical/asymmetrical verb second hypothesis'. E. Lang and G. Zifonun (eds), *Typologisch*. Berlin: Walter de Gruyter.

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

Ramchand, G. and Svenonius, P. (2006). 'Mapping a parochial lexicon onto a Universal Semantics'. To appear in Biberauer, Th. (ed), *The limits of syntactic variation*. Amsterdam: Benjamins.

Rizzi, L. (1986). 'Null Objects in Italian and the Theory of pro.' *Linguistic Inquiry* 1**7**: 501-557.

Sportiche, D., (2005). 'Cyclic NP Structure and the Interpretation of Traces'. In Broekhuis, H., Corver, N. ,Koster, J. ,Huybregts, R. and Kleinhenz, U. (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*. Berlin/New York: Mouton de Gruyter,.

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.

Von Stechow, A. (2005). 'Semantisches und morphologisches Tempus: Zur temporalen Orientierung von Einstellungen und Modalen'. *Neue Beiträge zur Germanistik* 4: 9-54.

Weerman, F. (2006), 'It's the economy, stupid! Een vergelijkende blik op *men* en *man*'. In M. Hüning et al. (ed). *Nederlands tussen Duits en Engels*. Leiden: SNL.

Zeijlstra, H. (2007). 'Functional structure and parametric variation: consequences of conflicting interface conditions'. To appear in K. Grohmann (ed), *InterPhases*. Oxford: Oxford University Press.

Hedde Zeijlstra
University of Amsterdam
Amsterdam Center for Language and Communication
Spuistraat134 (lsg NTK)
1012 VB Amsterdam
The Netherlands
zeijlstra@uva.nl
http:// http://home.medewerker.uva.nl/h.h.zeijlstra