Deriving the Functional Hierarchy[☆]

Gillian Ramchand and Peter Svenonius

CASTL, University of Tromsø – The Arctic University of Norway

Abstract

There is a tension between Chomsky's recent Minimalist theory and the cartographic program initiated by Cinque. Cinque's cartography argues for a large number of fine-grained categories organized in one or more universal Rich Functional Hierarchies (RFH). The subtlety of the evidence and the richness of the inventory virtually force an innatist approach.

In contrast, Chomsky argues for a minimal role for UG (MUG), shifting the burden to extralinguistic cognition, learning, and what he calls third factor principles such as principles of efficient computation. In this paper we reconcile the austere MUG vision of Chomsky with the impressive empirical evidence that Cinque and others have presented for RFH.

We argue (building on previous work) that some Cartographic work overstates the universality of the orders observed, and furthermore conflates several different sources of ordering. Ordering sources include scope, polarity, and semantic category.

Once these factors are properly understood, there remains an irreduceable universal functional hierarchy, for example that which orders epistemic modality and tense over root modality and aspect, and that which orders the latter over argument structure and Aktionsart (as discussed in much previous work).

This residual core functional hierarchy (CFH) is unexplained so far by work which follows MUG. Rather than simply stipulating the CFH as part of UG, we reconcile CFH with MUG by detailing what nonlinguistic cognition must look like in order for MUG to derive the CFH. We furthermore show how an individual language develops a language-specific RFH which is consistent with the universal CFH, illustrating with a detailed account of the English auxiliary system.

Keywords: functional hierarchy, functional categories, cartography, syntax-semantics interface, universal grammar

1. Introduction

1.1. The problem

The MINIMALIST PROGRAM strives to go beyond "explanatory adequacy" (an explanation of how language can be learned) to develop a plausible account for how human linguistic ability could have evolved (Chomsky 2005 inter alia). In this context it is conjectured that UG is sparse and minimal. Phase heads (e.g. C, and v) are the locus of important features driving derivations, and non-phase heads (e.g. T and V) are necessary for their operation. Anything else is, according to one interpretation of the Minimalist Program, unlikely to be due to UG, but must instead be due to external factors (e.g. 'general cognition').

The CARTOGRAPHIC enterprise, on the other hand, proposes to map the actually occurring functional heads in the world's languages, discovering extraordinarily rich structures in every extended projection, in every language (Cinque 1999). The impressive uniformity (variation seems to be largely restricted to the inventory of features, not their hierarchy) leads to the conclusion that the hierarchy must be based on innate factors. The hierarchy is furthermore restricted to a specific subdomain of cognition (e.g. diminutives, not 'dangerous things'), which suggests that it is part of UG.

1.2. Why it matters

Why Minimalism needs Cartography. Minimalists ignore the cartographic enterprise at their peril. It is common practice for minimalist work to posit an occasional Voice or Applicative or Focus head as needed, and to continue to assume that the sparse C-T-v-V architecture is sufficient, with minor modifications.

> Chomsky (2008:9): "C is shorthand for the region that Rizzi (1997) calls the "left periphery," possibly involving feature spread from fewer functional heads (maybe only one), ..."

But in a theory based on Minimalist principles, the flapping of butterfly wings in one place can cause a typhoon in another: When mechanisms are pared down to a minimum, each has tremendous consequences. Therefore it is vital to know what mechanisms regulate the combinations of heads beyond the phase-non-phase pairs C-T and v-V. How are features arranged at the edge? Are they contained in one or several heads? Does this arrangement bear on the order of operations? What are the properties of nonphase heads? And so on.[1]

Why Cartography needs Minimalism. Linguistic theory cannot rest on its maps. Cartography is in desperate need of a theory of the functional hierarchy. Although the data is quite rich, it seriously underdetermines the possible analyses. Are there categories which are ordered and others which are not (e.g. negation, agreement)? For those which are ordered, is there a total order or only a partial order? Can categories be missing from the middle of a sequence, or are they always present in some guise? What is the relationship among functional hierarchies in distinct extended projections? These questions cannot be answered by simple examination of the data, and require a theory.[2]

---

[1] See Shlonsky (2010) and Cinque and Rizzi (2010) for discussion of this point.

[2] Cartography is usually associated with an atomistic approach to features, where each syntactic head carries only one semantically interpretable feature (see Cinque and Rizzi 2010), but the same questions

1.3. The solution

We adopt (as working hypothesis) the Minimalist conjecture that a fine-grained hierarchy of functional heads cannot be part of UG; that is, it cannot be innate and specific to language.

We are persuaded that Cartographic work shows that there are fine-grained hierarchies of functional heads in each language, and that they are similar to each other (i.e. the clausal hierarchy of English is similar to those of Japanese, Navajo, Kîîtharaka, etc.)

We conclude that these hierarchies emerge in some highly constrained way. In this paper we offer a proposed account of how this happens.

Our approach is three-pronged. First, we adopt a fundamental triparition of the clause into a V-domain, a T-domain, and a C-domain (Platzack 2000; 2010, Grohmann 2003)[3] and provide this with a formal semantic grounding on a conceptual backdrop; we take events (e), situations (s), and propositions (p) to be conceptual primitives recruited by the language faculty, and we take the hierarchy of C > T > V to follow from the interaction of (i) the way these conceptual primitives are organized in the wetware and (ii) the way they are harnessed by the syntactico-semantic system.
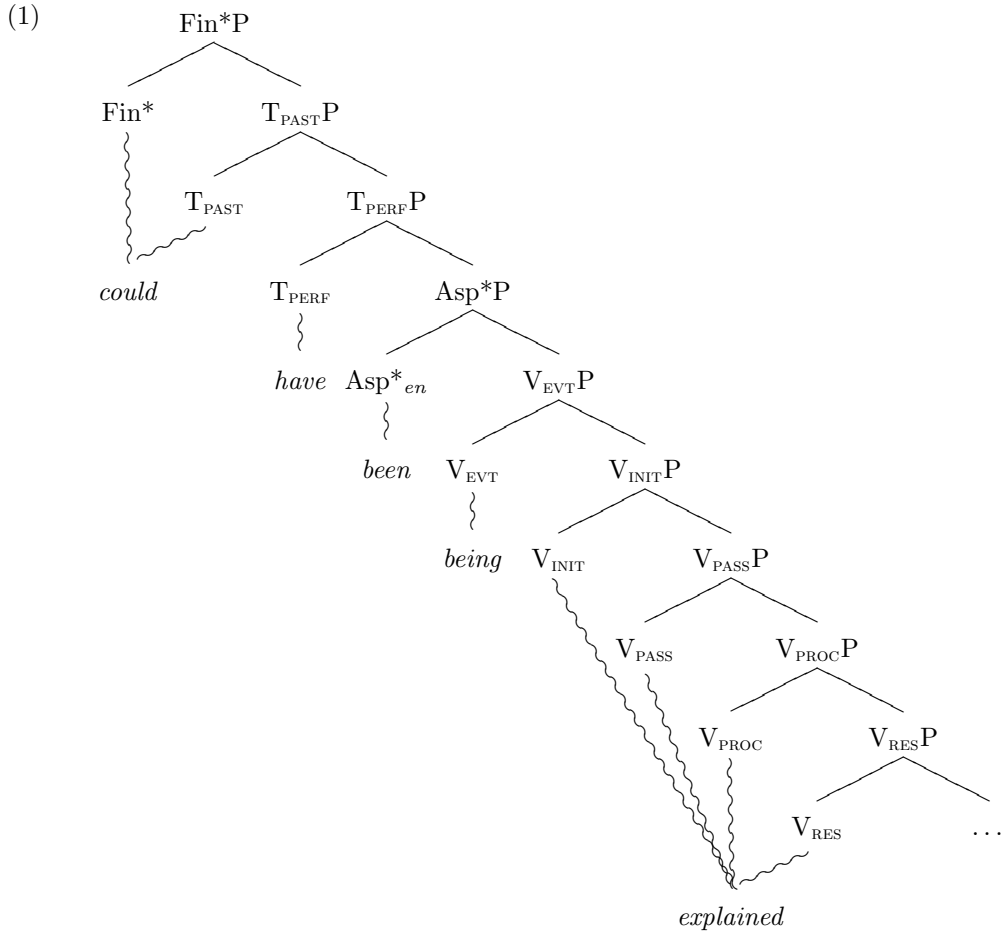
Second, we show that in some cases, the hierarchy is not in fact fixed; in other cases, there are independent factors giving rise to hierarchical effects.

Finally, we are left with a residue: Strict hierarchy which does not follow from the e-s-p triparition, nor from independent factors. For these cases we posit selectional restrictions, for example when a functor like the progressive is restricted to combining with dynamic eventualities.

To illustrate the general approach, we apply our starting assumptions and methodology to one concrete empirical domain where ordering is rich and rigid in English——classic auxiliary ordering, as treated in Chomsky (1957) and illustrated in (1).

---

we pose here apply to feature 'geometries' of the type proposed, for example, by Harley and Ritter (2002) and Cowper (2005), when those are construed as constraints on feature bundles.

[3]Wiltschko (to appear) proposes four domains: discourse linking, anchoring, point of view, and classification, lining these up approximately with C, T, Asp, and V respectively. In §3 we discuss the relationship between Wiltschko's domains and the ones we propose.

(1)

```
                        Fin*P
                       /    \
                   Fin*      T_PAST P
                    {        /      \
                    {    T_PAST      T_PERF P
                    {              /        \
                  could       T_PERF        Asp*P
                               {          /      \
                             have     Asp*_en      V_EVT P
                                       {          /      \
                                     been     V_EVT       V_INIT P
                                               {         /       \
                                            being    V_INIT       V_PASS P
                                                      {          /       \
                                                      {      V_PASS       V_PROC P
                                                      {       {          /       \
                                                      {       {      V_PROC       V_RES P
                                                      {       {       {          /     \
                                                      {       {       {       V_RES      ...
                                                      {       {       {        {
                                                         explained
```

The diagram in (1) shows two different kinds of elements, which are often not sharply distinguished. First, there are the syntactic elements, Fin[iteness], T[ense], Asp[ect], and V[erb], which stand in dependency relations created by Merge as reflected by the straight lines. The observed hierarchy constraining these dependency relations, such that for example Asp*$_{en}$ may dominate V$_{EVT}$ (Perf[ect] can take Prog[ressive] as a complement) but not vice versa, is the central explanandum of this paper.[4]

The other kind of element included in the diagram are the exponents, *might*, *have*, *been*, *being*, and *explained*. We use squiggly lines to represent exponence, a relation distinct from complementation; an exponent may spell out or lexicalize one or more

---

[4]The details of the labels we employ, and their denotations, will be taken up in section 4. In particular, note that V$_{PASS}$ is the label for the passive participial ending *-ed*, and we will locate it immediately above V$_{PROC}$, or big V. The other non-standard label here is V$_{EVT}$, which for reasons that will become clear later, we use as the label for the *-ing* participle ending. At this point the labels are not crucial, we simply observe that there is a rigid ordering of Perf over Prog over Passive which needs to be accounted for.

4

syntactic nodes.

Following Chomsky (1957) and subsequent work, a single word like *interviewed* can be associated with more than one element in the syntactic tree. Thus *interviewed* is the spell out of both V_{PASS} and the various thematic subcomponents of V; as is typical for words, suffixes (*-ed*) spell out structure higher than what are suffixed to.

We also assume that monomorphemic items like *might* can spell out more than one syntactic node, as indicated here. The other details of this tree will be discussed later in the paper. The treatment of certain morphologically complex forms such as *being* as syntactically simple is not crucial to our account. The main point is that syntactic hierarchy strictly constrains the morphological exponence, but the operation of spell out may introduce mismatches, so that it cannot be assumed that each morphophonological word corresponds to a single syntactic node (see Lasnik 1995 for a discussion of the current relevance of the affix hopping analysis).

For concreteness, we will assume head movement for word formation, so that we could loosely say that *explained* has moved in (1) to V_{INIT}; more precisely, we would say that V_{RES} has moved to V_{PROC}, which has moved to V_{PASS}, which has moved to V_{INIT}, and the cluster of $[V_{RES}\text{-}V_{PROC}]$ is lexicalized by *explain*, and V_{PASS} is lexicalized by *-ed* (and the exponent of V_{INIT} is phonologically null) (alternatively, exponents 'span' complement sequences of the heads which they spell out, see Svenonius 2012 for discussion).

The ingredients we need for this account are:

(2)   a.   A Cartographic contribution—the ordering of syntactic nodes in the (conceptually grounded) functional sequence, for example giving us the order of T_{PERF} over V_{EVT} (Perf over Prog)

        b.   A selectional contribution—for example the selection of Asp*_{en} by T_{PERF}, rather than some other featural instantiation of Asp*

        c.   A default rule for the spell out of heads in the eventive domain when those heads cannot be filled by raising. This gives us the illusion of *be* 'selecting' for the passive phrase and the progressive phrase.

        d.   A featural stipulation on English modals that they exist only in a morphological form that includes a Fin* feature, like the other tensed morphological forms. This needs to be a stipulation because it is an idiosyncratic fact about English (we give this real semantic content via world anchoring.)

We take up each of these points in detail in the paper, with the motivation of the cartographic contribution being the centerpiece.

2. Ordering and the English auxiliary system

As is well known, the ordering of the English auxiliaries is rigid (cf. Chomsky 1957), as illustrated in (3).

(3)   {T, Mod} $\prec$ Perf $\prec$ Prog $\prec$ Pass $\prec$ V

        a.   He could have been being interviewed.

        b. *John is having returned.

        c. *John is being hunting.

        d. *John seems to have had already eaten.

Various attempts have been made to explain this order. For example, Schachter (1983) suggests that the semantics of the tense and aspect functors is sufficient to impose the order. He suggests that the meaning of the progressive is 'incomplete instantiation of an action or state,' and that the meaning of the perfect is a 'nonspecific relative past.' He suggests that the unacceptability of an example like (3b) is due to an incompatibility between these two meanings: "The notion of an incomplete instantiation that is also a *prior* instantiation simply does not make sense" (ibid, p. 161, emphasis in the original).

But this is at odds with native speaker intuitions about (3b). The example does make sense, in fact it is clear what is should mean; it should mean something like "John is in a state characterized by his having returned." Despite the coherence of the thought, the example is strongly unacceptable. Schachter goes so far as to suggest that context can improve such examples, suggesting that "Supposing, for example, that every time I see you, you have just returned from a vacation, it seems to me that it might be possible for me to say [(4)]".

(4) *Whenever I see you, you're always just having returned from a vacation. (Schachter 1983: 161)

Schachter marks the example with a question mark. However, we mark it as ungrammatical, for the simple reason that it is not entirely acceptable; if a pragmatically plausible and easily interpretable example is nonetheless degraded, then it must be grammar which is responsible for its deviance.[5]

Most modern syntactic representations of the phrase structure of the English verbal extended projection simply assume a templatic ordering of Perf over Prog over Pass (Bjorkman 2011, Sailor 2012, Aelbrecht and Harwood 2013, Bošković 2014), when these elements need to be explicitly represented. Linguists differ with respect to whether they simply represent Perf, Prog and Voice as functional heads (Bjorkman 2011 and Sailor 2012) and handle the inflectional facts via 'affix lowering' or AGREE, or whether they in addition assume separate functional heads hosting *-en* and *-ing* (Bošković 2014 and Harwood to appear b). The minimalist assumption seems to be that some kind of selection is at work, and does not represent a universal functional sequence, and these projections are left out even for English when the literal perfect or progressive forms are not expressed in the sentence. There are reasons to feel dissatisfied with this state of affairs. First of all, the account as it stands barely rises above the level of description, since the labels for the functional projections Prog and Perf are tailor-made for just progressive and perfect respectively, and no attempt is made at a higher level analysis or generality for their function. Secondly, because the current accounts do not aspire to the next level of abstractness, even the question of comparison to other languages and speculations concerning universality or language specific idiosyncrasy are impossible to assess. Thus, the deep questions about what is responsible for this rigid ordering are

---

[5]Note that an absolutive adverbial, controlled by the subject, can contain the perfect under *-ing*:

(i) John was deeply tanned, just having returned from a vacation.

The absolutive is not progressive but is a distinct use of *-ing* (cf. *Having long arms, John can touch the ceiling* versus **John is having long arms*). Schachter's example has a parse in which the *-ing* phrase is an absolutive adverbial predicated directly of the subject using the copula; this might be the parse which is only mildly ungrammatical, rather than a parse which involves the progressive.

never even asked in a meaningful way; they are essentially sidestepped by the stipulation of a deliberately locally descriptive template. As should be clear from the introduction, we consider it an open and empirical question whether the orderings we observe in natural language emerge from (i) universal hierarchical effects, (ii) independent semantic/scopal facts, or (iii) language particular selectional rigidities, and a strategy that starts off by assuming (iii) gives itself no chance to capture crosslinguistic generalizations of the type explored and documented by Cinque and others.

We start off this section by arguing first of all that there is an important syntactic and semantic joint between progressive and perfect in English that should be represented explicitly by an abstract cut-off point in the phrase structure. The existence of this cut-off point will in turn motivate more abstract labels for the functions instantiated by perfect and progressive in English, and we will use it to argue for a broader conception of what drives the hierarchical ordering of perfect over progressive.

2.1. The Progressive vs. Perfect

With respect to a number of different linguistic tests, the progressive can be shown to have a tighter bond with the main verb and its arguments, in several ways which we discuss in this subsection.[6] We discuss expletive constructions, VP-fronting and specificational pseudoclefts, and British nonfinite *do*-substitution, showing how each in turn identifies a domain distinguishing the perfect from the progressive.

2.1.1. Expletive associates

Harwood (2013) (citing and extending Milsark 1974) notes that the thematic subject of a verb in the expletive *There*-construction in English remains low in the clause and is moreover confined to positions either adjacent to the main verb, passive or progressive participles. It can never surface to the left of the perfect participle.

The examples in (5) with the full complement of possible auxiliaries, show that there is only one position in the sequence for an expletive associate, between Perf -*en* and Prog -*ing*.

(5) a. *There could have been being <u>a truck</u> loaded.
     b.   There could have been <u>a truck</u> being loaded.
     c. *There could have <u>a truck</u> been being loaded.
     d. *There could <u>a truck</u> have been being loaded.
     e. *There <u>a truck</u> could have been being loaded.
     f.   <u>A truck</u> could have been being loaded.

Even when the progressive itself is not present, we see that the position to the left of the perfect participle is still unavailable, while the position to the left of the main verb and passive participle is fine, as we see in (6).

---

[6]This is the same conclusion independently arrived at by Harwood (2013), on the basis of some of the same kinds of evidence we show here. His evidence also includes an extended argument based on classical VP ellipsis, and the idea that ellipsis always targets a phasal spell-out domain. We are not actually sure that VP ellipsis is *directly* sensitive to zones the way Harwood suggests, and our view of the zones involved does not allow phasal 'flexibility' the way his account does. Part of the discrepancy between our account and his is that we take the semantic characterization of the zonal distinction as primary and axiomatic.

(6) a. There could have been <u>a truck</u> loaded.
    b. *There could have <u>a truck</u> been loaded.
    c. *There could <u>a truck</u> have been loaded.
    d. <u>A truck</u> could have been loaded.

Similarly, leaving out the perfect and building sentences with just the progressive and the passive as in (7), shows exactly the same restriction: the 'low' subject position can surface to the left of the main verb, passive participle or progressive participle.

(7) a. *There could be being <u>a truck</u> loaded.
    b. There could be <u>a truck</u> being loaded.
    c. *There could <u>a truck</u> be being loaded.
    d. <u>A truck</u> could be being loaded.

This suggests that there is a constituent to the right of the position of the associated DP which includes the verb and Prog if it is present (as well as the passive, if present), but which excludes Perf.

2.1.2. VP fronting and pseudoclefts

There are also differences in the acceptability of VP-fronting and specificational pseudo clefts depending on the nature of the constituent targeted (the observations go back to Akmajian and Wasow (1975), and have been discussed more recently by Aelbrecht and Harwood (2013), Harwood (to appear a), Sailor (2012) to name a few). In (8) we see the constituent headed by *-ing* undergoing fronting, and in (9) we see it forming a grammatical cleft. Crucially, in these two examples, the constituent selected by the perfect auxiliary, and that selected by the modal, cannot be targeted in these constructions.

(8) If Mary says that the cakes will have been being eaten, then ...
    a. *...   [eaten], they will have been being.
    b. ...   [being eaten], they will have been.
    c. *... [been being eaten], they will have.
    d. *...[have been being eaten], they will.

(9) A: John should have been being praised. B: No, ...
    a. *...   [criticized] is what he should have been being.
    b. ...   [being criticized] is what he should have been.
    c. *... [been being criticized] is what he should have.
    d. *...[have been being criticized] is what he should.

When the progressive is not present, we see that in fact the constituent consisting of the passive participle can also be fronted like the progressive participle phrase, but still the perfect participle phrase and the infinitival phrase selected by the modal are not legitimate targets.

(10) If Mary says that the cakes will have been eaten, then ...
    a. ...   [eaten], they will have been.
    b. *... [been eaten], they will have.
    c. *...[have been eaten], they will.

The examples in (11) show that when both the progressive and passive are present in the absence of the perfect, it is still the *-ing* phrase that fronts. The fact that the passive participle phrase does not front on its own seems to indicate that what is being targeted here is the maximal phrase of a certain type.

(11)   If Mary says that the cakes will be being eaten, then ...
    a. *...   [eaten], they will be being.
    b. ...   [being eaten], they will be.
    c. *... [be being eaten], they will.

These facts show that there is a privileged boundary at the point between Perfect *-en* and Progressive *-ing* which is not dependent on the surface presence of any specific aspectual feature or morphological exponent. The facts can be modeled by assuming that the main verb, and passive participle and progressive participle when they exist lie within one distinguished domain which is targeted by these fronting operations. The maximal such domain is what is fronted in 'VP -fronting', and what is clefted in the pseudocleft construction.

2.1.3. British nonfinite do-substitution

Finally, we turn to a novel argument from British nonfinite *do*-substitution, which we argue exposes the same essential division. Some background description of the facts is first in order. In British English, *do* is an abstract pro-form that substitutes not just for eventive verbs but for stative verbs as well, after an auxiliary.

(12)   a.   John might leave, and Mary might do also.
        b.   John might really like oysters, and Mary might do also.

Although British English *do* can replace stative verbs as we have seen, it is confined to main verbs and never substitutes for an actual auxiliary.[7]

(13)   a.   John might have seen the movie, and Mary might (*do) also.
        b.   John might be singing a song, and Mary might (*do) also.

However, even within these constraints, not all nonfinite forms may be substituted for by *do*:

(14)   a.   John might leave, and Mary might do also.
        b.   John has left, and Mary has done also.
        c.   John is leaving, and Mary is (*doing) also.
        d.   John was arrested, and Mary was (*done) also.

British nonfinite *do* can substitute for an infinitive modal complement or a perfect participle, but not for a progressive or passive participle; hence it, too, motivates a cut between Perf and Prog. This diagnostic is in some sense the converse of the previous one: the very constituents that could participate in the fronting constructions are the ones that British nonfinite *do* cannot substitute for.

---

[7]Note that the mismatched reading in (13a) where *do* is construed as substituting for a main verb in non-finite form after the modal auxiliary, is marginally possible here, but is irrelevant and will be ignored in what follows. The reading where it substitutes for the auxiliary phrase is robustly ungrammatical.

2.1.4. Two Domains

We thus have robust evidence for two distinct domains from three independent sets of grammatical facts. In each case, the facts point to a joint between the progressive participial phrase and the perfect participial phrase when they exist (and we have seen that the joint exists even when the morphological evidence is not fully articulated).

Let us recap: with respect to independence (mobility), and a thematic position for the external argument, we found that progressive, passive and main verb formed a unit to the exclusion of modals and the perfect; with respect to substitution by the pseudo-auxiliary verb *do* in British English, complements of passive and progressive auxiliaries patterned together in being not replaceable by *do*, while the complements of the perfect and modals could be so substituted.

Thus, with respect to a crude macro division of the clause into a VP-domain and a TP-domain, we find evidence that the progressive and passive forms lie within the former unit, while modals and the perfect lie within the latter. British English nonfinite *do*-substitution is a pro-form for the higher, but crucially not the lower domain.[8] If we follow standard assumptions about passive being located in VoiceP, the most conservative representation for what we find in the data can be illustrated by the tree in (15). Note that the generalization requires reference to the constituents lexicalized by *-ing* and *-en*, and not those lexicalized by the auxiliaries themselves, so we have labelled these as such, purely descriptively at a first pass.

(15)
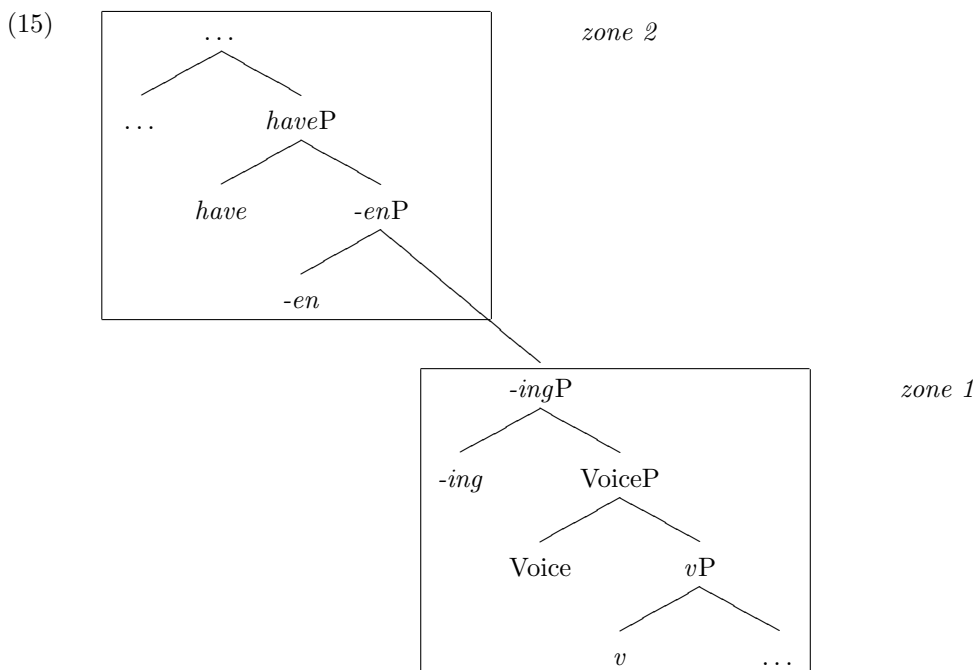
The pressing questions at this point concern the defining property of these two zones

---

[8]This makes the difference between the British English dialects and the more restrictive ones, such as the American, quite simple to state: standard dummy *do* support in the more restrictive dialects has only finite instantiations, British English possesses a non-finite version of this *pro*-form as well.

and why they should be hierarchically ordered in this way. The existence of two zones in this part of the clause is not in itself a new or surprising result. It is implicitly part of every phrase structural representation that assumes a VP-TP-CP partition at a coarse level of granularity, and of work on locality and phases. The novelty here comes from actually attempting to line up auxiliary ordering with these zones in an explicit way, placing the progressive participle within the VP domain and the perfect participle outside of it. Our argumentation has not involved the classic kinds of evidence usually adduced for phase boundaries with regard to escape hatches for movement (e.g. Legate 2003), for the simple reason that the argument can be made independently of such evidence, on the basis of uncontroversial data from English. However, the evidence we have used points to two important zones, and moreover to the conclusion that these do underpin the facts about cyclicity and locality of operations that syntacticians have noticed independently.[9]

There are further suggestive facts that point to the progressive being inside the VP zone of the clause. Under the assumption that selectional restrictions are strictly local (Baltin 1989), the fact that the progressive places selectional restrictions on the Aktionsart of the verb phrase it combines with is further evidence that Prog is low enough in the extended projection to select for the nature of the event structure described by the verb. As is well known (see e.g. Dowty 1979), the progressive in English combines with dynamic verbal projections and not stative ones (16).

(16)   a.   John is dancing the tango.
        b.   *John is knowing the answer.

In contrast, the Perfect does not constrain the Aktionsart of its complement. In (17), we see that the perfect has no restrictions at all: it can combine with *any* main verb in the English language. It is true that the meaning of the perfect changes subtly depending on the type of main verb, but the perfect itself does not choose what it can combine with. For example, the universal reading of the perfect as in (17c) below is facilitated by a time-frame adverbial which then forces an imperfective interpretation, and the result reading in (17a) can only arise with telics.

(17)   a.   John has destroyed the castle. (result)
        b.   John has eaten sushi. (existential)
        c.   John has known Sue for three years. (universal, stative)

The different readings are plausibly due to compositional semantics, whereas the perfect *-en* itself can blindly combine with any VP it likes. This is in contrast to the progressive which selects, and either rejects or coerces. The progressive thus must be local to the event building domain, while the perfect must be outside it.

The opposite kind of effect can be seen from the temporal domain. We assume that T is the locus for anchoring a temporal interval associated with the event to the utterance time by establishing a precedence or overlap relationship to it (see Klein 1994). In the perfect tense, the relationship established to the utterance time is not necessarily congruent with the event's notional run time. In (18a), we see that the perfect auxiliary is identified with the utterance time by morphological present tense but the actual event's

---

[9]As mentioned earlier, Harwood (2013) argues that VP ellipsis targets (flexible) phases, and argues that the ellipsis evidence shows that the progressive is inside the first phase.

run time (the writing of the letter) must be prior to that point. In (18b), the perfect auxiliary bears morphological past tense, but again the actual event of car-washing must precede that moment, and does not even need to overlap with it at all.

(18) a. John has (now) written the letter.
      b. When I met up with him, John had just washed his car the day before.

The same is true of root modals. While root modality expresses a permission, ability, obligation on an individual rooted at a particular time, the run time of the event denoted by the VP that the modal combines with is always projected *forward* from that permission/ability/obligation (see Condoravdi 2002 for discussion).

(19) (Now) John may go to the party tomorrow.

So in each of these cases, the topic situation (to put it in the terms of Klein 1994) is not identical to the event (or its run-time) per se, but bears a more indirect relationship to it. Both the modal and the perfect auxiliary introduce topic situations which are anchored to the utterance time, and which are distinguishable in principle from the VP event (although of course related to it in a particular way).

In the progressive, on the other hand, the event run time and the tense specification of the progressive auxiliary cannot be so distinguished. The progressive picks out a mereological subpart of the event as described by the VP, a relation that is possible without the presence of explicit temporal information. We propose that the more indirect relationships between the topic situation and the event that are observed with modals and the perfect are possible because they inhabit a zone where situations can be explicitly related to each other via temporal information, and not just via mereology. Thus, while semantic analyses of both the progressive and the perfect can be found in terms of the building of 'derived eventualities' (cf. Parsons 1990), only the perfect derived eventuality can be temporally disjoint from the VP event, and is truly distinguishable from it. We take this as further suggestive evidence that the perfect actually occupies a higher domain, one where temporal information is relevant, but which bears a more indirect relationship with the event. We will propose that the progressive is actually inside the core event building domain of the clause, *before* the introduction of temporal information. Thus, it is impossible to link the progressive situation to the utterance time without also entailing the linking of the event to that same utterance time.[10]

This examination of the perfect versus the progressive in English has exposed a pervasive zonal difference between the two forms that shows up in selection, interpretation, movability, substitutability and relationship to the external argument. Simply enforcing an ordering on them by a local selectional mechanism fails to predict these systematic correlations. Instead, we propose that the ordering of perfect over progressive is due to the fact that they occur on either side of an important joint in the phrase structural

---

[10]Notice that even in cases where the English present tense on the progressive auxiliary gets a 'planned future' interpretation, as in

(i)   We are playing football tomorrow.

The in-progress situation and the whole 'play football' event are *both* projected into the future. This is unlike the situation with the root modals.

make-up of the clause—broadly speaking, between the event related VP domain and the time related TP domain.

2.2. The Perfect versus The Modals

Turning now to the ordering we find in English between the perfect and root modals, we may ask whether this ordering too exposes an important juncture in the phrase structure of the clause. One problem that arises immediately is that in English modals have no nonfinite forms, limiting the range of diagnostics which can be applied to them.[11] One indication that English is misleading in this regard is that we can find closely related languages which do have nonfinite modals. We illustrate with Norwegian here, where modals can in fact be placed under the Perfect. First, in (20), we see the English-like order Mod ≺ Perf.

(20) a. Kari kan ha   gått på ski.
         Kari can have gone on ski
         'Kari might have gone skiing'
      b. Ola må   ha   måket.
         Ola must have shoveled
         'Ola must have shoveled snow'

In (21), Norwegian exhibits the option of placing a nonfinite modal under the perfect, ruled out in English due to the absence of participial modals.

(21) a. Kari har kunnet   gå på ski til jobb hver dag.
         *Kari has could.*PTCPL *go on ski to work every day*
         'Kari has been able to ski to work every day'
      b. Ola har måttet   måke sne i hele dag.
         *Ola has must.*PTCPL *shovel snow in whole day*
         'Ola has had to shovel snow all day'

Norwegian also allows one modal to be embedded under another, again impossible in English because the modals lack infinitive forms.

(22) a. Kari må   kunne   gå på ski.
         *Kari must could.*INF *go on ski*
         'Kari must be able to ski'
      b. Ola kan måtte   måke.
         *Ola can must.*INF *shovel*
         'Ola might have to shovel'

---

[11]For the moment, we set aside the interesting example of *have to* here, which raises additional issues. On the face of it, *have to* appears to support what we establish here using Norwegian, but the exact structure is unclear, along with the status of *to*.

(i)   a. Ollie has   to   have   cleared   the driveway.
              MODAL INF.T PERF.AUX PARTICIPIAL V
      b. Ollie has   had   to   clear   the driveway.
              PERF.AUX MODAL INF.T INFINITIVE V
      c. Ollie might have to   clear   the driveway.
              MOD   MOD INF.T INFINITIVE V

13

The data from Norwegian seems to show that the rigid ordering in English is a morphological selectional fact, and does not reflect anything deep about the semantics.[12] We have already seen that with respect to temporal marking, modals and perfects pattern together in establishing a more indirect relation to the underlying event. And we know that neither of them seems to select for aktionsart (like the progressive) or argument structure type (like the passive). So far the evidence points to modal and perfect being in the same zone, but not for a deep ordering between them.

At this point we need to bring tense and epistemic modality into the discussion, because a closer look at the Norwegian data reveals that it is only circumstantial modality that freely orders with respect to the perfect.

Looking at the Norwegian examples in (23), we see that only the root modal meaning is possible under the perfect auxiliary, not the epistemic possibility or epistemic necessity interpretations.

(23)   a.   Kari har kunnet   gå på ski til jobb hver dag.
               *Kari has could.*PTCPL *go on ski to work every day*
               'Kari has been able to ski to work every day' (root only; *'has possibly skied')
    b.   Ola har måttet   måke sne i hele dag.
               *Ola has must.*PTCPL *shovel snow in whole day*
               'Ola has had to shovel snow all day' (root only; *'has apparently shoveled')

Norwegian examples of Mod ≺ Mod also show this: an epistemic modal can dominate a root modal, and one root modal can dominate another, but a root modal can never dominate an epistemic modal, as the examples in (24) show.[13,14]

(24)   a.   Kari må   kunne   gå på ski.
               *Kari must could.*INF *go on ski*
               'Kari must be able to ski'
    b.   Ola kan måtte   måke.
               *Ola can must.*INF *shovel*
               'Ola might have to shovel'

The Cinquean hierarchy encodes this fact explicitly, where Epistemic > T > Circumstantial (> ...). It however also encodes circumstantial modality over perfect.[15]

---

[12]McCawley (1971) discussed this in the context of explaining English auxiliary ordering. He suggested that the reason that modals are at the top of the hierarchy in English is that they only have finite forms; but he failed to explain why finite forms have to be at the top of the hierarchy, i.e. why is T, or Fin in our terms, at the top rather than somewhere in the middle? We address this in §3.

[13]Similarly with English *have to*; it has an Epistemic reading only when unembedded

(i)   a.   John has to be in the library. (Epistemic possible)
    b.   John might have to be in the library. (No epistemic reading possible)

[14]See Eide (2005) for more detailed discussion of Norwegian modals.

[15]Cinque's position on the relative height of deontic modality and the perfect has changed over time. In Cinque (1999: 130), $Mod_{obligation}$ and $Mod_{ability/permission}$ are placed above $Asp_{perfect}$, but in Cinque (2004: 133), $Mod_{obligation}$ and $Mod_{ability/permission}$ are lower than $Asp_{perfect}$. On our account, both orders are in principle possible, though independent factors may rule out the one or the other in individual contexts, for example in English the fact that modals are always finite will force the first order.

We tentatively conclude from Norwegian that epistemic modals are indeed universally ordered over root modality, but that the ordering of root and perfect are in fact *not* rigid. Our view is largely compatible with that of Hacquard (2006), who proposes that epistemic modals are bound by the higher speech event, while circumstantial modals are bound by a lower event at the level of aspect. For her then, there is no ordering statement per se, rather, she offers an explanation for why different sorts of meanings emerge when the very same modal is merged high or low in the clause (cf. also Butler 2006). In the next section, we will propose an analysis that is, like hers, grounded in the compositional semantics, but it will differ in that it will be framed in the context of a basic distinction between events and situations.

3. Sortal domains as conceptual underpinnings of hierarchy

We have seen that the ordering of auxiliaries in English stems from a number of different sources. In one case (PERF over PROG), we argued that the order arose because of a major phrase structural zoning relating to the semantics of tense and events; we suggest that the way languages express tense and event structure is universally constrained in certain ways which we will elucidate in this section. In the other cases, we have proposed that the facts that order modal over perfect are morphologically specific to English. Finally, we have so far left open the discussion regarding epistemic modality and its relation to tense and the other modals.

Given what we have seen so far, the difference between the VP domain and the TP domain is something that we want to reify in the phrase structure, and it is only at this crude level of granularity that we need to make the cut between perfect and progressive. However, we wish to emphasize that even in its most pared down form, the ordering of CP> TP > VP still represents a templatic residue of cartography in so far as it needs to be stipulated as opposed to derived. Having convinced ourselves of the reality of VP versus TP zonal distinction in the auxiliary system of English, we wish to take a step back now and ask what this juncture could possibly signify, which we consider to be an important first step in understanding how such zoning could possibly emerge from independent properties of the mind/brain.

Clearly, to do so requires leaving the comfort zone for most syntacticians. Nonetheless, we feel that we need a concrete hypothesis about the origins of the functional sequence in order to push forward the agenda that we outlined at the outset of this paper, reconciling Cartographic findings with the Minimalist goal of moving beyond explanatory adequacy.

Hinzen (2006) argues persuasively that syntactic phenomena cannot be explained by appeal to 'deep' or underlying semantic structures, in the fashion of Generative Semantics (including more recent models such as that of Jackendoff 2002). Instead, Hinzen proposes, complex semantic meaning must be built by syntax.

However, there are underlying cognitive biases which can be assumed to be independent of syntax, such as certain biases regarding the separation of percepts into domains (e.g. events, which have participants, versus objects, which can be participants in events) or the organization of perceived states of affairs in terms of Figure and Ground. Such aspects of the cognitive backdrop for language, we suggest, lead to certain structures being universally realized.

We propose that the purely formal phrase structural zones we have evidence for in language correspond to certain arrangements of basic ontological semantic notions that

are universal to our species (at least). In particular, we will develop a model of events versus situations which characterizes to the best of our ability the distinction between the V-domain and the T-domain.

Events. There are compelling reasons to recognize EVENTS in semantic representations, for example they can be quantified over (*John knocked twice*). The tradition of recognizing events in linguistic semantics goes back to Davidson (1967). However, note that we depart substantially from Davidson in our particular conception of events, which for us are atemporal (see Gawron 2006 on events mapped onto scales other than time).

Basic characteristics of events, or eventualities (including states), include the following:

(25) Characteristics of events (or eventualities)
   a. People have consistent intuitions about what percepts constitute a single event; an instance of a potentially distinct event-type may be a SUBEVENT in a larger event (Wolff 2006)[16]
   b. Causation and resultativity are relations among subevents; possibly they are both specific instances of a more general 'leads to' relation (Ramchand 2008)
   c. Thematic roles are relations between individuals and events (cf. Higginbotham 1985, Parsons 1990)
   d. Stativity and dynamicity are possible properties of events or subevents (cf. Bach 1986, Jackendoff 1990)

We suggest that the V-domain as commonly recognized in syntax is that part of the syntactic structure which denotes an event description. This can be represented as follows, if the semantic content of *verb* is VERB.

(26) $[\![_{\text{VP}} verb]\!] = \lambda e.\text{VERB}(e)$

A standard formal semantics identifies type $\langle e \rangle$ 'entity' with the semantic type of arguments. We assume that such arguments can be of distinct SORTS, including a sort for individuals and a sort for events. We thus reserve e for the sort of events, though this is also type $\langle e \rangle$ in the traditional, broader sense.

Situations. Situation semantics (Barwise and Perry 1983, Kratzer 1989, inter alios) originally emerged as an alternative to possible world semantics. Situations are partial specifications of states of affairs. We distinguish them from events, but also from possible worlds and from propositions, which we will discuss shortly.

Descriptively, what we see in linguistic semantics are situations with the following properties.

(27) Characteristics of situations

---

[16]Wolfram Hinzen (personal communication) informs us of unpublished evidence by Jill de Villiers suggesting that these intuitions are impaired when the language faculty is engaged, suggesting that humans need access to the language faculty in order to discriminate events. We have not had the opportunity to review this work, but will simply point out here that on our view, syntax is used to build complex events, for example consisting of V_{INIT} and V_{PROC} (actor-directed activities), even if the possibility of construing their combination as constituting a single 'sort' of thing is governed by underlying cognitive bias.

a. Situations are elaborations of eventualities (Kratzer 2008) (hence they presuppose the existence of an eventuality, so the eventuality is closed — either existentially closed or bound by some other kind of operator)
b. Situations have a TIME parameter, unlike events (Giorgi and Pianesi 1997)
c. Situations have a WORLD parameter, unlike events (Lewis 1986, Austinian topic situations)
d. Situations can have topics (the case where the Austinian topic situation is based on an individual, or a description of an individual)

The central claim for our purposes (that is, for deriving a portion of the functional hierarchy) is that situations contain events (or eventualities), in a way that is different from events containing situations. In other words, an event is a *constitutive part* of a situation, while a situation cannot be a constitutive part of an event (though an event can embed a situation, in syntactic recursion).

Formally, we can say that a situation is an event plus something else. Wiltschko (to appear) proposes a notion of *anchoring* for situations, which according to her can be parametrized by language over anchoring to times, anchoring to locations, or anchoring to discourse participants (see also Ritter and Wiltschko 2009). Adapting that proposal slightly, we could say that a situation is an 'anchorable' entity, since it contains parameters (such as time and world) that can be directly related to utterance parameters.

Ultimately, we stipulate that events have a central place in the constitution of a situation. Other parts of the situation, which have a different status, include times, worlds, and grammatical functions (subject, object, etc., possibly as an extension of the notion of topic mentioned in (27d), though we will not be able to discuss grammatical functions further in this paper.)

Alongside the sort e for events, then, we posit a sort s for situations. We suggest that the traditional syntactician's T-domain is exactly that part of the clause which denotes a situation description. If the semantic content of some T-domain functor *aux* is AUX, then the semantic interpretation of a TP can be represented as follows:

(28) $[[_{TP} aux \dots]] = \lambda s.\text{AUX}(s) \wedge \dots$

It is important to the workings of the system that events are not visible to operators in the TP domain. We ensure this by existentially closing the event when it is embedded in a situation. There are two aspects to this sortal shift that need to be built in. One is the fact that the sort of variable in the lower domain is not accessible to direct modification in the higher domain. However, there is an important constraint on the way linguistic symbolic descriptions are constructed which enforces monotonicity of informational complexity. In other words, the lower sortal information is not lost, but embedded or clothed within the description of the higher sort. The higher sort is always an elaboration, or precisification of the information given by the lower sort. This means in particular that situations are more complex than events. They contain *in addition* to their event parameter, place holders for temporal and world parameters.[17] In fact, we

---

[17]Here we leave aside the possibility of locational parameters even though they are often assumed in the philosophical literature (cf. Lewis 1986), since they do not crucially play into our ontology for English. See Ritter and Wiltschko (2009), Wiltschko (to appear) for arguments that locations are important to situation anchoring in some languages.

believe this is part of a general principle of semantic compositionality and posit it as a principle of Compositional Coherence, as stated here.

(29) **Compositional Coherence**: If X embeds YP, then the denotation of XP is a monotonically coherent elaboration of the denotation of YP.

This notion is at the heart of our architecture, and deserves closer formal treatment. The intuition here is that there is a cognitively real primitive of informational elaboration that constrains sortal shifts here, underpinning the embedding of event under situation and not vice versa. One issue is that it is still unclear to us whether the device of 'existential closure' for lower eventuality variables is the appropriate notation for these kinds of elaborations, but we must leave that question aside for now. In addition, informational monotonicity needs to be stated carefully enough so that it still allows for things like negation. A discussion of genuine recursion is also missing where new events *can* be derived from situational descriptions, once the functional sequence contains recursive embedding, or is 'restarted' in some sense (this principle would argue that it *must* restart in such cases). We leave discussion of these issues for further work.

The shift between events and situations is a sortal one in our view, but the transition we describe can actually be related to a long semantic tradition with tense-aspect semantics, starting with Reichenbach (1947). In the aspectual literature, it is clear that a distinction needs to be made between relations established between an event (E) and an abstract reference time (R) on the one hand, and R and the speech time (S) on the other (Klein 1994, Giorgi and Pianesi 1997, Demirdache and Uribe-Etxebarria 2000). The diagram below comes from Giorgi and Pianesi (1997).

(30)

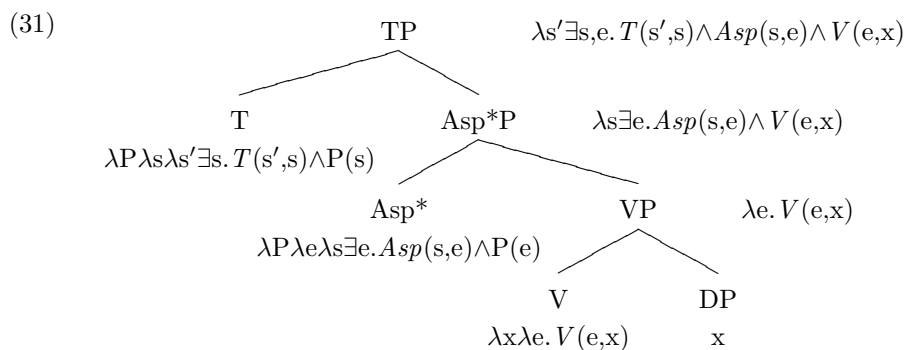| RELATION 1 (tense):		RELATION 2 (aspect):	
$S \prec R$	future	$R \prec E$	prospective
$S \succ R$	past	$R \succ E$	perfect
$S = R$	present	$R = E$	neutral

In those works it is assumed that both relation 1 and relation 2 are relations between intervals, and no sortal distinction is made between E and R. The relation between E and R is most often labelled Asp(ect), where the relation between R and S is Tense. In the sortal view we are exploring in this paper, we essentially saying that Relation 2 in the table above is more semantically significant than usually acknowledged: this is the point of sortal transition where the mereological event domain is booted up to the more complex sort corresponding to situations and where an explicit temporal parameter therefore becomes associated with the event. To make clear the parallels to this tradition, we will label the functional head in our phrase structure that effects the transition Asp*.[18]

---

[18]We use the asterisk to distinguish this Asp from other kinds of auxiliaries or inflections or markings that might pretheoretically be called aspectual in the literature. By Asp* we mean specifically the functional head that combines with an event description without a temporal parameter to deliver a situational description with temporal parameters. Within this abstract general definition, there can of course be different ways of doing this furnished by different morphemes, depending on *how* the temporal reference situation is related to the underlying event. The different morphemes that sit in Asp* probably correlate with many instances of inflection that have been called imperfective or perfective aspect in the literature. In this respect, we are also in essential agreement with Klein (1994) and the subsequent work in this tradition.
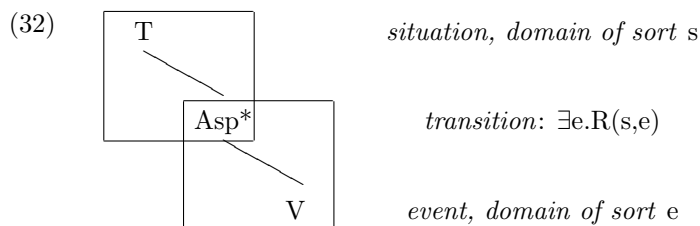
We think this is a satisfying reinterpretation of the Reichenbachian view for a number of reasons. Firstly, there is no real logical reason why tense forms in language should require a two step process of temporal relations to relate an event to the speech time. If an event has a time, and the speech time is the deictic anchor, why doesn't language just relate the event directly to the speech time? Why does it seem to go through this intermediate 'placeholder' which Reichenbach called the reference time? Under the sortal view, the two step process becomes required: events do *not* inherently come with intervals so they need to be converted to the situational sort first, derivationally speaking (by embedding under Asp*), and then related to the speech time (by T).

Asp* is formally relational: It relates its complement, the event description, to the situation of which that event is a constitutive part. We could represent the situation as an argument in the specifier of Asp*, along the lines proposed by Wiltschko (to appear) (see also Percus 2000), but since that will play no further role in the specifics of our proposal, we do not explicitly represent it in our tree diagrams.

Thus, to reiterate, we assume that the locus of Relation 2 in the above table is an aspectual head, Asp*, while the locus of Relation 1 is the tense head, T (cf. Klein 1994, Demirdache and Uribe-Etxebarria 2000). We furthermore assume that at the transition point Asp*, the event sort is embedded in a situation (formally, it is related to a situation and existentially closed). This is represented in the following tree.

(31)

TP $\quad\quad \lambda s'\exists s,e.\, T(s',s)\wedge Asp(s,e)\wedge V(e,x)$

T $\quad\quad$ Asp*P $\quad\quad \lambda s\exists e.\, Asp(s,e)\wedge V(e,x)$
$\lambda P\lambda s\lambda s'\exists s.\, T(s',s)\wedge P(s)$

Asp* $\quad\quad$ VP $\quad\quad \lambda e.\, V(e,x)$
$\lambda P\lambda e\lambda s\exists e.\, Asp(s,e)\wedge P(e)$

V $\quad\quad$ DP
$\lambda x\lambda e.\, V(e,x)$ $\quad\quad$ x

Below is a compressed representation of the same analysis, where the boxes represent the accessibility of the e and s arguments:

(32)

T $\quad\quad\quad\quad\quad$ *situation, domain of sort* s

Asp* $\quad\quad\quad\quad\quad$ *transition*: $\exists e.R(s,e)$

V $\quad\quad\quad\quad\quad$ *event, domain of sort* e

So for example, if a sentential adverbial (S-Adv) like *always* or *already* is a property of situations, then that S-Adv can merge in the T domain, but cannot merge in the V domain, where it will have no interpretation.

And if a verb-phrase adverbial (V-Adv) like *completely* or *well* is a property of events, then that V-Adv will be interpretable in the V domain, but cannot be attached outside
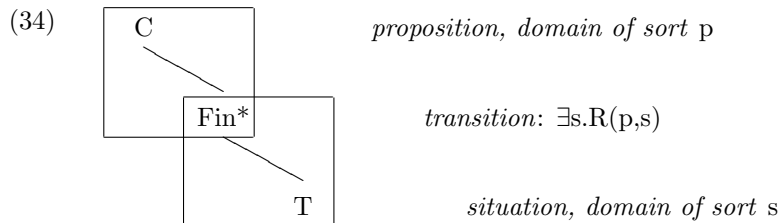
the existential closure of e at Asp*P.

Although we have not so far spent much time discussing the CP domain, we propose that it too corresponds to a primitive semantic sort in our ontology. We call this sort the proposition, though the term has been used for many different purposes by semanticists and philosophers (see in particular Hinzen 2006 for a critique of the Russelian notion of a mind-independent proposition). Despite the historical burden the term carries, we find it best evokes the notion we want to express. We give it our own narrow interpretation here.

(33)   Characteristics of propositions.
   a.   Propositions are elaborations of situations; thus they presuppose a situation, which is existentially closed
   b.   Propositions, unlike situations, are anchored to the utterance context, having 'Force' in the discourse (Bianchi 2003, Ritter and Wiltschko 2009, Wiltschko to appear)
   c.   It is only at the level of the proposition that speaker-oriented parameters come into play (Giorgi 2010).

Propositions are built from situation descriptions, just as situation descriptions are built from event descriptions. Compare Wiltschko's notion of 'discourse linking'; adapting that model, we could say that a 'discourse-linked' situation is a proposition.

Thus, analogous to the box-diagram in (32), we have a representation showing the C-domain over the T-domain. A head relates the situation description to the proposition; we propose that this is a central function of the head Fin[iteness], which we label Fin* in order to highlight the parallel with Asp* (and just as Asp* could be modeled as taking a situational variable as its outer argument, Fin* could take a propositional variable as an argument in its specifier).

(34)

proposition, domain of sort p

transition: $\exists s.R(p,s)$

situation, domain of sort s

Putting the two boxes together, we have a clause consisting of three domains:

(35)

C	proposition, domain of sort p
Fin*	transition: $\exists s.R(p,s)$
T	situation, domain of sort s
Asp*	transition: $\exists e.R(s,e)$
V	event, domain of sort e
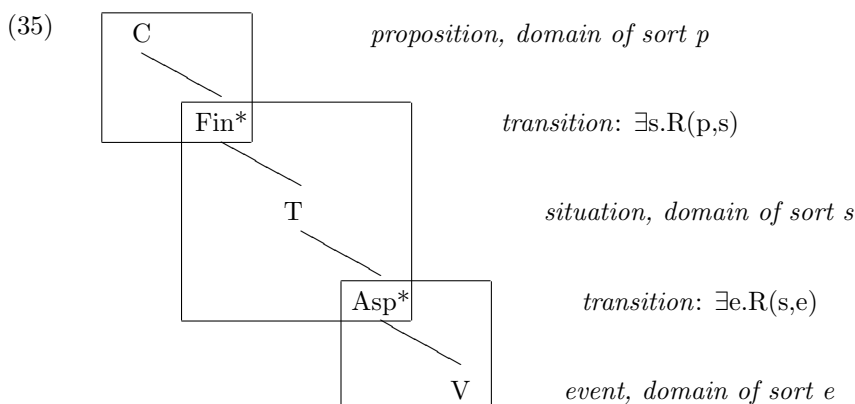
Just as functors and operators in the T-domain could not interact scopally with event-level material, operators in the C-domain cannot modify the content of the situation description, because it is closed at the level of Fin*.

Wiltschko (to appear), on the basis of considerations similar to ours, posits a zone for discourse linking, analogous to our proposition (C), and a zone for 'anchoring,' which corresponds to our situation (T). Her zone for 'classification' corresponds to our event (V), but she also posits a distinct zone for 'point of view' between the event and the situation, corresponding to the category of Aspect. She locates the perfect there, as well as the perfective and other aspectual distinctions. We think that our model gives us the right degree of granularity, with aspectual distinctions being variously made in the T-domain, in Asp*, and in the V-domain, without positing a fourth semantic sort, but further investigation may prove otherwise.

In this section, we have outlined a theory of the core semantics of propositions. We suggest that a primitive distinction be made among three basic sorts: the sort of timeless eventuality descriptions, the locus of force dynamics and thematic relations among arguments; the sort of time-anchored situation descriptions, where basic modal ordering sources become available; and the sort of propositions, which are anchored to a discourse context, making it possible to take a speaker's perspective into consideration. We hold that propositions are built from situation descriptions, which are built from eventuality descriptions, and that this fact lies at the root of the syntactic structure of the clause, in which C dominates T, which dominates V.[19]
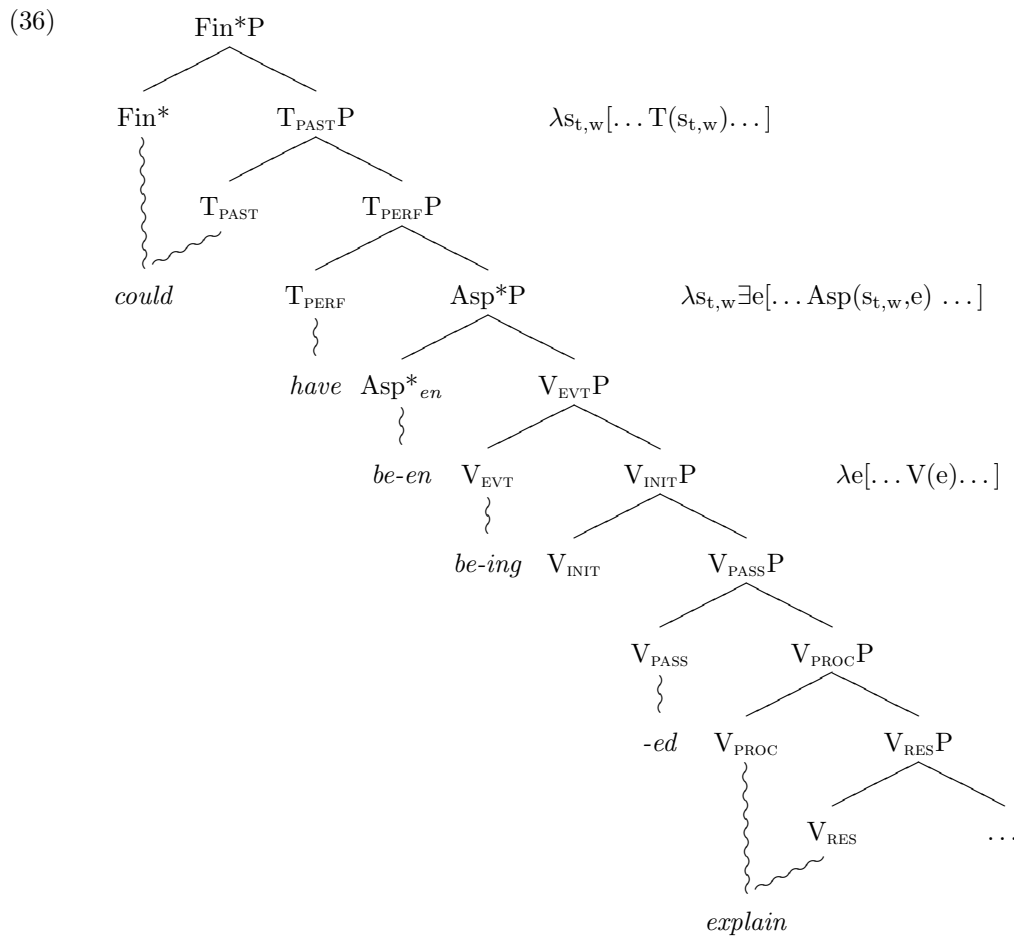
4. English Auxiliaries, once more

In this section we return to the concrete example of English auxiliary order, showing what aspects of that order can be explained by the model outlined in the previous section.

Empirically, we motivated a cut-off point between the constituent headed by the -ing participle, and that headed by the -en participle of the perfect. Conceptually, we

---

[19]A reviewer asks about the often proposed parallelism between the clause and the noun phrase (explored, for example, in Svenonius 2004 and in Wiltschko to appear). Our proposal would lead us to expect zones in the extended noun phrase and the extended projections of other categories much as in the clause, but perhaps based on different primitive concepts, for example regions and paths in the PP domain, as suggested in work in progress.

motivated a cut-off between the event sort and the situational sort, mediated by the transition point which effected that sortal shift and which we labelled Asp*. In lining up the two results, we still find a degree of indeterminacy: we could claim that the head lexicalized by the *-en* participial morphology must be *above* Asp*, but we would also get the facts if *-en* were a possible lexicalization of Asp*, perhaps a featural variant of Asp* which could be represented Asp*$_{en}$. We go for the latter position here, for concreteness. The *-en* participle also spells out part of the passive structure. This suggests that the perfect and the passive share some syntactic component (an issue to which we return in §4.2).

A tree with the maximum number of modals would therefore have a structure roughly as follows (adapted from (1)).

(36)

Fin*P
├─ Fin* — *could*
└─ T$_{PAST}$P $\lambda s_{t,w}[\dots T(s_{t,w})\dots]$
   ├─ T$_{PAST}$
   └─ T$_{PERF}$P
      ├─ T$_{PERF}$ — *have*
      └─ Asp*P $\lambda s_{t,w}\exists e[\dots Asp(s_{t,w},e)\dots]$
         ├─ Asp*$_{en}$ — *be-en*
         └─ V$_{EVT}$P
            ├─ V$_{EVT}$ — *be-ing*
            └─ V$_{INIT}$P $\lambda e[\dots V(e)\dots]$
               ├─ V$_{INIT}$
               └─ V$_{PASS}$P
                  ├─ V$_{PASS}$ — *-ed*
                  └─ V$_{PROC}$P
                     ├─ V$_{PROC}$
                     └─ V$_{RES}$P
                        ├─ V$_{RES}$ — *explain*
                        └─ …

In the interest of maximum explicitness, we assign labels and sorts to all the elements in the auxiliary system and give denotations for them. We do this because part of our claim is that in certain cases the sortal denotations are inseparable from the hierarchical order of syntactic categories given. This is what we mean when we say that certain universal aspects of the functional sequence are grounded in conceptual sorts and the

necessary relations among them. At the same time, although we have tried to be faithful to the results of much recent semantic work in the areas of progressive, perfect and modality, justifying these denotations in all their specificity goes way beyond the scope of this short article. We hope that we can be explicit enough to demonstrate the workability of the general approach, even though the details remain to be nailed down and would require a much longer exposition.

An important aspect of the type of analysis we propose here is that we do give explicit denotations for functors like Asp*$_{en}$ and V$_{\text{EVT}}$ ("-ing"), and attempt to give a compositional account for the formatives instead of assuming that they arise syncategorematically as a result of some selectional or morphological matching. The evidence given in the first part of this paper shows that there is something substantive about the meaning of the -ing participle that is sortally specific and needs to be represented explicitly as a head in the structure.

Note that while the perfect auxiliary T$_{\text{PERF}}$ (have) is given a specific denotation and position in the structure, the progressive and passive auxiliaries are inserted under the same node as the inflections they bear. Here we follow the common assumption that be can be inserted to carry inflection (essentially following Warner 1986, Lasnik 1995, and Bjorkman 2011). In the absence of an 'active' overt main verb with uninterpretable features at the end of the first phase, be must be inserted. This allows us to say that the passive auxiliary and the progressive auxiliary are the same (dummy) element.

4.1. The event zone, sort e

The assumption of this model is that all the heads in this zone have denotations that make them descriptions of the event sort. We have left specifiers out of the diagram in order to focus on the semantic denotation of the extended projection of V, the 'spine' of the clause. Along that spine, everything below Asp* is an eventuality description, semantically $\lambda e.P(e)$, a set of events. Events are built up from eventuality descriptions (interleaved with argument positions). In the lowest part of the VP, subevents combine to form complex events via a generalized causative relation (cf. 'implicates' in Hale and Keyser 1993). Also in the event building zone are formatives like passive (V$_{\text{PASS}}$) which manipulate the argument structure of the causally complete event description.

We give a denotation for V$_{\text{PASS}}$, spelled out as the participial morphology, which does this by creating a derived stative event description (e$'$) from the event description (e) it combines with, where e$'$ is the eventuality that emerges as a direct result of e (Ramchand and Svenonius 2004). The progressive head is also in this zone. It creates a derived stative eventuality description which is a mereological subpart of the (dynamic) event that it combines with. Events can be made complex via causativization and applicativization in this zone, but also simplified by taking mereological subparts. As is well known, the range of possibilities is tightly constrained by apparently syntax-external factors, such that something like a thematic hierarchy is respected. The important issue for us here is that heads within this zone build events of increasing complexity, but remain within the event sort.[20]

---

[20]Note that the information from the core event description is never entirely lost, as is well known: the interpretation of the progressive derived event is dependent on the interpretation of the whole non-progressivized event in ways that are non-extensional, and the passive derived event still contains an existential entailment over the unexpressed external argument.

For perspicuity, we label all the heads within the event sortal zone V, with some appropriate subscript.

The head we label V_{PROC} (cf. Ramchand's 2008 *process*) is a description of a process, a dynamic event with a participant argument that UNDERGOES the process. This projection is a component of all dynamic verbs (following the theory of Ramchand 2008). The interpretation of the argument (not shown in the above tree) is provided by a neo-Davidsonian thematic relation as follows (cf. Parsons 1990):

(37) V_{PROC}: $\lambda x \lambda e[\text{Process}(e) \ \& \ \text{Undergoer}(e,x)]$

The node labeled V_{INIT} implies an external argument (thus, if it is absent, the verb is unaccusative). When it combines with a process eventuality description, as depicted here, the combination is interpreted as a causal one, so that the whole macro-event is a caused-process eventuality description. We assume that the causal interpretation is the most basic one, cognitively, for a macro-event composed of two subevents, and that the interpretation of the external argument as an 'initiator' (causer or agent) follows from that.

(38) V_{INIT}: $\lambda P \lambda y \lambda e'[\text{CausedProcess}(e', e) \ \& \ P(e) \ \& \ \text{Initiator}(e', y)]$

The node labeled V_{PASS} is a 'passive' head, indicating the site of merger of the passive participial morphology. The key fact about this participle is that it is sensitive to the argument structure of the verb that it combines with, and crucially, that it derives a predicational form that is 'true' of the internal argument.[21]

(39)   a. The fallen leaves.
        b. The broken stick.
        c. The destroyed castle.
        d. *The danced man.
        e. *The written author.

In building a verbal passive, the participle combines with a dummy verb *be* and apart from suppressing any direct DP expression of the external argument, essentially retains the aktionsartal (e.g. state vs. event) classification of the base verb. We thus define V_{PASS} as being a head which selects for V_{PROC} (or V_{RES}), creating a predication over the single internal argument.

(40) V_{PASS}: $\lambda P[\exists x (x \text{ is the 'external argument' of } P)]\lambda e \lambda x[P(e) \ \& \ \text{HOLDER/UNDERGOER}(e,x)]$

Although the semantics of V_{PASS} is rather weak, we assume that it has big implications for the structure in that it prevents the merge of V_{INIT}. Also, since the passive participial form does not, by hypothesis, have any temporal (finiteness) feature, raising to Asp* will have to be supported by a dummy verb, in English, *be*. We assume that *by*-phrases in English are felicitous because of the presupposition introduced by the passive morphology that the described event actually *has* an existing external argument participant. This

---

[21]As is well known, there are poorly understood constraints on the felicity of attributive modification, even when the internal argument is being targeted. I will assume that these factors are independent of the general requirement on internal argumenthood.

restricts actual passive formation to complex events in English (i.e. ones in which the passive actually has an effect on the argument structure of the output), although this is not a universal fact about passive cross linguistically.

The properties of passive vary considerably from language to language, so the specific properties of this head must be learned; but the fact that it affects argument structure means that learners will consistently posit it as part of the eventuality description, hence inside the (extended) VP.

Above the passive and initiation heads, we posit a head V_{EVT}, which is where we place the English progressive morpheme *-ing*. This head selects for a dynamic eventuality and creates a derived eventuality description as the 'in-progress' state corresponding to the dynamic event it combines with (essentially following Parsons 1990).

(41) V_{EVT}: $\lambda Q \lambda z \lambda e' \exists e[Q(e) \ \& \ \text{InProgressState}(e',e) \ \& \ \text{Holds}(e',z)]$

Again, this is a language-specific head, posited only in case the target language manifests a progressive; but if learners understand a progressive operator to be modifying the eventuality description, then they will posit it as part of the V-domain, as here. An operator in the T-domain would not be able to directly modify the eventuality, only higher-order properties such as its temporal run-time.

In the V domain, we see a strict order $V_{\text{EVT}} > V_{\text{INIT}} > V_{\text{PASS}} > V_{\text{PROC}} > V_{\text{RES}}$. For the structure of the lexical verb, where $V_{\text{INIT}} > V_{\text{PROC}} > V_{\text{RES}}$, Ramchand (2008: ch. 3) proposes that there are only two distinct theta-assigning functors, the dynamic V_{PROC} and a stative one. When the stative one dominates V_{PROC}, it is interpreted as initiation, and when it is dominated by V_{PROC}, it is interpreted as result. Conventions of event composition prevent dynamic–dynamic or stative–stative sequences from being distinguished, so the only complex possibilities are $V_{\text{PROC}} > V_{\text{RES}}$ (with a result) and $V_{\text{INIT}} > V_{\text{PROC}}$ (with an initiator) and the combination of the two (with both).

Concerning Prog versus Pass, we think there is good language internal evidence that V_{PASS} is sensitive to argument structure in English and explicitly selects for the internal shell of a potentially complex thematic domain. However, there is also strong language internal evidence that the progressive participle does *not* care about argument structure and attaches to the fully formed event projection regardless of the number of arguments.

(42)   a.   The falling leaves.
         b.   *The destroying castle.
         c.   The dancing man.
         d.   The writing author.

The argument that the *-ing* participle abstracts over is always the one that would have been the SUBJECT of the corresponding verbal predication, *regardless of thematic role*.[22] Thus, the label V_{EVT} in our decomposition is given to any projection which operates on an *already fully articulated event structure*, as opposed to V_{RES}, V_{PROC}, V_{INIT} and V_{PASS} which are part of the event building phase which introduces arguments. We claim that PROG is a language specific formative that is a species of V_{EVT} and has learnable

---

[22]As noted before, there are additional felicity constraints on attributive modification, but these are plausibly independent of the general pattern.

and language specific selectional behaviour that indicates that it operates outside the argument structure zone.[23]

We motivated a sortal distinction between events and situations based on a convergence of many types of independent evidence, which we use to provide the underpinning for some sets of phrase structural rigidities. We have explained ordering restrictions *within* the event domain in terms of the logic of semantic composition and language specific selection for individual morphemes. This exemplifies the division of labour we outlined in the introduction. We would like to stress here that there is no a priori way of telling which types of orderings emerge from which factors. Our position is that making these kinds of analytic decisions is only made possible by looking at phenomena and orderings in detail, both within and across languages.

4.2. The embedding of events in situations

As already noted, we use the '*' diacritic to distinguish Asp* from the other heads in the system that one might want to call 'aspectual' in a pretheoretic sense.[24] Asp* selects for an eventuality description (delivered by VP) and builds a situational description that has time and world parameters, based on it. Our overarching constraint of conceptual coherence or monotonicity requires that the situational description so built include the VP event as one of its parameters. (This is the equivalent of the more traditional notion that the time variable overlap with the run time of the event.)

The Asp* head must be present in every clause, since it represents the transition from event descriptions to situation descriptions. In some cases, the morphology of Asp* is null; for example in the simple present, or in the imperative.

In the English perfect, a participle is used. We assume that the participial morphology spells out Asp*, the lowest level at which temporal dimensions are available. The spell-out of this head has a range of allomorphs, including *-(e)n* with 'strong' verbs like *grow*, hence it is sometimes called the 'EN' participle. A denotation is given in (43).

(43) Asp^*_{en}: $\lambda P \lambda s \exists e[P(e) \ \& \ \text{Transition}_1(s, e)]$

Transition is defined above as a relation between two events. We need a version here (Transition$_1$) that is basically the same except that the result event is in addition provided with a specific time parameter. Semantically, it builds a reference situation that encloses the derived event e' for which Transition(e', e) above is true. We think that the construction of a reference situation of this type is central to the meaning of the perfect, however it is ultimately best analyzed. In addition, empirically, it seems to us that the specific form in English imposes some additional requirements on that reference situation—in particular, we think that the English perfect is 'realis' and stipulates the world parameter of the reference situation as w*; depending on one's theory of the perfect it may also introduce a 'perfect time span' interval for the reference situation anchored

---

[23]It seems clear that 'progressive' per se is not a universal linguistic category, but it remains an open question whether something more abstract in the position of V_{EVT} *is*. One possibility is to identify it with the EvtP of Travis (REFS), and/or unify it with the projection that linguists have been calling Voice. These issues go beyond the scope of our concerns here.

[24]Not to be confused with the conventional '*' in e.g. *t**, the time of utterance, or *w**, the world in which the utterance occurs.

as the early end by the event and at the top end by the eventual anchoring to tense (cf. Iatridou et al. 2001).

The same participial form is used in English for the perfect and the passive. In a DM-type morphology, this kind of situation is captured by underspecification of the exponent, or in this case, the exponent class (i.e. all the morphemes which spell out the participles of the various verbs in English; *-(e)d* for *explain*, *-(e)n* for *grow*, a suppletive stem for *bring–brought*, etc.). The exponent class then is specified to substitute for a head containing the Transition relation, but underspecified for the sort of its arguments (and also underspecified for the presence of the component of the passive which demotes the external argument).[25]

Another variant of Asp* is the imperfect, which bears no overt morphology in English. It locates the reference situation, with its world and time parameters, temporally somewhere inside the temporal trace function of the event. We can model it explicitly as a head Asp*$_{\text{IMP}}$: it is an instantiation of Asp*, meaning that it takes an event-denoting complement and returns a situation-denoting constituent (hence it is in complementary distribution with other Asp*'s), but it denotes a distinct semantic relation between s and e, hence the distinguishing subscript IMP.

(44)   Asp*$_{\text{IMP}}$: $\lambda P \lambda s \exists e [P(e)$ & the temporal parameter of s is *contained in* $\tau(e)$ (the temporal trace function of e)]

4.3. The situation zone, sort s

Within the situational zone, we consistently use the label T for perspicuity, with an appropriate subscript.

T heads select for situational descriptions and deliver an updated or derived situational description. There are a number of lexical items in our system that have the syntactic category label T.

One of them is T_{PERF}, the perfect auxiliary, which selects for a situation and then builds a complex derived stative situation based on it, related to a notional 'holder'. As before, we leave the nature of the 'result' relationship and the 'holding' relationship deliberately vague in order to accommodate the different types of readings of the perfect. We have found no evidence that the different readings of the perfect occur at different heights in the phrase structure. The absence of double perfects strongly argues against such an approach (e.g. *John has had arrived in his time machine*, or (3d)).

(45) T_{PERF}: $\lambda Q \lambda x \lambda s' \exists s [Q(s)$ & s' is a stative situation that begins as a consequence of s & Holder(s', x)]

The perfect auxiliary must combine with a situational description, which makes the constituent headed by Asp* a well formed complement. V_{EVT} (*-ing*) is not a well formed complement for T_{PERF} because it denotes in the sort of events. Conversely, V_{EVT} cannot attach to T_{PERF} because it denotes a situational description and V_{EVT} requires an event description. In this way, the sortal denotations given in this system have the consequence that what we were calling Perf and Prog in the beginning of this paper can only combine in one, the attested, order. The fact that T_{PERF} selects specifically for the *-en* participle

---

[25] See Larsson and Svenonius (2013) on the analysis of the passive and the perfect participles.

should not need to be stated as an explicit morphological or selectional fact, but should fall out as a consequence of the fact that it is the only non-finite form that denotes a situation with the appropriate temporal and world parameters.

The above denotation will allow perfect *have* to combine with Asp*$_{en}$, but for constructions in which deontic modals are embedded under the perfect, an attachment site for -EN in the T domain must be available as well, and so we give the denotation for it here:

(46) T_{en}: $\lambda P \lambda s' \exists s[P(s) \ \& \ \text{Transition}_2(s', s)]$

Once again, the Transition relation will need to be adapted to accommodate a relationship between two situations: in outline, $\text{Transition}_2(s', s)$ will mean that *s′ is the result situation of s in w*, with temporal parameter t.* The two entries could be unified with judicious use of underspecified categories.

Next, we give a very general schematic for the denotation of modality. Although we have not discussed the semantics of modals in any great detail in this paper, there is one core issue that we need to establish if the general project here is to succeed. Recall that our empirical results suggested that root modals and the perfect actually inhabit the same zone in the cartography of the verbal extended projection. This, at first blush seems to fly in the face of the standard conception of tense and modality as operating in different dimensions—tense requires a semantic model expanded along the temporal dimension, for modals the model is *further* expanded along the world dimension.

However, as should already be clear, we are departing considerably from the standard conception. Following Kratzer (2008) (building in turn on Lewis 1986), we root our understanding of propositions as a whole not as sets of possible worlds, but are rather as characteristic functions of (sets of) *situations*. Situations are the key variable that all operators, whether temporal or modal, modify equally. The situational sort, as we have conceived it, itself has time and world parameters, in addition to the event description that it encloses. Situations are smaller and more specific than worlds, and have no transworld reality except via the 'counterpart' relation of Lewis (1986), but they are also larger than worlds, or events for that matter, in that they represent a richer information structure. Once this move is made, then it is clear that the situational sort is the appropriate level for both temporal and world parameter modification. This is not the place to advance and justify a specific analysis of modality. We assume however that such an analysis can be given, following the Kratzerian move, in terms of the 'accessibility' of potential situations from an anchor situation (instead of, as previously, accessibility of possible worlds), since this notion can now be stated in terms of the world *parameter* of the situations in question. We give the denotation of the modal head very abstractly here (essentially identical to that given in Kratzer 2008).

(47)   a.   ⟦ T_{might} ⟧^{c} $\lambda Q \lambda s \exists s'$: $[\text{Acc}(s)(s')][Q(s')]$
       b.   ⟦ T_{must} ⟧^{c} $\lambda Q \lambda s \forall s'$: $[\text{Acc}(s)(s')][Q(s')]$
             (after Kratzer 2008: 67)

The situational description the modal combines with is claimed to be accessible with a certain probability from the anchor situation introduced by the modal. The different modals will of course come specified with different Accessibility relations (Acc) and quantificational force.

4.4. From Situations to Propositions

In our general architecture we include a head that we label Fin* whose job is to effect the transition from the situational sort to the sort of propositions. Recall that propositions, in our narrow sense, are a semantic sort that is an enrichment of the situational sort to include a relationship to the utterance situation ('discourse linking' in the terms of Wiltschko to appear). Let us call the utterance situation s*, which can also be thought of as the richly structured Kaplanian context (Kaplan 1989 containing speaker and hearer and other deictic information).

Fin* needs to combine with a situational description to create a proposition by binding off the situational variable s, anchoring it (expressing a relationship to) the utterance situation s* (the Kaplanian context). This job is classically done by tense information in a language like English, and we give denotations for present and past tense in (48).

(48)   a.   Fin*$_{pres}$: $\lambda R \lambda p[\ p = \text{Assertion}(\exists s[R(s)\ \&\ s_t = s^*_t])]$
       b.   Fin*$_{past}$: $\lambda R \lambda p[\ p = \text{Assertion}(\exists s[R(s)\ \&\ s_t \neq s^*_t])]$

In general, modals also carry anchoring information (it is just that they anchor via the world parameter, rather than the temporal parameter). In the world domain, non-equality of the world parameter is irrealis; in the temporal domain it is past (cf. also Iatridou 2000):

(49)   a.   Fin*$_{realis}$: $\lambda P \lambda s \lambda p[p = \text{Assertion}(\exists s[R(s)\ \&\ s_w = s^*_w])]$
       b.   Fin*$_{irrealis}$: $\lambda P \lambda s \lambda p[p = \text{Assertion}(\exists s[R(s)\ \&\ s_w \neq s^*_w])]$

We will assume that the proposition is a relationship between a situation and an assertor and contains information about the speaker and speaker attitude as well as encoding of familiarity and novelty of the information to the members of the utterance situation (participants in the speech act.)

At this point, we can also be a little bit more specific about the evidence that epistemic modals are higher up in the structure than root modals. Evidentials and epistemics also sit in the propositional sortal zone because they quantify over semantic objects that are rich enough to include information about the speaker (speaker knowledge and evidence) (see Ritter and Wiltschko 2009, Sigurðsson 2004, Bianchi 2003, Giorgi 2010). In Ramchand (2012), a more fleshed out analysis along these lines is given, in which epistemic modality involves quantification over situations with fixed world and time variables, but free speaker-oriented parameters. This is contrasted with circumstantial modality, which requires quantification over situations with free world and time variables. Under the view proposed in Ramchand (2012), the domain for the accessibility relation is constrained by the semantic sort of the modal complement. In this way, one derives the result that the very same modal gets a circumstantial interpretation when merged within the situational sortal domain, but an epistemic interpretation when merged in the propositional sortal domain (a different implementation of the same intuition pursued in Hacquard 2006 and Butler 2006).

5. Adverbs

In this final section, we consider the evidence from adverbs to see whether they provide grounds for introducing more functional heads than we have so far assumed in

our minimalist cartography of the English verbal extended projection. In languages like English, the ordering of functional elements motivates something on the order of a half-dozen positions in the extended projection of the verb. A few languages have richer morphology and motivate twice or even three times as many positions, but it is rare for a single language to offer evidence for more than ten morphological slots of relevance to the functional hierarchy, and when a language does show rigidly ordered morphological slots, it is possible that morphological factors contribute to the rigidity of the ordering.

For this reason the adverb evidence presented by Cinque (1999) is vital; he argues that the majority of the functional heads attested in TAM morphology in other languages are corroborated by adverbs in languages like English and Italian, and that furthermore the orderings manifested by those adverbs corroborate the pairwise orderings of the corresponding functional heads.

Subsequent work has supported Cinque's findings concerning adverb ordering in other languages (e.g. Nilsen 1997 for Norwegian, Beijer 2005 for Swedish, Alexiadou 1997 for Greek, Rackowski and Travis 2000 for Malagasy, etc.).

Since the adverb facts seem to support a richer universal hierarchy than we have deduced, we must examine the data carefully. Our account predicts that adverbs will be restricted to domains according to what sorts of elements they modify; an event-modifying adverb should be confined to the e-domain, and if preverbal will therefore follow a situation-modifying adverb, which is confined to the s-domain (cf. also Ernst 2002).

Under close examination of pairs of adverbs, we find several different situations, listed in (50):

(50) a. Cases where order is rigidly determined by sortal domains
      b. Cases where order is flexible within a sortal domain, as predicted by our account
      c. Cases where an adverb can be inserted in either of two sortal domains, but with the difference in meaning predicted by our account
      d. Cases where an extrinsic factor restricts the ordering

We discuss these in turn in the following subsections.

5.1. Order forced by sortal distinctions

There are fairly straightforward cases in which order is forced by sortal distinctions. For example, evidential, epistemic, and speaker-comment adverbials must precede anything that modifies the situation (in the preverbal space), and situation adverbs must precede event-modifying adverbs (cf. Jackendoff 1972, McConnell-Ginet 1982, Ernst 2002).

(51) a.   John fortunately already knows that.
      b. *John already fortunately knows that.
      c.   John already quietly declined.
      d. *John quietly already declined.

This kind of example is probably the best understood, and has been treated repeatedly in the previous literature, so we will not dwell on it further here, noting only that the general evidence is consistent with the number of sortal domains we are already assuming here.

5.2. Order flexible within a sortal domain

As an example of (50b), flexible order, consider the facts of *twice*, originally discussed in Andrews (1983) and illustrated in (52).

(52) a. John intentionally knocked on the door twice
      b. John twice intentionally knocked on the door.
      c. John knocked on the door intentionally twice.
      d. John knocked on the door twice intentionally.
      e. ??John intentionally twice knocked on the door.

Examples (52c) and (52d) on their own might seem to pose a problem for the rigid ordering of a functional sequence that hosts the relevant adverbs. However, Cinque (1999) argues that the conflict is illusory.

> "The paradox however, is not real, as there is evidence that *twice* belongs to a class of adverbs (*many, few, etc. times, often, rarely, frequently*, etc.) that are systematically ambiguous between two interpretations, each associated with a different position. The higher position quantifies over the entire event (saying how *frequently* it takes place). In [(52c)], for example, it says that there were two events of knocking on the door (intentionally). The lower position, instead, just indicates the *repetition* of the act denoted by the verb. So [(52d)] says that there was a single event of (intentional) repetition of the act of knocking on the door."
> Cinque (1999:26)

As Cinque notes, the right core orders and scopes are derivable from the single underlying structure in (53), with one position for *intentionally* and two positions and interpretations for *twice*.

(53) John (twice$_1$) [$_{XP}$ intentionally [$_{YP}$ knocked (twice$_2$) on the door.]]

Thus far we completely agree with the argumentation. We disagree with the next step which argues that there are therefore two functional heads Asp$_{freq}$ and Asp$_{rep}$ which are ordered on either side of Mod$_{volitional}$ (proposed on Cinque's p. 106). We agree with the idea that there are (at least) two base positions for *twice*. It also seems to us that the two relevant positions must be *within* the event sortal domain, below what we have been calling Asp* in the sections above. The reason for this is that both positions and interpretations are possible under a passive auxiliary, as shown here, which has the same pattern of judgements as (52) above.

(54) a. John has been twice intentionally insulted.
      b. John has been intentionally insulted twice.
      c. John has been insulted twice intentionally.
      d. John has been insulted intentionally twice.
      e. ??John has been intentionally twice insulted.

We think that one should build semantic representations that 'update' the event variable after functional composition, to reflect the increased complexity of the event description. In this way, can can capture the scopal effects and sensitivity of the adverb to the

particular sister it modifies. In (55), we derive the reading for *twice > intentionally*; in (56) we show the derivation for *intentionally > twice*.

(55)   John $[_{VP''}$ twice$_1$ $[_{VP'}$ intentionally $[_{VP}$knocked on the door $]]]$
    a.   VP = ⟦knocked on the door⟧
          = λe[knock(e) & on-the-door(e)]
    b.   VP$'$ = ⟦**intentionally** [knocked on the door]⟧
          = $\lambda e'$[intentional(e$'$) & $\exists e$[e $\subset$ e$'$ & $[_{VP}$knock(e) & on-the-door(e)]]]
    c.   VP$''$ = ⟦**twice** [**intentionally** [knocked on the door]]⟧
          = $\lambda e''$[twice(e$''$) & $\exists e'$[e$'$ $\subset$ e$''$ & $[_{VP'}$intentional(e$'$) & $\exists e$[e $\subset$ e$'$ & $[_{VP}$knock(e) & on-the-door(e)]]]]]

(56)   John $[_{VP''}$ intentionally $[_{VP'}$ $[_{VP}$knocked] twice$_2$ on the door]].
    a.   VP = ⟦knocked on the door⟧
          = λe[knock(e) & on-the-door(e)]
    b.   VP$'$ = ⟦**twice** [knocked on the door]⟧
          = $\lambda e'$[twice(e$'$) & $\exists e$[e $\subset$ e$'$ & $[_{VP}$knock(e) & on-the-door(e)]]]
    c.   ⟦**intentionally** [**twice** [knocked on the door]]⟧
          = $\lambda e''$[intentional(e$''$) & $\exists e'$[e$'$ $\subset$ e$''$ & $[_{VP'}$twice(e$'$) & $\exists e$[e $\subset$ e$'$ & $[_{VP}$knock(e) & on-the-door(e)]]]]]

Crucially, every time a new variable is introduced, the lower one is existentially closed. This gives surface scope without the need to change semantic sorts. The updating of the event variable is not free, but is constrained by the general principle of compositional coherence we have already assumed is necessary to constrain the sortal transitions. In other words, we do not lose information, but we enrich the event description by some extra information to create a new derived event description.

Recall our general principle of semantic compositionality in (29), repeated here:

(57)   **Compositional Coherence**: If X embeds YP, then the denotation of XP is a monotonically coherent elaboration of the denotation of YP.

5.3. Adverb ordering flexibility due to sortal underspecification

We also find cases where a single adverb is underspecified, allowing insertion in two different sortal domains. In such cases we predict a difference in readings compatible with the different content of the sortal domains. The following is a straightforward example.

(58)   a.   John lifted the box easily.
      b.   John is easily the best box lifter in the room.

Such cases are lexically restricted, unlike the situation for *twice*, whose behavior is typical of frequentative adverbs.

(59)   a.   John lifted the box laboriously.
      b. *John is laboriously the best box lifter in the room.

In our story this difference in *easily* is a case of sortal ambiguity, and must be distinguished from the *twice* situation, where we deny that any articulation of the functional sequence is required or desirable.

To formalize the underspecification, we either need to identify some features that the different sortal arguments have in common or else posit abstract categories which encompass more than one class of sortal argument. To make such a proposal would require some rather extended discussion of adverb classes; at this point we refer the reader to Ernst (2002) for extensive discussion of this and other ways in which adverbs exhibit flexibility in their position and mode of combination.

5.4. Other factors

We also find cases where adverb ordering is constrained by other factors, as discussed by Nilsen (2003), Ernst (2007; 2009). For example, take the order *probably ≺ once*, noted by Cinque:

(60)   a.   John was probably once married.
          b. *John was once probably married.

Since *once* binds t and *probably* quantifies over w, and t and w are parameters of s, these are both located in the same sortal domain. However, worlds are not independent of times. In (60a), the likely worlds each contains a 'once'-time in their past at which John was married; but in (60b), we are looking for a 'once'-time in the past at which John's being married was likely. We suggest that this is formally possible but pragmatically anomalous (thus the '*? in (60b) should technically be a #). It should be acceptable in the right context, for example a context in which it is perceived that the probability of a certain possible state of affairs changed dramatically after another, actual, state of affairs obtained.

Suppose that in 2006, the Rolling Stones were in negotiations with a concert promoter in Norway, who seemed to be on the cusp of signing them to play a concert in Tromsø as part of their European tour. Then, before negotiations were concluded, the guitarist Keith Richards fell out of a tree and injured his head, requiring surgery and causing the tour to be shortened. Negotiations were cut off, and the Rolling Stones didn't play in Tromsø. Years later, the dejected concert promoter might felicitously utter (61).

(61)   The Rolling Stones were once probably going to play in Tromsø.

6. Conclusion

Taking the Minimalist Program seriously, we are forced to reject the rich functional hierarchy as an axiomatic part of UG; there is no plausible evolutionary scenario to support the natural selection of a language faculty with such a highly structured organization of functional features.

But taking the results of the Cartographic enterprise seriously, we are forced to seek a source for the rich functional hierarchy. This holds no matter how much of the specifics of cartography one accepts—even the pared-down C-T-v-V is a functional hierarchy in need of explanation.

We have argued that the rich functional hierarchy has multiple sources, and we suggest that progress will be impeded if the functional hierarchy is ignored (as in some Minimalist work) or taken for granted (as in some Cartographic work).

The most important source that we identify is grounded, we argue, in extralinguistic cognition: A cognitive proclivity to perceive experience in terms of events, situations,

and propositions (with analogous ontologies for other extended projections). Granted, we have little direct evidence for these posited proclivities apart from the explananda themselves; but at present we do not know of plausible alternatives. By making concrete proposals about the nature of the cognitive underpinnings of the syntactic categories, we hope to push investigation forward in the direction of testable hypotheses with replicable results.

To be concrete, we have suggested that an EVENT is recognized as a special kind of object, with thematic participants, and that an event combined with a certain special kind of parameter (an 'anchor' in Wiltschko's to appear sense), for example a time, is a SITUATION, a conceptually different 'sort' of object from an event. A situation includes an event as a privileged part, somewhat like the way a person includes a body. Similarly, a situation merged with another special kind of parameter (a 'discourse link' in Wiltschko's sense) is a PROPOSITION, which is again a different sort from, and constitutively includes, the situation.

Additional sources of ordering, we suggest, must be distinguished from the ordering of the sortal domains. We invoked a variety of such sources at different points. For the order of the English modals over the perfect, we suggested in §2.2 that this is because the modals are lexically specified to include Fin*, the transition to the C-domain. When we turn to a language in which the modals have nonfinite forms, we see that epistemic modals are nonetheless rigidly ordered over the perfect. We suggested in §4.4 that this is because epistemic modals are linked to the proposition, hence are at least as high as Fin*, whereas the perfect, being temporal, is interpreted in the T-domain. Since deontic modals are also interpreted in the situational zone, they are freely ordered with respect to the perfect, subject to language-specific constraints.

For the order of the progressive over the passive, we suggested in §4.1 that this was because of selectional restrictions of the progressive and passive heads.

In other cases, we have suggested that an apparently universal ordering might arise illusorily as an effect of labeling. For example, we adopted Ramchand's (2008) proposal that states and processes are freely combinable in either order, but that a state leading to a process is understood as an initiation, and a state to which a process leads is understood as a result; this gives the illusion of a hierarchy INITIATION > PROCESS > RESULT (§4.1). We suggested in §5.2, in our discussion of *twice*, that a functor with a meaning of repetition might get different interpretations in different positions, corresponding to Asp_{FREQ} and Asp_{REP}, without that motivating an underlying difference or any rigidity of hierarchy.

The end result, we think, suggests that cartographic findings are fully compatible with Minimalist principles. But we make an even stronger claim here, namely that we cannot actually make good on minimalist demands for explanatory adequacy without taking cartographic patterning into account. Conversely, cartographic generalizations need to be decomposed into their component parts in order to answer the deepest questions concerning the interplay between the universal and the particular in language patterning.

Aelbrecht, L., Harwood, W., 2013. To be or not to be elided: VP ellipsis revisited, ms. University of Ghent; available at ling.auf.net/lingbuzz/001609.

Akmajian, A., Wasow, T., 1975. The constituent structure of VP and AUX and the position of the verb *be*. Linguistic Analysis 1 (3), 205–245.

Alexiadou, A., 1997. Adverb Placement: A Case Study in Antisymmetric Syntax. John Benjamins, Amsterdam.

Andrews, III, A., 1983. A note on the constituent structure of modifiers. Linguistic Inquiry 14 (4), 695–697.

Austin, J. L., 1950. Philosophical Papers. Oxford University Press, Oxford.

Bach, E., 1986. The algebra of events. Linguistics and Philosophy 9 (1), 5–16.

Baltin, M. R., 1989. Heads and projections. In: Baltin, M. R., Kroch, A. S. (Eds.), Alternative Conceptions of Phrase Structure. University of Chicago Press, Chicago, pp. 1–16.

Barwise, J., Perry, J., 1983. Situations and Attitudes. MIT press, Cambridge, Ma.

Beijer, F., 2005. On the relative order of adverbs in the I-domain: A study of English and Swedish. Ph.D. thesis, Lund University.

Bianchi, V., 2003. On finiteness as logophoric anchoring. In: Guéron, J. (Ed.), Temps et Point de Vue/Tense and Point of View. Université de Paris X, Nanterre, pp. 213–246.

Bjorkman, B., 2011. BE-ing default: The morphosyntax of auxiliaries. Ph.D. thesis, MIT, Cambridge, Ma.

Bošković, Ž., 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. Linguistic Inquiry 45 (1), 27–89.

Butler, J., 2006. The structure of temporality and modality (or, Towards deriving something like a Cinque Hierarchy). Linguistic Variation Yearbook 6, 161–206.

Chomsky, N., 1957. Syntactic Structures. Mouton, 's-Gravenhage.

Chomsky, N., 2005. Three factors in language design. Linguistic Inquiry 36, 1–22.

Chomsky, N., 2008. On phases. In: Freidin, R., Otero, C. P., Zubizarreta, M. L. (Eds.), Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud. MIT Press, Cambridge, Ma., pp. 133–166.

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

Cinque, G., 2004. Restructuring and functional structure. In: Belletti, A. (Ed.), Structures and Beyond: The Cartography of Syntactic Structures, vol. 3. Oxford, New York, pp. 132–191.

Cinque, G., Rizzi, L., 2010. The cartography of syntactic structures. In: Heine, B., Narrog, H. (Eds.), The Oxford Handbook of Linguistic Analysis. Oxford University Press, Oxford, pp. 51–65.

Condoravdi, C., 2002. Temporal interpretation of modals: Modals for the present and for the past. In: Beaver, D. I., Casillas Martínez, L. D., Clark, B. Z., Kaufmann, S. (Eds.), The Construction of Meaning. CSLI, Stanford, Ca., pp. 59–88.

Cowper, E., 2005. The geometry of interpretable features: Infl in English and Spanish. Language 81 (1), 10–46.

Davidson, D., 1967. The logical form of action sentences. In: Rescher, N. (Ed.), The Logic of Decision and Action. University of Pittsburgh Press, Pittsburgh, Pa, pp. 81–95.

Demirdache, H., Uribe-Etxebarria, M., 2000. The primitives of temporal relations. In: Martin, R., Michaels, D., Uriagereka, J. (Eds.), Step by Step. Essays on Minimalist Syntax in Honour of Howard Lasnik. MIT Press, Cambridge, Ma., pp. 157–186.

Dowty, D. R., 1979. Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ. Reidel, Dordrecht.

Eide, K. M., 2005. Norwegian Modals. Mouton de Gruyter, Berlin.

Ernst, T., 2002. The Syntax of Adjuncts. Cambridge University Press, Cambridge.

Ernst, T., 2007. On the role of semantics in a theory of adverb syntax. Lingua 117 (6), 1008–1033.

Ernst, T., 2009. Speaker-oriented adverbs. Natural Language & Linguistic Theory 27 (3), 497–544.

Gawron, J. M., 2006. Generalized paths. In: Georgala, E., Howell, J. (Eds.), Proceedings of SALT XV. CLC, Ithaca, NY, pp. 135–150.

Giorgi, A., 2010. Towards a Syntax of Indexicality. Oxford University Press, Oxford.

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

Grohmann, K., 2003. Prolific Domains: On the Anti-locality of Movement Dependencies. John Benjamins, Amsterdam.

Hacquard, V., 2006. Aspects of modality. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Hale, K., Keyser, S. J., 1993. On argument structure and the lexical expression of syntactic relations. In: Hale, K., Keyser, S. J. (Eds.), The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger. No. 24 in Current Studies in Linguistics. MIT Press, Cambridge, Ma., pp. 53–109.

Harley, H., Ritter, E., 2002. Person and number in pronouns: A feature-geometric analysis. Language 78 (3), 482–526.

Harwood, W., 2013. Being progressive is just a phase: Dividing the functional hierarchy. Ph.D. thesis,

University of Ghent.

Harwood, W., to appear a. Being progressive is just a phase: Celebrating the uniqueness of progressive aspect under a phase-based analysis. Natural Language and Linguistic Theory.

Harwood, W., to appear b. Rise of the auxiliaries: A case for auxiliary raising vs. affix lowering. The Linguistic Review.

Higginbotham, J., 1985. On semantics. Linguistic Inquiry 16 (4), 547–593.

Hinzen, W., 2006. Mind Design and Minimal Syntax. Oxford University Press, Oxford.

Iatridou, S., 2000. The grammatical ingredients of counterfactuality. Linguistic Inquiry 31 (2), 231–270.

Iatridou, S., Anagnostopoulou, E., Izvorski, R., 2001. Observations about the form and meaning of the perfect. In: Kenstowicz, M. (Ed.), Ken Hale: A Life in Language. MIT Press, MIT, Cambridge, Ma., pp. 189–238.

Jackendoff, R., 1972. Semantic Interpretation in Generative Grammar. No. 2 in Current Studies in Linguistics. MIT Press, Cambridge, Ma.

Jackendoff, R., 1990. Semantic Structures. No. 18 in Current Studies in Linguistics. MIT Press, Cambridge, Ma.

Jackendoff, R., 2002. Foundations of Language. Oxford University Press, Oxford.

Kaplan, D., 1989. Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In: Almog, J., Perry, J., Wettstein, H. (Eds.), Themes from Kaplan. Oxford University Press, Oxford, pp. 481–563, [originally circulated in 1977].

Klein, W., 1994. Time in Language. Routledge, London and New York.

Kratzer, A., 1989. An investigation of the lumps of thought. Linguistics and Philosophy 12, 607–653.

Kratzer, A., 2008. Modals and conditionals again, ms. University of Massachusetts, Amherst.

Larsson, I., Svenonius, P., 2013. English and Scandinavian participles and the syntax-morphology interface, paper presented at the 25th Scandinavian Conference of Linguistics in Reykjavík.

Lasnik, H., 1995. Verbal morphology: *Syntactic Structures* meets the Minimalist Program. In: Campos, H., Kempchinsky, P. (Eds.), Evolution and Revolution in Linguistic Theory: Essays in honor of Carlos Otero. Georgetown University Press, Washington, D.C., pp. 251–275.

Legate, J. A., 2003. Some interface properties of the phase. Linguistic Inquiry 34 (3), 506–516.

Lewis, D. K., 1986. On the Plurality of Worlds. Blackwell, Oxford.

McCawley, J. D., 1971. Tense and time reference in English. In: Fillmore, C. J., Langendoen, D. T. (Eds.), Studies in Linguistic Semantics. Holt, Rinehart, and Winston, New York, pp. 97–113.

McConnell-Ginet, S., 1982. Adverbs and logical form: A linguistically realistic theory. Language: Journal of the Linguistic Society of America 58 (1), 144–184.

Milsark, G. L., 1974. Existential sentences in English. Ph.D. thesis, MIT.

Nilsen, Ø., 1997. Adverbs and A-shift. Working Papers in Scandinavian Syntax 59, 1–31.

Nilsen, Ø., 2003. Eliminating positions: Syntax and semantics of sentential modification. Ph.D. thesis, Universiteit Utrecht, Utrecht.

Parsons, T., 1990. Events in the Semantics of English: A Study in Subatomic Semantics. MIT Press, Cambridge, Ma.

Percus, O., 2000. Constraints on some other variables in syntax. Natural Language Semantics 8 (3), 173–229.

Platzack, C., 2000. Multiple interfaces. In: Nikanne, U., van der Zee, E. (Eds.), Cognitive Interfaces: Constraints on Linking Cognitive Information. Oxford University Press, Oxford, pp. 21–53.

Platzack, C., 2010. Den fantastiska grammatiken: En minimalistiska beskrivning av svenskan. Norstedts, Stockholm.

Rackowski, A., Travis, L., 2000. V-initial languages: X or XP movement and adverbial placement. In: Carnie, A., Guilfoyle, E. (Eds.), The Syntax of Verb Initial Languages. Oxford University Press, New York, pp. 117–141.

Ramchand, G., 2008. Verb Meaning and the Lexicon. Cambridge University Press, Cambridge.

Ramchand, G., 2012. Argument structure and argument structure alternations. In: den Dikken, M. (Ed.), The Cambridge Handbook of Generative Syntax. Cambridge University Press, Cambridge, pp. 265–321.

Ramchand, G., Svenonius, P., 2004. Prepositions and external argument demotion. In: Solstad, T., Lyngfelt, B., Krave, M. F. (Eds.), Pre-proceedings of Demoting the Agent: Passive and other Voice-related Phenomena. University of Oslo, Oslo, pp. 93–99.

Reichenbach, H., 1947. Elements of Symbolic Logic. Macmillan, New York.

Ritter, E., Wiltschko, M., 2009. Varieties of INFL: Tense, location, and person. In: Hans Broekhuis, J. v. C., van Riemsdijk, H. (Eds.), Alternatives to Cartography. Mouton de Gruyter, Berlin, pp. 153–202.

Sailor, C., 2012. Inflection at the interface, ms. UCLA.

Schachter, P., 1983. Explaining auxiliary order. In: Heny, F., Richards, B. (Eds.), Linguistic Categories: Auxiliaries and Related Puzzles. Vol. 2. D. Reidel, Dordrecht, pp. 145–204.

Shlonsky, U., 2010. The cartographic enterprise in syntax. Language and Linguistics Compass 4 (6), 417–429.

Sigurðsson, H. Á., 2004. The syntax of person, tense and speech features. Italian Journal of Linguistics 16, 219–251.

Svenonius, P., 2004. On the edge. In: Adger, D., de Cat, C., Tsoulas, G. (Eds.), Peripheries: Syntactic Edges and their Effects. Kluwer, Dordrecht, pp. 261–287.

Svenonius, P., 2012. Spanning, ms. University of Tromsø, available at ling.auf.net/lingBuzz/001501.

Warner, A. R., 1986. Ellipsis conditions and the status of the English copula. York Papers in Linguistics 12, 153–172.

Wiltschko, M., to appear. The Universal Structure of Categories: Towards a Formal Typology. Cambridge University Press, Cambridge.

Wolff, P., 2006. Dynamics and the perception of causal events. In: Shipley, T., Zacks, J. (Eds.), Understanding Events: How Humans See, Represent, and Act on Events. Oxford University Press, Oxford, pp. 555–587.