The Workspace as a source of Hierarchy in Extended Projections

Peter Svenonius
CASTL, University of Tromsø – The Arctic University of Norway

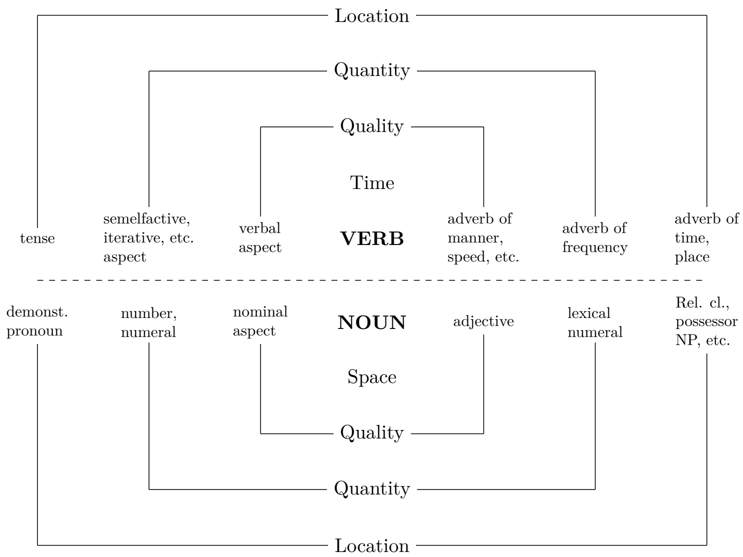
Abstract

Extended projections (EPs) in natural languages have several properties which have not yet been explained. (i) EPs conform to a Hierarchy of Projections (HoP), a crosslinguistically similar hierarchical arrangement of semantically grounded categories. (ii) Each token of an EP is linear, in the sense that it has a single dimension (with specifiers and adjuncts stripped out, it is a string). (iii) HoPs are rooted in a lexical category with conceptual content at the bottom, and a succession of functional elements above. I argue that these properties motivate a particular architecture of the workspaces in which sentences are constructed. I model the workspace as a Finite State Automaton (FSA) with a monotonicity property which underpins hierarchy. The FSA starts from a lexical category (cf. (iii)), 'projecting' it into an EP. Transitions correspond to applications of Merge, and states are stages in the derivation. The sequence of states in a path from start to accepting state is a string, the EP (cf. (ii)). The HoP is then the entire FSA, arranged with the start at the bottom and the final state, the complete clause or noun phrase, at the top (cf. (i)).

1. What is an extended projection?

There is a broad consensus in the field for a coarse-grained universal hierarchy in the clause, for example C > T > V, see for example Ramchand and Svenonius (2014) or Wiltschko (2014) for two examples of implementations. These coarse-grained collections of features are often referred to as DOMAINS, as in the V-domain, the T-domain, and the C-domain. The motivation and outlines of the hierarchy of domains match the layered structures of functionalist work such as Dik (1989) and Hengeveld (1990), and illustrated in (1) from Rijkhoff (2002), so at a certain level of abstraction, the hierarchical structure of the clause is a consensus that transcends frameworks, an all too rare phenomenon.

(1)



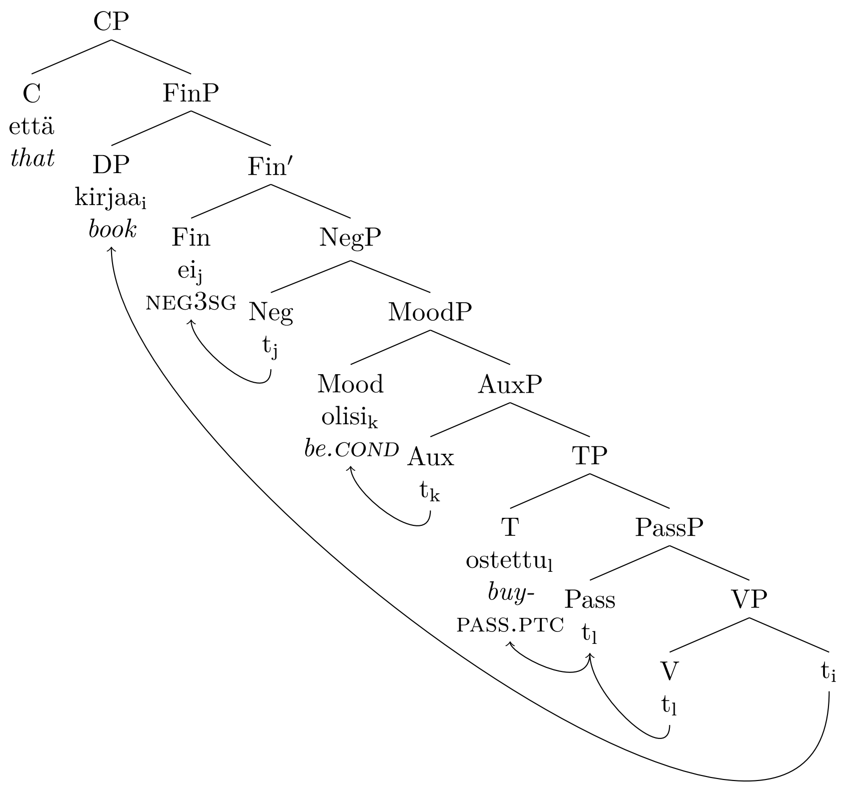
 Nordlyd 48.1: 115-132, Special issue on Extended Projections, edited by Peter Svenonius.
Published at the University of Tromsø – The Arctic University of Norway.
http://septentrio.uit.no/index.php/nordlyd https://doi.org/10.7557/12.7982

A related hierarchy is usually recognized in the noun phrase, and something like it is necessary to capture the observations of Cinque (2005). Again, this is recognized in functionalist work as well; the lower half of Rijkhoff's diagram in (1) represents the noun phrase, the upper half the clause.

1.1. Fine-grained language-specific hierarchies of projections

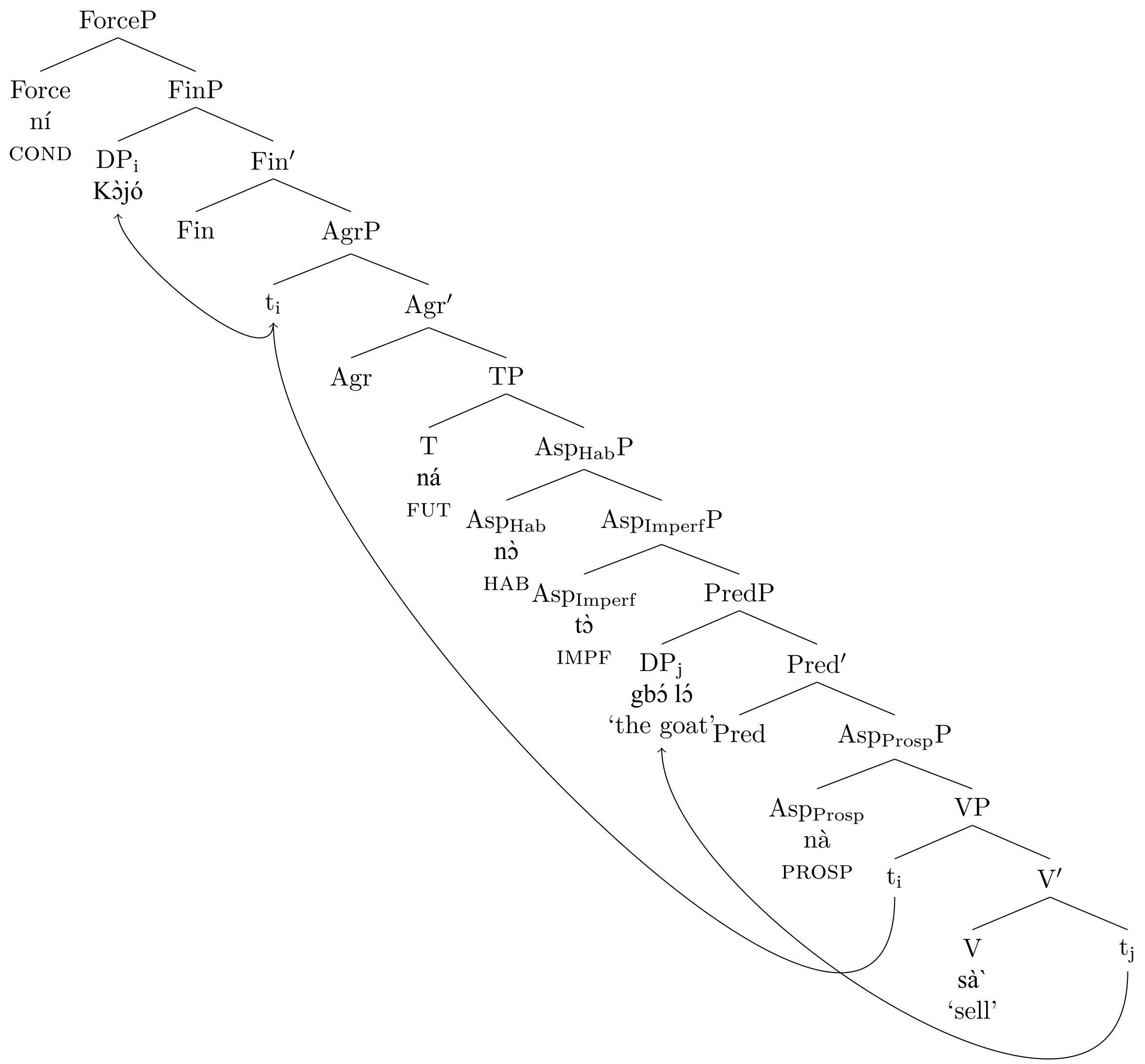
The detailed study of individual languages convincingly establishes more fine-grained hierarchical organization. For example, Holmberg et al. (1993) propose the following "maximal expansion" of the clause for Finnish, illustrated with 'that the book would not have been bought':

(2)



Aboh (2004:157–190) proposes the following structure for a Gungbe clause meaning 'if Kojo will often be about to sell the specific goat...'.

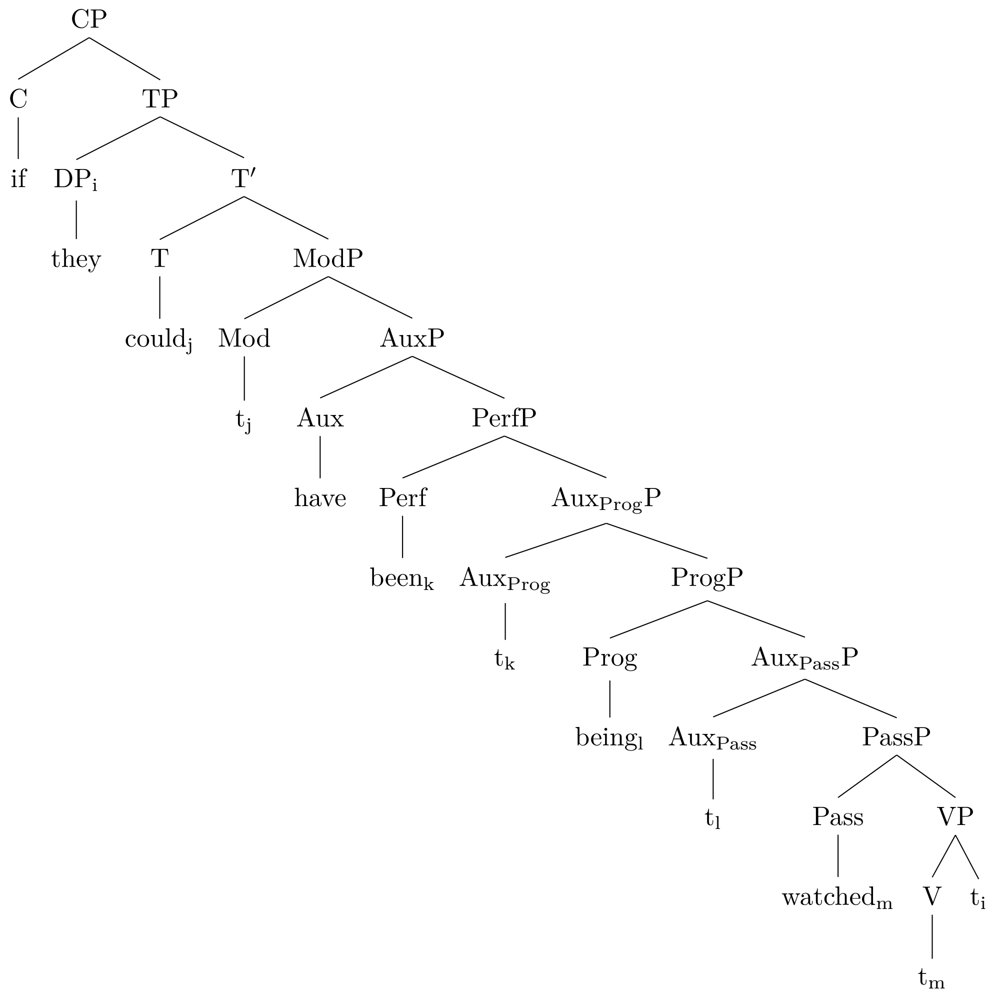
(3)



It is easy to find other examples of clause structures requiring ten or more heads for a single sentence. Work on Athabaskan languages (e.g., Rice 2000) and on Iroquoian languages (e.g., Lounsbury 1949) identify ten or more 'slots' for distinct classes of inflectional morphemes. Crippen (2012) identifies 18 prefix slots and five non-derivational suffix slots for Tlingit.

Even in English, if each auxiliary counts as a distinct head, and if *could* is analyzed as *can* plus past tense, then *If they could have been being watched* involves ten heads in one clause.

(4)



One hypothesis, advanced in particular by Cinque (1999), is that all of these language-specific hierarchies of projections (HoPs) are subcases of a single universal hierarchy, a fine-grained "functional sequence" (a term used by Gruber 1997:178).

1.2. Properties of Extended Projections

I will use Grimshaw's (1991, 2005) term EXTENDED PROJECTION for a structure which conforms to an HoP. In common parlance, an extended projection is a hierarchically organized collection of functional categories, normally with a lexical category or acategorial root at the bottom, in which the higher take the lower as dependents. Some properties which extended projections have been observed to have are listed in (5).

(5) a. **Hierarchy**: There is a Hierarchy of Projections. Each token of an extended projection (EP) conforms to a Hierarchy of Projections (HoP), a hierarchy of semantically grounded syntactic categories.

b. **Linearity**: Each token of an EP is linear, in the sense that it has a single dimension (without specifiers and adjuncts it is a string).

c. **Rootedness**: HoPs are rooted in a lexical category associated with conceptual content at the bottom, with a succession of functional elements above.

Note that while individual tokens of EPs are assemblies of syntactic formatives (heads), the HoP is stated over categories. In many languages, positions can be identified in the verbal EP in which a set of formatives is in complementary distribution, for example the modals of English. In Cup'ik as described by Woodbury (2024), there is a category Realization1 which includes *-yaaqe-* 'in vain,' *-ngate-* 'seem,' and *-ksaite-* 'not yet'. This category occupies a position between a set of aspectual markers and the tense. If these three elements are in complementary distribution, they belong to the same category.

Languages apparently vary on this score; where two concepts are mutually exclusive in one language, they may be combined in another (e.g., modals are mutually exclusive in English but may be combined in Norwegian). Typically, when they can be combined, the order is strict (Cinque 1999; for example when modals are combined a deontic modal cannot appear outside an epistemic modal).

1.3. Extended projections are like words

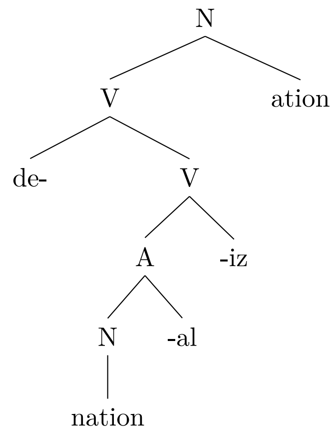
Extended projections are in an important way like words. In many cases, an extended projection spells out as a single word. For example, in the sentence *Flowers bloomed*, there are two extended projections, an extended projection of N spelled out as the word *flowers* and an extended projection of V spelled out as the word *bloomed.* In languages with richer morphology than English has, it is common that an extended projection corresponds to a single word.

In other cases, an extended projection spells out as multiple words. For example in *The birds are singing*, there are two extended projections, the extended projection of N, spelling out as *the birds*, and the extended projection of V, spelling out as *are singing.* In *The migratory birds are singing loudly* there are four, the two from the previous example plus the adjectival extended projection *migratory* and the adverbial extended projection *loudly*, assuming that these are adjuncts (or specifiers); the extended projection consists only of the heads which stand in a transitive complementation relation with the lexical head at the bottom.

Just as with words, it is possible to construct paradigms for extended projections of a given lexical root. For example, some descriptive grammars of English provide tables in which forms like *are singing* and *have sung* are listed in paradigm tables.

It is well known that words can have internal structure, for example in [*de*-[*nation-al-iz*]]-*ation*, the verb-to-verb prefix *de-* must be structurally located between the deadjectival verbalizing suffix *-ize* and the deverbal nominalizing suffix *-(a)tion*, as illustrated in (6) (cf. e.g., Carden 1983).

(6)



But the hierarchical structure itself (N-A-V-V-N, reading upward from the bottom) is nonbranching, in that it contains no node immediately dominating more than one branching node. Arguably rare or absent are complex affixes which must be composed before combination with the base. For this reason, models of morphology do not build in mechanisms for affixes to combine with other affixes, prior to combining with a base. Instead, all operations apply to the base, sequentially.

Consider the Icelandic definite plural *hest-ar-nir* 'the horses' (nominative), where *-nir* marks definiteness; the indefinite plural is *hest-ar* 'horses' (nominative). The definite form has the structure [*hest-ar*]-*nir* rather than *hest*-[*ar-nir*], as can be seen when there is an irregular plural, the definite is always regularly and compositionally combined with the base including the plural, for example in the irregular plural *faðir∼feður* 'father∼fathers (nominative)', where the root ends in /r/ (the accusative singular form, where no suffix containing /r/ is expected, is *föður*), the definite form is *feður-nir* 'the fathers (nominative)'.

There are cases for which it cannot be proven that a structure is nonbranching in the sense used here. For example, the definite suffix *-nir* can be decomposed into a definite *-n-* followed by a masculine plural nominative agreement *-ir*. There is no semantic motivation to distinguish between a strictly left-branching bracketing [[*hest-ar*]-*n*]-*ir* and the alternative with a branching noncomplement [*hest-ar*]-[*n-ir*], because the information encoded in *-ir* is redundant. But semantically motivated cases overwhelmingly involve linear structure.[1]

The fact that non-compound words are built from linear structures does not mean that the full paradigm of a given root must be small. With just ten inflectional categories which each can independently either be present or not present, there are 2^{10} or 1024 forms. A language with much productive morphology can have exceedingly large paradigms, and if derivational morphology is included, there can be millions of forms. An example of an inflected Turkish verb is given in (7).

(7) in-dir-il-e-mi-yebil-ecek-ler
descend-CAUS-PASS-ABLE-NEG-ABLE-FUT-3PL
'They will be able to resist being brought down' (Turkish, Hankamer 1989:395)

Hankamer (1989) calculates that the number of verb forms that can be formed from a single verb root in Turkish is 1.8 million ($<2^{21}$), without any recursive procedures (though this would include some morphology which would usually be characterized as derivational). For Turkish nouns, Hankamer calculates over 9 million forms without recursion.

In morphology, locality is closely linked to adjacency. For category selection, structural adjacency is relevant, while morphophonological processes may be sensitive to linear adjacency (cf. Embick 2010, Fábregas 2013 for some discussion of cases in which different kinds of adjacency appear to matter). Portmanteau spell-out of two heads may require both linear and structural adjacency. Illustrating with (6), *-al* is structurally and linearly adjacent to both *nation* and *-ize*. The prefix *de-* is linearly adjacent to *nation*, but not structurally, and it is structurally adjacent to *-ize* and to *-ation*, but not linearly. In these terms, the plural suffix *-en* is selected when it is linearly adjacent to a stem such as *ox* (structural adjacency may also be necessary). If another noun is derived from that stem by a suffix, for example *ox-hood*, the plural is no longer adjacent to the *ox* and the special plural is no longer conditioned; the plural would be *oxhoods*, not **oxhooden*. Nonlocal dependencies are always due to syntax. For example in Icelandic *hest-ar-n-ir* 'the horses', the inner suffix *-ar* and the outer suffix *-ir* match in masculine gender, plural number, and nominative case, and are separated by the definite *-n-*; but it is syntax and not morphology that is responsible for copying ϕ features from the inner structure to the outer head, and case features from the outer structure to the inner head.[2]

[1] Compounds, on the other hand, can clearly involve branching non-heads, for example in *football player*. Some cases which appear to involve branching affixes actually involve phrasal incorporation.

[2] An operation like Agree, which copies features from one place to another across an intervening head, would

1.4. Finite state grammar

Thus, the conditions on the selection of morphemes are local, in the sense that each choice point in growing a complex word is determined by the immediately prior stage of the derivation, with no reference to earlier stages. This is exactly how a finite state grammar works, and so there is a sense in which word formation is finite state.[3] Alternations between prefixation and suffixation may cause structural adjacency to skip nodes, but this is only relevant for the realization of allomorphs, which in a Late Insertion model like Distributed Morphology is separated from the actual construction of the word, or Extended Projection.

Chomsky (1957:24) pointed out that finite state grammars are inadequate for capturing natural language grammar. However, in the same place he comments that a finite state grammar "represents in a way the minimal linguistic theory that merits serious consideration. A finite state grammar is the simplest type of grammar which, with a finite amount of apparatus, can generate an infinite number of sentences."

Features triggering agreement and merge of specifiers will take the grammar of the clause beyond finite state, as is required, but on my proposal here (and in Svenonius 2012), the production of extended projections does not go beyond finite state.

Thus, I argue that the properties of EPs suggest a particular architecture for their generation. The FSA starts from a lexical category (cf. (5c)), 'projecting' it into an EP. Transitions correspond to applications of Merge, and states are stages in the derivation; at each stage, the root of the tree created up to that point has a category, so states can be given category labels. The sequence of states in a path from start to accepting state is a string, the EP (cf. (5b)). The HoP for a language is then the entire FSA, arranged with the start at the bottom and the final state, the complete clause or noun phrase, at the top (cf. (5a)).

2. Workspaces

Up to this point I have laid out some properties of EPs, including their linear structure, their adherence to a hierarchy, and their rootedness (cf. (5)), and I have pointed out how they are similar to words in all of these respects. Here I will suggest how these properties might follow from some plausible architectural assumptions concerning Workspaces. Svenonius (2016a), building on earlier ideas about Workspaces (such as Uriagereka 1999 and Lohndal 2014), suggests that each EP is built in a distinct Workspace. This suggestion can be formalized in the following way.

(8) *Rootedness Hypothesis*: A verbal EP Workspace is initialized by the activation of a V-domain formative (e.g., a verbal root); the EP which is built in that Workspace contains that formative, and all material which is Merged in the Workspace is Merged to the top of that EP

This partly restates Chomsky's (1993) Extension Condition, which requires that every instance of Merge expand the existing structure at the top; material introduced earlier becomes successively more embedded (the Extension Condition is empirically motivated by observations of cyclic effects). If a Workspace contains exactly one active node at any time, following the *Root-*

undermine the restrictiveness of the locality condition on morpheme selection if wantonly invoked. The fact that declension class, for example, is not visible in positions nonlocal to a stem suggests that Agree cannot copy declension class features. In the vast majority of clear cases, Agree copies only ϕ (person, number, gender) and/or case. Further investigation is necessary to understand the conditions under which Agree is extended to additional features.

[3] Cf. Koskenniemi (1983) and Roark and Sproat (2007), though they are mainly concerned with surface strings of exponents, while I am concerned with underlying hierarchies of heads in an EP. See Oseki and Marantz (2020) concerning why this matters.

edness hypothesis in (8), then a complex specifier or adjunct will have to be built in a distinct Workspace and imported.

Uriagereka (1999) suggests that for an expression to be imported, it must be spelled out. This derives a certain kind of islandhood for specifiers and adjuncts, since all specifiers and adjuncts will have to be imported in this way and therefore will undergo spellout prior to being embedded. However, this has broader ramifications. For example, if the case features of an argument are valued by a case licenser after A-movement, and the morphology of the argument reflects its case, then the morphological exponence of the argument cannot be fixed prior to importation into the theta position in the verbal EP. I will therefore modify Uriagereka's assumptions. Importation of the output of one EP Workspace into another does not require morphophonological spell out. It fixes syntactic structure but not the valuation of unvalued features nor their morphological exponence.

To distinguish the islandhood of specifiers and adjuncts from the non-islandhood of complements, (8) is intended to force specifiers and adjuncts, but not complements, to be constructed in distinct Workspaces. It is therefore distinct from phase theory, in which vP is a phase despite being a complement of T, and a typical adjective or adverb is not a phase.

The Linear property of EPs, (5b), appears to reflect an interpretive difference between heads and non-heads. The linear property of EPs is the fact that all heads in an EP are nonbranching; they are distinct from branching specifiers and adjuncts in that they require no preassembly before being introduced into the EP, so they need not be imported.

If EPs could be non-linear, then an aspectual collocation like *was going to* could be a head, with a structure like [[was going to] stay]. In fact, languages tend to parse such structures linearly, [PAST [be [going [to [stay]]]]], or [PAST [be [-ing [go- [to [stay]]]]]]. This could be made to follow if there is a basic interpretive difference among the different structural elements. For example, if specifiers are semantically arguments, and adjuncts are semantically conjuncts, and heads in an EP are semantically functors, then if the prospective aspectual semantics expressed by *be going to* is most naturally identified as a functor, a learner will be inclined to analyze it as a head, or a collection of heads.

Any specifier or adjunct which is an EP consisting of more than one formative must be imported from its own EP WS. It is an open question whether there are also simple specifiers or adjuncts which need not pass through an EP WS. In any case we can distinguish between heads in the WS numeration and specifiers and adjuncts which are imported into the WS, either from other EP WSs or from some store of non-EP objects. If complements cannot undergo internal merge, as suggested by Adger (2013), then anything which is internal merged would have been imported at an earlier stage of the derivation.

Assuming that auxiliaries and modals are included in the same EP as the main verb they support, such auxiliaries and modals would have to be analyzed as purely functional material, lacking a lexical root.

If all elements which are imported from other Workspaces are specifiers or adjuncts, and lexical roots are identified with Workspaces, that seems to leave no room for structural complements which contain lexical roots. However, certain kinds of complements, for instance bridge and restructuring complements, are typically not islands, which is problematic if importation from a distinct Workspace implies islandhood (as suggested by Uriagereka 1999, it will be recalled). Svenonius (2016a) argues that restructuring arguments are structurally complements (distinct from the typical direct object, which is a specifier).

A structural complement is a sister to a head, or a first-Merged dependent of a head, a relation which is correlated in the Workspace with external Merge of a head. We can distinguish two kinds of complement terminologically. EP-internal complements are lower than the head combining with them in the HoP; for example Voice over V, or a manner verb over a resultative particle. These are unproblematically constructed in a single EP. Embedded complements are different: embedded complements contain material which is as high or higher as the embedding head in the Hierarchy of Projections. One example of an embedded complement is an outer causative of an agentive verb, in which Voice embeds Voice, as in the Hiaki example in (9). Another example is a restructuring complement in which V embeds some functional material, as in the Cup'ik example in (10).

(9) Maria hitevi-ta uusi-ta hitto-tua.
Maria doctor-ACC child-ACC treat-CAUSE
'Maria is making the doctor treat the child' (Hiaki, Blanco 2011)

(10) pik-a-qa-ar-lu-ku
own-IND.3SG.OBJ-1SG.A-say-APPOS-3SG.OBJ
'Saying to him/her, "It's mine!"' (Cup'ik, Woodbury 2024:116)

In both examples, there is language-internal evidence showing that the complex verb is a single word. In Hiaki, the causative combines with a special, listed, bound form of the verb stem (Blanco 2011). In Cup'ik, phonological diagnostics like stress show that there is no prosodic boundary between the restructuring verb and its inflected clausal complement (Woodbury 2024).

It is a robust crosslinguistic characteristic of EPs that they manifest strict hierarchies; thus, examples like these show that a single word can contain more than one EP. On the other hand, there is something special about embedded EPs which are complements, as opposed to specifiers and adjuncts.

Embedded complements allow clause union effects like clitic climbing, clustering, and long passive (Wurmbrand 2001). These kinds of effects are not seen when adjunct or specifier EPs are imported.

On the other hand, embedded complements are distinguishable from single EPs by a morphological and lexical boundary. As noted, Hiaki causatives show selection for a root allomorph, but there is nevertheless a morpheme boundary between the verbal root and the causative. The allomorph selection is not lexical in the sense of allowing roots to select idiosyncratic listed allomorphs of the causative (as is common for inner causatives, where there is no Voice recursion). Nor does it allow fully idiosyncratic causative forms of the roots, something which again is common for inner causatives. Voice recursion is never fully lexicalized or idiomatic; it must be regularly derived. No listed verb idiosyncratically has two agents.

This applies quite generally to restructuring. If a verbal form embeds something which is as high or higher than it in the Hierarchy of Projections, then the morphological form must be regular.

I propose to accommodate these observations by allowing a Workspace to "Reset". At a certain point in the projection of a verbal EP, the derivation may cycle back to a lower entry point, namely a restructuring verb (such as an outer causative). The fact that the derivation is still in the same Workspace means that the derivation can still produce a single word including material from the first EP cycle as well as from the second. But the resetting requires exponent selection for the material up to the reset point.

(11) *Reset*: A restructuring head allows a loop back to a Workspace Cycle in the same Workspace, restarting the EP in the V-domain. The material from the first Workspace Cycle undergoes lexical insertion at the point of Reset.

A derivation which takes place in a single Workspace is linear and can spell out as a single word. Within a Workspace Cycle, the derivation is strictly monotonic, but the derivation can be 'reset' to start a second cycle in the same Workspace.

Thus the projection line of a sentence spanning multiple clauses can be built in a single Workspace, broken into separate Workspace Cycles. Each Workspace Cycle will have a distinct Workspace numeration. The Workspace Cycle has consequences for span-based lexical insertion which includes lexical roots as well as lower material, as in Son and Svenonius (2008) and Svenonius (2016b).

The other property of EPs is the property of Hierarchy, (5a). It does not follow immediately from the fact of there being a Workspace with a single Root node and a single lexical root in which a WS numeration of heads is linearly assembled that that assembly should be subject to a Hierarchy of Projections. Therefore, we must identify the source or sources of hierarchy, to which I turn in the next section.

3. Sources of hierarchy

In this section, I discuss a number of potential sources of hierarchy among syntactic categories in EPs. No single identified source of hierarchy appears to be capable of fully explaining the observed tendencies of languages to have hierarchically structured EPs. I conjecture, therefore, that a property of human language acquisition favors monotonically structured paths for EP construction.

If there is a functional sequence, in the sense of a universal hierarchy of functional categories which constrains the forms of the language-specific HoPs, which in turn constrain the construction of EPs in individual sentences, then what is the source of the functional sequence? In chemistry, atomic weight determines a total order on the elements. Could there be a similar property behind the functional sequence? Imagine a periodic table of functional categories, based on Cinque's (1999) TMA categories. For example, there are four categories in Cinque (1999) which situate a reference point subsequent to the situation described in the complement, namely T_{past}, $T_{Anterior}$, Asp_{term} (terminative), and Asp_{retro} (retrospective). There are also three categories which situate a reference point prior to the complement situation, namely T_{future}, $Asp_{proximate}$, and $Asp_{prospective}$, and possibly also a fourth, $Asp_{continuous}$. Three of the 'prior' heads are immediately dominated by three of the 'subsequent' heads, suggesting a table organized as follows.

(12)

$\prec$	$\succ$	$\forall$	$\exists$	Gen	Asp
T_{past}	T_{future}	Mod_{nec}	Mod_{poss}	Asp_{hab}	Asp_{repI}, Asp_{freqI}, Asp_{celI}
$T_{anterior}$					
Asp_{term}	Asp_{cont}	Asp_{perf}			
Asp_{retro}	Asp_{prox}	Asp_{dur}		Asp_{prog}	
	Asp_{prosp}	Mod_{oblig}	Mod_{perm}		
		Asp_{compl}			Asp_{celII}, Asp_{repII}, Asp_{freqII}

In the original periodic table of chemical elements, empty cells corresponded to predictions, and Nobel prizes could be won by isolating the predicted elements for the first time. In that light, one response to this pattern would be to start looking for more categories in the second row, including a prior-type head for the second column, and for a subsequent-type head for the fifth row of the first column.

However, it seems unlikely that many researchers will feel this call. Prospects for something like an atomic weight to explain the functional sequence seem dim.[4]

Culbertson et al. (2020) propose to explain the universality of the hierarchy of Demonstrative > Numeral > Adjective > Noun (Greenberg 1963, Cinque 2005) in terms of acquired conceptual associations based on real world properties. Their idea is that nominal concepts such as *dog* and *wine* tend more strongly to be associated with adjectival concepts like *red* and *spotted* than with the kinds of quantities referred to by numerals. For example *wine* is more likely to be associated

[4]There is a long history of suggestions about how HoP might follow from some property of features, for example Ackema et al. (1993), Koeneman (2000; 2010), Larson (2021), but none have achieved a level of predictiveness.

with *red* than *dog* is, and vice-versa for *spotted*. Some nominal concepts are strongly associated with some numbers, for example *eyes* with *two* or *fingers* with *ten*, but numerals as a whole are not as closely associated with individual nouns as adjectives are. In information-theoretic terms, an adjective is, on average, a better predictor of a noun than a numeral is. Demonstratives show an even less skewed distribution than numerals, and are an even less good predictor of nouns. So for the three classes of nominal modifier, adjective, numeral, and demonstrative, distance from the noun corresponds negatively with strength of conceptual association.

The Culbertson et al. scale works only at a coarse-grained level similar to the level of domains. It does not predict, for example, the order of adjective classes. Among adjectives, for example, size tends to precede color when both adjectives are prenominal: *big blue table* versus ?*blue big table*. But it is not obvious that things in the world are more closely associated with typical colors than they are with typical sizes, so it is not clear that the Culbertson et al. scale can be applied to adjective classes.

Scontras et al. (2017) argue that for classes of adjectives, distance from the noun corresponds to a measure of subjectivity. They argue that size descriptions are more subjective than color descriptions, explaining the relative order of *big* and *blue*. Scontras et al. use several different measures of subjectivity to test this, for example subjects' ratings of how likely it is that two people could disagree on whether an adjective applies without either of them being wrong.

Thus, it is likely that there are several different sources for hierarchy, as argued by Ramchand and Svenonius (2014). They propose that the coarse-grained hierarchy of domains, C > T > V, is due to a containment relation among three semantic objects, namely EVENTS (the atemporal domain of thematic relations and force dynamics, including causation) are contained in SITUATIONS (drawing on Barwise and Perry (1983), a situation has a time and may have a topic and other attributes which are not possible in an event), which are in turn contained in PROPOSITIONS (taking this term to describe something with an anchor to the context of the utterance, lacking from the situation). Ramchand and Svenonius (2014) argue that this containment is the reason for the hierarchy of clausal domains C > T > V.

Containment may be responsible for hierarchies at finer-grained level as well. Wurmbrand argued at the 2022 Tromsø Workshop on the Origins of Extended Projections for a containment relationship within propositions, where a commitment to the content of a proposition necessarily implies a judgment about the proposition, with the result that a Commitment projection dominates a Judgment projection (and strong Judgment dominates weak Judgment for the same reason).

Another example of hierarchy by containment is the hierarchy of first and second person (participant) over third person, if third person is the absence of participant features (Bobaljik 2008). This hierarchy manifests itself among other ways in the fact that omnivorous person agreement always prefers participants to nonparticipants (Preminger 2014). This can be captured if first and second person features are added to something which is interpreted as third person in their absence; for example, a person probe might probe for D, and participant either dominates D (so that the marked category contains the unmarked category as a subconstituent) or participant is a feature of D (assuming that categories can have features; more on this below).

In other cases, however, containment does not seem to correlate with hierarchy. Consider, for example, the hierarchy of nominative over accusative. In agreement systems, nominative is always preferred over accusative, in the sense that if a single argument is agreed with, it is the nominative, not the accusative. This is similar to the generalization about person. However, in featural analyses of case, the nominative is always the least marked, and in a structural breakdown of case like that proposed by Caha (2009), accusative properly contains nominative. Again, the marked contains the unmarked, but this time it is the unmarked case which appears to be 'higher' and occupies a higher position than the marked case.

Thus it seems that there is currently no explanation for (5a), the fact that there are HoPs which govern the constructions of EPs. I offer the admittedly speculative proposal that the development of EP WSs as part of human cognitive development and language acquisition favors monotonicity. In other words, when a property licensing category combination is at issue, the language learner is more likely to build that combination into the grammar if it leads to a monotonic EP WS, and less likely if it leads to a non-monotonic EP WS. Monotonicity in this case means that the categories are arranged in a hierarchy, so that every step in the projection of an EP is "upward" (cf. Graf 2019).

4. Finite state generation of extended projections

The derivation of a sentence involves several different operations, some of which are structure-building and others of which are not. Agree copies feature values, it does not build structure of the kind that can be detected by constituency tests, but Merge does. Structure-building may introduce specifiers or adjuncts. These are constructed in distinct workspaces prior to being introduced as specifiers or adjuncts, and at a given stage of a derivation where a specifier is to be introduced, any of an infinite number of possible specifiers could be; the same is true of adjuncts.

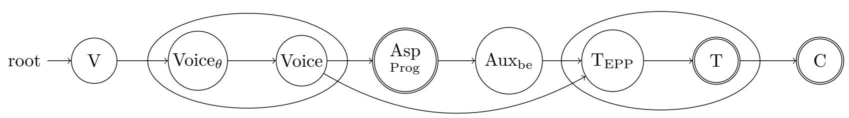
Structure-building which extends the EP, on the other hand, consists in adding a functional head to an already preexisting structure, and is highly constrained. In general only a small number of options are possible at any given stage of extending an EP, often only one.

Consider the derivation of *The dogs are barking*, taking it to be a CP with a null declarative C. The DP *the dogs* is constructed in one workspace. In another workspace, the extended projection of the verb is built. The first stage of structure building will combine the verb V, or its root, with the next lowest head in the EP. Suppose for the sake of argument that that lowest head is (active) Voice, which introduces the external argument. An agentive verb like *bark* always has Voice, and there are never any heads between V and Voice, so the projection from V to Voice is deterministic.

Voice introduces the external argument, so the second structure-building stage of the derivation in this workspace is the introduction of *the dogs* from the other workspace. In the third structure-building step in this workspace, Asp_{Prog} is introduced, and will eventually be pronounced as an *ing* suffix on the verb. An alternative derivation passing through Voice could skip progressive aspect to move on to simple tense, deriving *The dogs bark.* Though there are multiple options at some points in the derivation, they are still highly constrained.

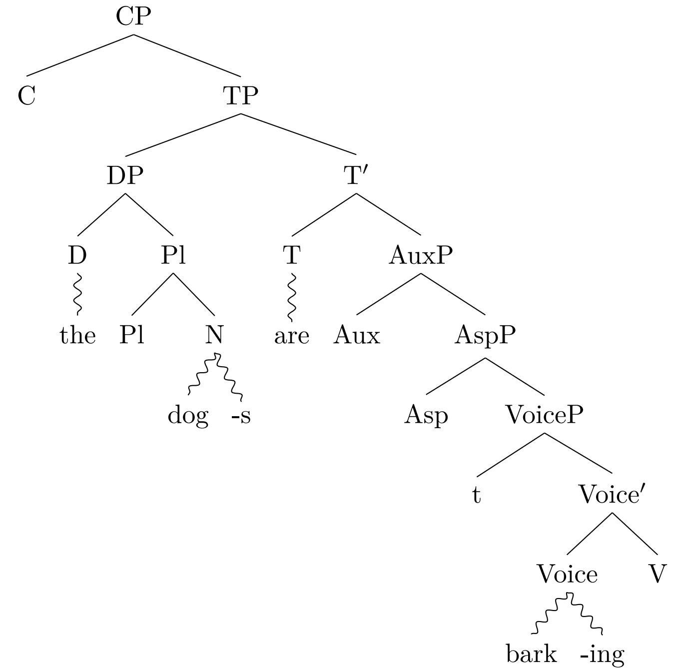
We can represent each structure-building stage of the derivation in the workspace of a single extended projection as a state in a finite state automaton.

(13)



Each state corresponds to a structural position in the "spine" of the syntactic tree.

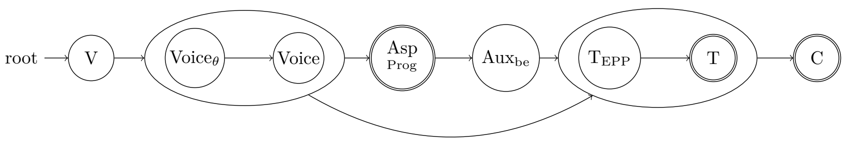
(14)



In (13), I mark the start as "root" to accommodate the possibility that V is built from a root; in any case V combines with Voice, and Voice has a property which requires it to merge with an external argument, signified here by the subscripted θ. The transition from Voice_θ to Voice represents the merge of the specifier, which is a distinct extended projection and is built in a separate workspace; the diagram in (13) shows only the steps in the derivation of an extended projection of V.

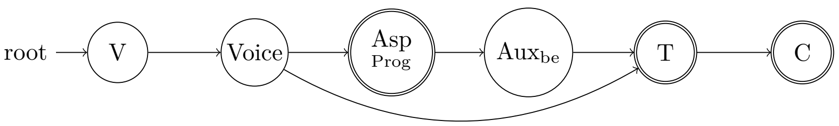
Voice_θ and Voice are a single category, as indicated by the ellipse drawn around both of them. The states not contained in ellipses are heads by themselves. To better distinguish between steps in the derivation that introduce specifiers from steps in the derivation that introduce heads, we could treat each category as a state, as in the following diagram, which now relegates the merge of specifiers to a procedure happening within a stage of EP growth by category change.

(15)



A further simplification is to allow the specifier-merging stages to remain implicit, as in the following diagram. Notice that T is an accepting state, but implicitly only after the EPP has been satisfied.

(16)



The transition which goes directly from Voice to T represents an alternative derivation, giving *The dogs bark*, i.e., without the progressive aspect.

Once progressive Asp is merged, the derivation in this workspace could in principle terminate, and the AspP consisting of *the dogs barking* could be introduced as a specifier or adjunct in another workspace (as in (17a) or (17b)) or as the thematic complement of a lexical head, as in (17c).

(17) a. The dogs barking is the least of our worries.

b. The thieves ran off, the alarms wailing and the dogs barking.

c. I heard the dogs barking.

For this reason, I have marked Asp with a double circle, indicating a possible final state in this FSA.

If the derivation is not terminated, then the only possible move is to pick up the auxiliary *be*. The auxiliary requires T, so the only transition from *be* is to T. T has an EPP feature which requires another round of Merge, as seen in (13), but collapsed in (16), where the state T corresponds to the whole category T, including TP and T′.

T is an accepting state assuming that TP can be embedded, and CP is an accepting state.

There are five converging paths from start to an accepting state: one through all the states, ending at C; one through all the states but C, ending at T (for embedded TP); one which goes to C but skips Asp and Aux, and one that goes to T but skips Asp and Aux; and finally, one ending at Asp (for embedding AspP).

Thus this FSA produces five different EPs, with a specifier in VoiceP, and one in TP if they contain TP (in English, these two specifiers will be the same DP). In each EP, the path through the FSA matches the "spine" of the syntactic tree.

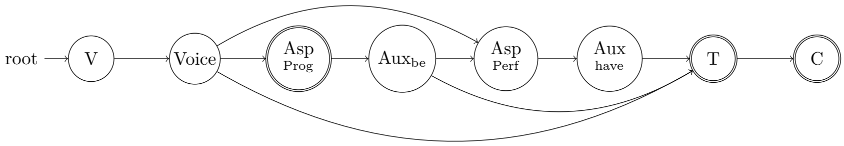
T agrees with the subject, and the noun phrase also receives case. One could add states to the diagram for non-structure-building operations, but there is reason to believe that they might have a different status, related to the difference between features and categories. As it stands, I can assume that every state in a diagram like (16) is structural in the traditional sense, for example it is a possible adjunction site. If states are added for every feature, then it is no longer clear that all states are structurally independent.

Adjunction is not triggered by features, but is freely introduced at any category level. It is subject to semantic interpretation; adjunction is interpreted as conjunction, and requires something like semantic type matching. If identical categories are combined, the head is indeterminate, and the structure is interpreted as coordination, often requiring a coordinate marker to be introduced. Therefore, adjuncts can never be the same category as the phrase to which they are adjoined, but must be the same semantic type.

Given these assumptions suppose that the lexicon of English contains a lexical entry for the word *often* which is category Adv_{Asp}, so-called because it is of the same semantic type as Asp, but has a different category (defined in terms of position in an HoP). This allows *often* to adjoin to Asp freely. In a derivation where it does, there is an additional step of structure-building compared to a derivation in which it doesn't. The FSA cannot represent these additional free steps of adjunction, because the FSA represents the sum total of all EP building. It is like the mental representation of a paradigm, and the adverbs do not belong to the paradigm.

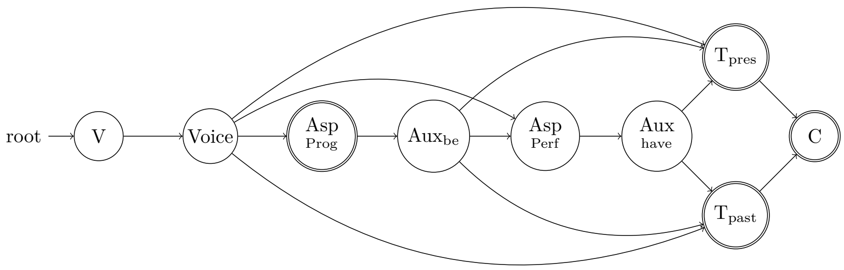
Suppose that a language happens not to have any modifiers that can attach to a given category, for example the auxiliary *be*, which combines with T in examples like *the dogs are barking*, but which has semantic content. This eliminates a test for structure, but does not mean that there is no structure there. Constituency tests will also fail, if the category is not an accepting state.

The perfect in English is like the progressive in consisting of two morphological parts, an auxiliary and an inflection, which are optional but codependent. Adding the perfect to the FSA for English gives the following.

(18) 

The perfect can be combined with the progressive, or used without it, and it is optional, like the progressive. Unlike the progressive and passive participles, the perfect participle is not easily embedded in other EPs, hence it is not an accepting state.

Now consider the past tense. Like the present tense, it can immediately dominate Voice, Aux_{be}, or Aux_{have}, and also like the present tense, it can be immediately dominated by C. We can add it to the FSA as follows.

(19) 

Since T_{past} and T_{pres} have transitions to and from all the same states, they are distributionally identical. If we define category by position in an FSA, advancing the tradition of Bloomfield's defining category by distribution, then T_{past} and T_{pres} are by definition the same category. If the FSA represents categories and not features, we can collapse T_{pres} and T_{past} to a single state, as was already depicted in (18).

This brief exhibition will serve to give a sense of the EP WS as a mental construct that is partly independent of the inventory of syntactic formatives. Syntactic formatives like T_{pres} and T_{past} are category T, because they have the same distribution. If additional formatives are added to the category, for example modals, they inherit the properties which have been established for the category.

The monotonicity property which I have suggested underpins the HoP is a property of EP WSs. If the EP WS is established as a generalization independent of the formatives which match its categories, then the properties of the grammar which it expresses need not be independently coded in the individual formatives. This breaks slightly with the popular Minimalist conception of cross-linguistic variation being located entirely in the properties of lexical items (or of formatives and exponents, in a Distributed Morphology).

5. Conclusion

I have addressed three properties of Extended Projections, namely Hierarchy, Linearity, and Rootedness, and have suggested that they are central to the architecture of the Workspaces that are responsible for the construction of Extended Projections.

I conceive of the EP Workspace as a Finite State Automaton. It starts at a verbal root, and follows transitions through a series of states to arrive at a clause. I have suggested that a tendency in the language learner to favor the property of monotonicity in the EP Workspace might go a long way to explaining the properties of Hierarchy, Linearity, and Rootedness.

Monotonicity is most closely related to Hierarchy, since a non-monotonic FSA would not generate a Hierarchical EP. I stipulated a source for Rootedness. An FSA that could start anywhere could still be consistent with monotonicity, if it could only proceed in one direction.

Linearity comes partly from the binary nature of Merge and the the asymmetry of headedness (or labeling), but since an EP is a collection of heads, Merge does not by itself explain why there are no complex heads. I suggested that the kinds of meanings that expand EPs tend to be added to the EP WS as categories, while specifiers and adjuncts express different kinds of meanings.

I have also introduced a notion of EP WS reset, which allows restructuring to show some properties of single clauses (for example morphological integration of predicates, long passive, clitic climbing) and some properties of biclausal structures (for example multiple argument structures, multiple case licensing possibilities, obligatory morphological independence of the two predicates).

References

Aboh, Enoch Oladé. 2004. *The Morphosyntax of Complement-Head Sequences: Clause Structure and Word Order Patterns in Kwa.* Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780195159905.001.0001.

Ackema, Peter, Ad Neeleman, and Fred Weerman. 1993. Deriving functional projections. In *North East Linguistics Society*, vol. 23, pp. art. 3 (1–15).

Adger, David. 2013. *A Syntax of Substance.* MIT Press, Cambridge, Ma. https://doi.org/10.1017/s0022226714000073.

Barwise, Jon and John Perry. 1983. *Situations and Attitudes.* MIT press, Cambridge, Ma.

Blanco, Mercedes Tubino. 2011. *Causatives in Minimalism.* John Benjamins, Amsterdam. https://doi.org/10.1075/la.179.

Bobaljik, Jonathan David. 2008. Missing persons: A case study in morphological universals. *The Linguistic Review* 25 1-2: 203–230.

Caha, Pavel. 2009. *The Nanosyntax of Case.* Ph.D. thesis, University of Tromsø.

Carden, Guy. 1983. The non-finite=state-ness of the word formation component. *Linguistic Inquiry* 14 3: 537–541.

Chomsky, Noam. 1957. *Syntactic Structures.* Mouton, 's-Gravenhage.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyser, pp. 1–52. MIT Press, Cambridge, Ma.

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

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36 3: 315–332. https://doi.org/10.1162/0024389054396917.

Culbertson, Jennifer, Marieke Schouwstra, and Simon Kirby. 2020. From the world to word order: Deriving biases in noun phrase order from statistical properties of the world. *Language* 96 3: 696–717. https://doi.org/https://doi.org/10.1353/lan.2020.0045.

Dik, Simon C. 1989. *The Theory of Functional Grammar.* Foris, Dordrecht.

Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology.* MIT Press, Cambridge, Ma. https://doi.org/10.7551/mitpress/9780262014229.001.0001.

Fábregas, Antonio. 2013. Diminutives as heads or specifiers: The mapping between syntax and phonology. *IBERIA: An International Journal of Theoretical Linguistics* 5 1: 1–44.

Graf, Thomas. 2019. Monotonicity as an effective theory of morphosyntactic variation. *Journal of Language Modelling* 7 2: 3–47. https://doi.org/10.15398/jlm.v7i2.211.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language. Report of a Conference Held at Dobbs Ferry, NY, April 13-15, 1961*, edited by Joseph H. Greenberg, pp. 73–113. MIT Press, Cambridge, Ma.

Grimshaw, Jane. 1991. Extended projections. Ms. Brandeis University.

Grimshaw, Jane. 2005. Extended projection. In *Words and Structure*, edited by Jane Grimshaw, pp. 1–73. CSLI, Stanford, Ca. Revised version of 1991 ms.

Gruber, Jeffrey. 1997. Modularity in a configurational theta theory. In *Projections and Interface Conditions: Essays on Modularity*, edited by Anna Maria Di Sciullo, pp. 155–200. Oxford University Press, Oxford.

Hankamer, Jorge. 1989. Morphological parsing and the lexicon. In *Lexical Representation and Process*, edited by William Marslen-Wilson, pp. 392–408. MIT Press, Cambridge, Ma.

Hengeveld, Kees. 1990. The hierarchical structure of utterances. In *Layers and Levels of Representation in Language Theory*, edited by Jan Nuyts, A. Machtelt Bolkestein, , and Co Vet, pp. 1–24. John Benjamins.

Holmberg, Anders, Urpo Nikanne, Irmeli Oraviita, Hannu Reime, and Trond Trosterud. 1993. The structure of INFL and the finite clause in Finnish. In *Case and Other Functional Categories in Finnish Syntax*, edited by Anders Holmberg and Urpo Nikanne, Studies in Generative Grammar, pp. 177–206. Mouton de Gruyter, Berlin.

Koeneman, Olaf. 2000. *The Flexible Nature of Verb Movement*. Ph.D. thesis, Universiteit Utrecht.

Koeneman, Olaf. 2010. Verb movement in Germanic and Celtic languages: A flexible approach. *Lingua* 120 2: 210–231. https://doi.org/10.1016/j.lingua.2008.05.006.

Koskenniemi, Kimmo. 1983. *Two-level Morphology: A general computational model for word-form recognition and production*. Ph.D. thesis, University of Helsinki, Helsinki.

Larson, Richard K. 2021. Rethinking Cartography. *Language* 97 2: 245–268. https://doi.org/10.1353/lan.2021.0018.

Lohndal, Terje. 2014. *Phrase Structure and Argument Structure: A Case Study of the Syntax-Semantics Interface*. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780199677115.001.0001.

Oseki, Yohei and Alec Marantz. 2020. Modeling human morphological competence. *Frontiers in Psychology* 11: 513740. https://doi.org/https://doi.org/10.3389/fpsyg.2020.513740.

Preminger, Omer. 2014. *Agreement and its Failures*. MIT Press, Cambridge, Ma. https://doi.org/10.7551/mitpress/9780262027403.001.0001.

Ramchand, Gillian and Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46 B: 152–174. https://doi.org/10.1016/j.langsci.2014.06.013.

Rijkhoff, J. 2002. *The Noun Phrase*. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780198237822.001.0001.

Roark, Brian and Richard Sproat. 2007. *Computational Approaches to Morphology and Syntax*. Oxford University Press, Oxford. https://doi.org/10.1093/oso/9780199274772.001.0001.

Scontras, Gregory, Judith Degen, and Noah D Goodman. 2017. Subjectivity predicts adjective ordering preferences. *Open Mind* 1 1: 53–66. https://doi.org/10.1162/opmi_a_00005.

Son, Minjeong and Peter Svenonius. 2008. Microparameters of cross-linguistic variation: Directed motion and resultatives. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, edited by Natasha Abner and Jason Bishop, pp. 388–396. Cascadilla, Somerville, Ma.

Svenonius, Peter. 2012. Merge, project, and bundle. Ms. CASTL, University of Tromsø.

Svenonius, Peter. 2016a. Complex predicates as complementation structures. In *Approaches to Complex Predicates*, edited by Léa Nash and Pollet Samvelian, pp. 212–247. Koninklijke Brill, Leiden. https://doi.org/10.1163/9789004307094_009.

Svenonius, Peter. 2016b. Spans and words. In *Morphological Metatheory*, edited by Heidi Harley and Daniel Siddiqi, pp. 199–220. John Benjamins, Amsterdam. https://doi.org/10.1075/la.229.07sve.

Uriagereka, Juan. 1999. Multiple spell-out. In *Working Minimalism*, edited by Samuel David Epstein and Norbert Hornstein, pp. 251–282. MIT Press, Cambridge, Ma.

Wiltschko, Martina. 2014. *The Universal Structure of Categories: Towards a Formal Typology*. Cambridge University Press, Cambridge. https://doi.org/10.1017/cbo9781139833899.

Woodbury, Anthony C. 2024. Constituency in Cup'ik and the problem of holophrasis. In *Constituency and Convergence in the Americas*, edited by Adam J. R. Tallman, Sandra Auderset, and Hiroto Uchihara, pp. 83–136. Language Science Press, Berlin. https://doi.org/10.5281/zenodo.13208542.

Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and Clause Structure*. Mouton de Gruyter, Berlin. https://doi.org/10.1515/9783110908329.