


Article

Lexicalised Locality: Local Domains and Non-Local Dependencies in a Lexicalised Tree Adjoining Grammar

Diego Gabriel Krivochen ^{1,*} and Andrea Padovan ² ¹ Dipartimento di Culture e Civiltà, Università di Verona, 37129 Verona, Italy² Dipartimento di Lingue e Letterature Straniere, Università di Verona, 37129 Verona, Italy; andrea.padovan@univr.it

* Correspondence: diegokrivochen@hotmail.com

Abstract: Contemporary generative grammar assumes that syntactic structure is best described in terms of sets, and that locality conditions, as well as cross-linguistic variation, is determined at the level of designated functional heads. Syntactic operations (merge, MERGE, etc.) build a structure by deriving sets from lexical atoms and recursively (and monotonically) yielding sets of sets. Additional restrictions over the format of structural descriptions limit the number of elements involved in each operation to two at each derivational step, a head and a non-head. In this paper, we will explore an alternative direction for minimalist inquiry based on previous work, e.g., Frank (2002, 2006), albeit under novel assumptions. We propose a view of syntactic structure as a specification of relations in graphs, which correspond to the extended projection of lexical heads; these are *elementary trees* in Tree Adjoining Grammars. We present empirical motivation for a lexicalised approach to structure building, where the units of the grammar are elementary trees. Our proposal will be based on cross-linguistic evidence; we will consider the structure of elementary trees in Spanish, English and German. We will also explore the consequences of assuming that nodes in elementary trees are addresses for purposes of tree composition operations, substitution and adjunction.



Citation: Krivochen, D.G.; Padovan, A. Lexicalised Locality: Local Domains and Non-Local Dependencies in a Lexicalised Tree Adjoining Grammar. *Philosophies* **2021**, *6*, 70. <https://doi.org/10.3390/philosophies6030070>

Academic Editor: Peter Kosta

Received: 15 July 2021

Accepted: 4 August 2021

Published: 18 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: locality; tree adjoining grammar; lexicalised grammar; tree composition

1. Introduction

The definition of local domains for the application of syntactic rules, as well as rules of semantic interpretation, has been a central topic in generative grammar since the first definitions of a transformational cycle [1–3]. In the early days of generative grammar, cycles were delimited by the presence of two designated nodes, S and NP, as the only categories that could take subjects. From a contemporary perspective, this definition implies a certain heterogeneity. Whereas NP is the projection of a lexical category, S does not fit in the uniformly endocentric template that has been the norm since the mid-1980s, and thus is neither a “lexical” nor a “functional” category in a strict sense. The mixed nature of bounding nodes remained in the Government and Binding (GB) framework, where barriers were nodes defining opaque domains for the application of syntactic operations [4]. For example, CP was a barrier for movement from VP (not L-marked, dominated the bounding category inflexion), as were VPs (not L-marked) and adjuncts (similarly, not L-marked). Note that C is a functional head, but V is not (adjuncts may be PPs or AdvPs, also lexical categories). The definition of locality and the distinction between lexical and functional categories remained relatively orthogonal to each other.

In the late 1980s functional categories started playing a much more crucial role in the syntactic derivation (in particular after the layered analysis of the VP in [5] and the “explosion” of the node Infl(ection) in [6]), and the research on locality became centred on functional heads and their projections. With the advent of the phase-theoretic framework, in [7] and much subsequent work, this shift towards functional categories as delimiters of local structure became even more obvious. Syntactic domains are delimited by the

presence of functional categories v^* and C , which defines the size of syntactic chunks where operations of feature checking/valuation and movement/internal merge take place. However, the consequences of this shift are far-reaching. The focus on functional categories as the skeleton of the clause (and the main locus of cross-linguistic variation) has found its maximal expression in the cartographic enterprise as well as nanosyntax (where features configure a finer-grained hierarchy at a smaller level of organisation). Because the order of projections in the phrasal skeleton is fixed (grossly speaking, the phrasal skeleton is an ordered array $C > T > V$, where C , T and V each stands for a possibly highly articulated sequence of functional heads, as in [8–11]), the occurrence of phase heads and non-phase heads follows a rhythm determined *a priori* [12–14] (p. 56).

In this paper, we will argue for a different approach to the locality in syntax, which follows the lead of lexicalised tree adjoining grammars (TAGs). This entails two departures from contemporary minimalist practice, which will be empirically motivated. First, that the grammar builds and operates over graphs, not sets. Second, that locality conditions defined in structural representations are intimately related to the distinction between lexical and functional material, such that lexical heads define local interpretative domains. We will provide arguments in favour of a model of the grammar where the size of syntactic domains is neither defined *a priori* nor is it determined by the presence of a specially designated functional head. Rather, cycles are defined as the *extended projections of lexical heads*. In order to do so, we need to present some preliminary definitions of lexicalised TAGs.

2. Main Tenets of the TAG Formalism

A TAG is essentially a rewriting system working on predefined *elementary trees* that can be augmented and combined with one another at the frontier or “counter-cyclically”; these two cases correspond to the (generalised) operations *substitution* [15,16] and *adjunction* [17,18], respectively. Trees may be either *elementary* or *derived*: the latter are obtained by means of applying composition operations to the former. In turn, two kinds of elementary trees are defined: *initial trees* and *auxiliary trees*. Initial trees are single-rooted structures that contain a non-terminal node that is identical to the root of an auxiliary tree. Auxiliary trees are also single-rooted structures, which contain a node in their frontier that is identical to their root: this allows for auxiliary trees to be adjoined to initial trees at the non-terminal that corresponds to the root of the auxiliary tree.

The root node of an initial tree $IT \gamma$ is labelled “S”, and its “frontier” is constituted by terminal nodes. Intermediate nodes (transitively dominated by S) are non-terminals, as seen in Figure 1.

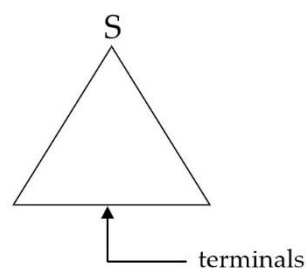


Figure 1. Initial tree.

The root node of an auxiliary tree $AT \beta$ is labelled “X”, and its frontier contains at least a node with the same label as the root [19] (p. 18). Intermediate nodes are non-terminals. This is diagrammed in Figure 2.

The crucial aspect of TAGs is the possibility of constructing *derived trees* from these elementary units under well-defined conditions. Let us define and illustrate the two operations that give us tree composition in a TAG.

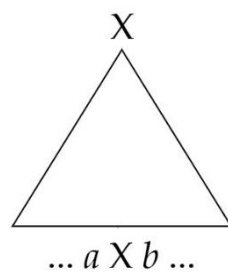


Figure 2. Auxiliary tree.

- Substitution:** let T be a tree with root S and T' a tree with root S' and a node labelled S in its frontier. Substitution inserts T into the frontier of T' at S . In other words, this operation can be conceived of as the rewriting of a node at the frontier of a piece of structure as another piece of structure. In this sense, substitution is a tail-recursive procedure. Substitution is diagrammed in Figure 3.

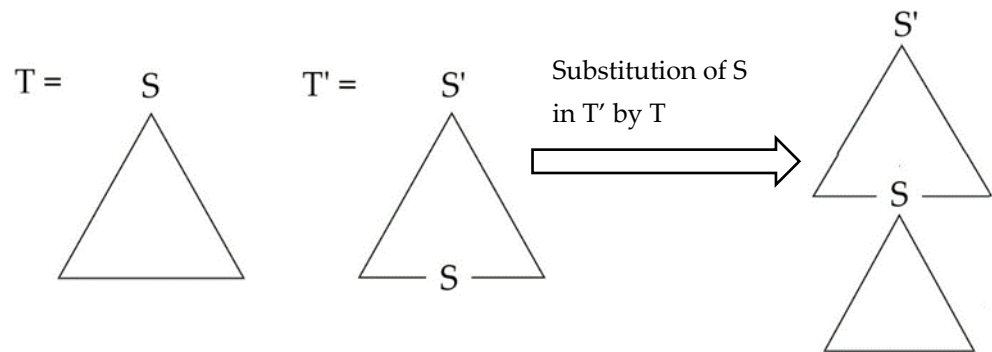


Figure 3. Substitution.

- Adjoining:** in this case, a piece of structure, the *auxiliary tree*, is inserted into a designated node (the *adjoining site*) in the *initial tree*. In other words, the operation of adjoining rewrites a node by “shoehorning” a piece of structure (i.e., the auxiliary tree) inside another structure (i.e., the initial tree). The auxiliary tree must have two distinguished nodes at its frontier (node Z in (Figure 4)), namely the *root node* and the *foot node*, which must carry the same label as each other, and which also must be the same label as the target for adjunction in the initial tree (here, that label is Z). Note that the foot node can be the daughter of the root node (as in the auxiliary tree (Figure 4a)) or show up much lower in the structure (as in the auxiliary tree (Figure 4b)).

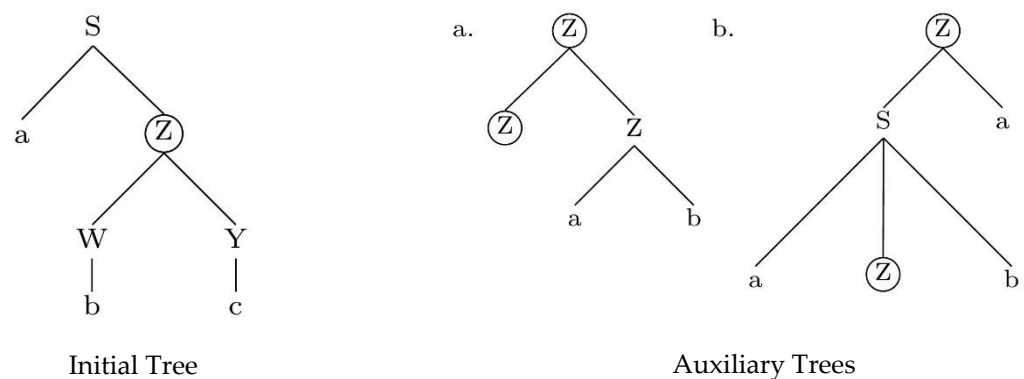


Figure 4. Initial and auxiliary trees. Auxiliary tree (a) features a recursive foot node immediately dominated by the root; auxiliary tree (b) features a foot node not immediately dominated by the root.

Both (Figure 4a,b) are admissible auxiliary trees (root and foot nodes are circled as in the original notation in [18] (p. 209)). Adjunction would insert an auxiliary tree (either

(Figure 4a) or (Figure 4b)) into the node Z in the initial tree, effectively pushing W and Y downwards. Suppose that we adjoin the auxiliary tree (Figure 4b) to the initial tree above. The resulting tree, referred to as a derived tree, would be (Figure 5).

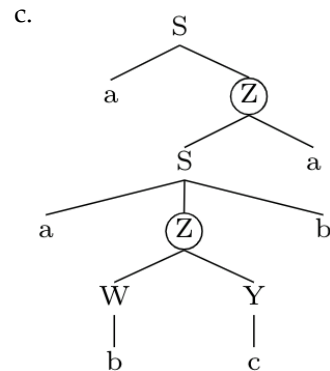


Figure 5. Derived tree.

A grammar featuring adjunction, as defined in [18] and related work, pushes the generative power of the grammar to mildly context-sensitive; it is worth pointing out that recent developments in so-called minimalist grammars [20] converge in mildly-context sensitive generative power.

Crucially, in both cases (substitution and adjunction), we are asking of the grammar that it be capable of identifying nodes that carry the same label. In other words, the grammar must be sensitive to nodes assigned to the same indexed category, even in distinct local structural chunks. The consequences of requiring that the grammar be capable of identifying nodes that are in some sense (to be specified below) identical even in distinct structural contexts have been, in our opinion, underexplored. The second part of our paper will deal with this issue in detail.

Conditions Imposed on TAG Derivations

The main question we want to address in this paper is how to define the size of elementary trees: what does an elementary tree (minimally) contain? Note that this question, as we posed it, makes no assumption that the size of elementary trees is uniform: as we will insist, a lexicalised TAG gives us the flexibility to define elementary trees of different sizes. Further, if elementary trees are defined in a lexicalised grammar, then it is entirely possible that different natural languages define trees of different sizes, depending on the properties of their lexical categories. This picture of cross-linguistic variation in terms of the size of elementary trees can be found in [21].

After introducing the building blocks of TAG, we can now expound on what has come to be known as the “fundamental TAG hypothesis” and the condition on elementary tree minimality (CTEM); a corollary on locality is also provided.

- **Fundamental TAG hypothesis:** Every syntactic dependency is expressed locally within a single elementary tree [19,21,22] (p. 233).
- **Condition on Elementary Tree Minimality (CTEM):** the heads in an elementary tree must form part of the extended projection (in the sense of [23]) of a lexical head. As a result, the lexical array underlying an elementary tree must include a lexical head and may include any number of functional heads that are associated with that lexical head, see [19] (p. 151).

In other words, elementary trees contain the nodes that form the extended projection of a single lexical head: this is what makes a TAG “lexicalised”. The size of elementary trees can be limited under an appropriately restricted formulation of the CTEM: what exactly does it mean for a head to be “lexical”, and what does its extended projection minimally contain? This will be one of the main issues to be explored in this paper.

- **Non-local Dependency Corollary:** Non-local dependencies always reduce to local ones once the recursive structure is factored out [21] (p. 237).

What the CETM states is that a particular lexical head H may entail a functional shell made of the (functional) projections that are assumed to surround it [21] (p. 239). For instance, V might include even T and C as they belong to the extended domain of V , i.e., the domain where selection, agreement, feature inheritance/percolation occur (in this sense, it may be that under certain formulations of the CETM, the extended projection of a V could coincide with its *s-projection*, in the sense of [24]). At first glance, the fact that a V is endowed with all that functional material might appear counterintuitive, especially from a phase-theoretic perspective, which crucially assumes that v^* and C are the two (main) phases. However, the idea of the extended projections of a head H belonging in the elementary tree of H allows us to have a natural definition of local domains that do not depend on designated nodes and is a direct consequence of CETM as formulated by Frank.¹ As long as an elementary tree does not contain more than one lexical head, it is a perfectly well-formed structure from a TAG perspective; this makes the connection between lexicality/functionality and locality much more natural than in minimalism. What becomes crucial in an LTAG is to have a proper definition of the lexical category, which is independently required in the grammar. After all, in a theory where functional material is the centre of attention, an intentional or extensional definition of what exactly is the functional lexicon of a language also defines, as the complement set, the “lexical” lexicon. In the present perspective, two notions become essential; lexicality and, as we will see below, predication.

Note that, under the CETM as presented above, V 's nominal arguments should define independent elementary trees since N is a lexical head (with D , Num , Deg and other functional heads being part of its extended projection). We will see that it is possible to modify the CETM in such a way that it becomes more restrictive with respect to the set of categories that can head elementary trees; this modification, crucially, is substantive, not formal. The generative power of grammar remains unchanged. In our opinion, the choice between a more or a less restrictive version of the CETM is an empirical issue and may depend on language-specific properties.

In order to have an empirical bite, the **Fundamental TAG hypothesis** and the **Non-local dependency corollary** require characterisation of the structural domain that comprises an elementary tree. Just as the context-free formalism tells us nothing about the kind of phrase-structure rules that should be present in grammar, the TAG formalism says little about the nature of elementary trees. In the words of [25] (pp. 2–3):

“TAG is a formal system that combines elementary trees, but it is not a linguistic theory and imposes no constraints upon the nature of [. . .] elementary trees”.

Consequently, from the perspective of formalism, the grammar of a language could simply consist of a list of an arbitrary finite set of (finite) trees. Such an approach would, however, fail to account for the generalisations that hold across the structures for a particular language, or across languages in general, unless independently motivated constraints are imposed on the nature of elementary trees: what linguistically relevant categories they can and cannot contain. Just as the theory of generalised phrase structure grammar (GPSG) [26] aimed to characterise the set of phrase-structure rules in the grammar of a language, a TAG-based theory of syntax must provide an independent characterisation of the set of elementary trees. Since the adoption of the TAG formalism commits us to the invariant adjoining and substitution operations, the definition of elementary trees in the grammar of a language is the locus of cross-linguistic differences, imposing substantial restrictions on the kind of variation that we may find. This is a particularly relevant point since different versions of TAGs may impose different restrictions. In a lexicalised TAG (that is, a TAG that adheres to principle (b) above), this entails that a source of variation may be precisely the definition of what constitutes a lexical category. In other words, the CETM gives us a framework in which to formulate a theory of configurational variation. The case for a

lexicalised grammar that we make in this paper is illustrated with the analysis of empirical data pertaining to Spanish auxiliary chains, which is contrasted with what an empirically adequate grammar for English auxiliary chains. LTAGs offer an interesting direction for minimalist inquiry in terms of grammatical analysis and variation.

In this paper, we examine two issues related to the TAG principles above:

1. What is the format of elementary trees? (Or, in other words, how many nodes does an extended projection need, and what are those nodes?).
2. How do non-local dependencies reduce to local ones? What exactly is a “local” dependency?

We will address these issues in the following sections.

3. What Is Inside an Elementary Tree?

The first question we need to address is what elementary trees may contain. The TAG formalism does not impose formal a priori conditions on the content of elementary trees, which means that the substantive definition of elementary trees becomes an eminently empirical matter. We may begin by laying out a meta-theoretical requirement: we aim at minimising the number of non-terminal nodes in elementary trees. We adhere to a principle of economy of expression [27]; it can be seen as a form of syntactic pruning. This means that, to the extent that it yields a descriptively adequate grammar, we want to minimise the number of projections, in particular, empty nodes and a priori functional material. This is consistent with the basic goals of lexicalised grammar but may be interpreted in different ways. One way of lexicalising a TAG is to make all lexical heads (predicates and arguments) have an extended projection [19,28,29] but not functional heads. In this case, a TAG is lexicalised if every elementary structure is associated with a single lexical item. Another is to have functional heads, as well as lexical heads, head elementary trees [28,30,31]. Finally, it is also possible to restrict the class of lexical heads that can project an elementary tree: not all lexical heads, but only lexical predicates [32,33]. In the latter case, an elementary tree is the smallest structure that contains:

- i. A predicative basic expression *p*.
- ii. Temporal and aspectual modifiers of *p* (e.g., functional auxiliaries, in [34]).
- iii. Nominal arguments of *p* (subjects, objects).

In this view, it is not enough for a head to belong to the set of commonly assumed lexical categories (N, V, A, Adv, P). In order to head an elementary tree (in TAG parlance, be the “anchor” of that tree), that head must also be a predicate: it must, in GB terms, have a thematic grid and a subcategorisation frame. Note that the elementary trees predicted by each approach are different: while a strict version of the CETM requires nominal arguments to head their own elementary trees if the focus is set on predication rather than just lexically, the elementary trees defined by the grammar look quite differently; we will compare these possibilities below. A theory of this kind is prefigured in [21], where it is claimed that “an elementary tree should be built around a single semantic predicate”; we merely implement the condition that the anchor of an elementary tree is only a semantic predicate. We can consider a very simple example, such as the sentence “John has run”, to illustrate each approach (the specific labels used here have no particular theoretical significance).

First, let us consider a proposal such as that in [30,31] ([21]; see also [29], p. 5), as in Figure 6.

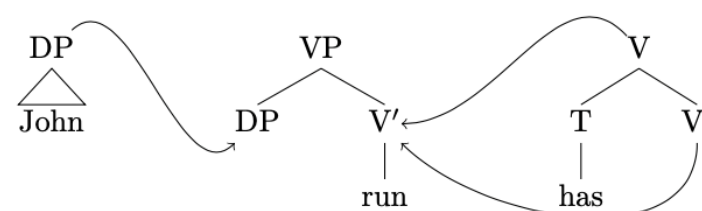


Figure 6. Elementary trees following [30,31].

Note that in (Figure 6), not only lexical categories N and V head their own elementary trees, but the auxiliary *have* also has, which is a T head (thus, functional).

Elementary trees can also be structured as in [19,21], see Figure 7.

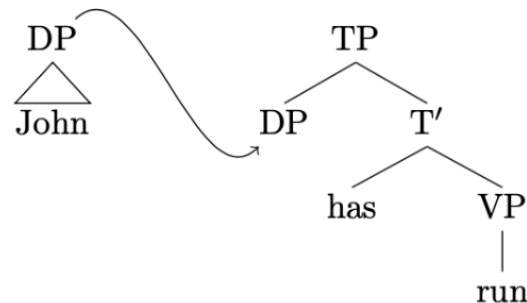


Figure 7. Elementary trees following [19,21].

In this case, only lexical categories head independent elementary trees: the DP is the extended projection of the lexical category N, and the TP is the extended projection of the lexical V (see also [24,25]).

Finally, the last option we will consider here is the proposal in [32,33], as in Figure 8.

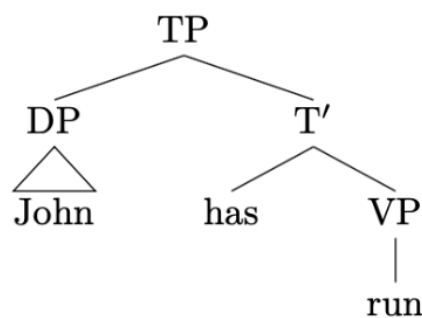


Figure 8. Elementary tree following [32,33].

The third possibility is that only a lexical predicate (as opposed to any lexical head) defines an elementary tree. Figure 8 contains only one such predicate (*run*), with perfective *have* and the subject DP being part of the extended projection of this predicate.

The following subsections will deal with the application of these systems to phenomena in Spanish and German, focused on the structure of auxiliary verb constructions and the possibilities that a lexicalised TAG offer for the development of a comparative agenda.

3.1. Elementary Trees in Spanish Auxiliary Chains

As we said above, we need to define exactly what it means for a node to be a lexical predicate: [34] and subsequent works focused on the syntax of the verbal domain, a lexical predicate is defined as a syntactic object that can both modify another object and be modified itself. In the case of perfect *have*, it modifies the lexical verb but cannot be modified itself. Suppose that we had the sentence “John will have run”: in this case, the temporal (future) information conveyed by *will* [35] modifies the event denoted by the lexical verb, as does the perfect aspect. Furthermore, *have* does not take nominal arguments or assign thematic roles: in this sense, it fits the traditional characterisation of auxiliaries as elements that lack argument structure and thematic grid. Thus, perfective *have* is a predicate, yes, but not a lexical one. ²

Whereas the functional nature of perfective *have* (or its counterpart *haber* in Spanish) is hardly contested, the case of modal auxiliaries is more complicated and thus requires more detailed discussion. Let us consider the derivation proposed by Frank for a sentence such as, “The senator should read a speech” as in Figure 9:

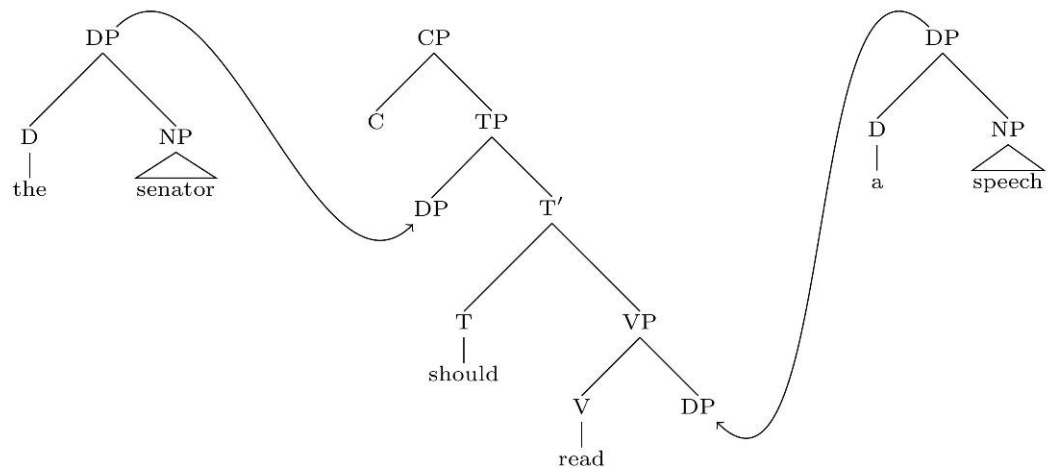


Figure 9. Derivation of ‘the senator should read a speech’ according to [21].

Here, the extended projection of read (the minimal elementary tree containing read) contains should and excludes the nominal arguments (since N is a lexical category, and as such, defines its own elementary tree). In a representation along the lines of [32,33] for the English example, there would be no substitution needed for the nominal arguments (since the definition of an elementary tree includes not only functional modifiers of the lexical predicate but also its arguments), but otherwise, the structure would look very much the same. However, when we consider the corresponding Spanish example, things become more complex:

El	senador	debería	leer	un	discurso.	
The	senator	must-COND	read-INF	a	speech.	(1)
The senator should read a speech.						

The main question here is whether the auxiliary *deber* (‘must’) belongs in the same elementary tree as the lexical verb *leer* (‘to read’). Under the definition of elementary trees suggested above, if there is a functional predicate that modifies *deber* but does not modify *leer*, then it means that *deber* is indeed a lexical predicate and, as such, may head its own elementary tree. Consider, then, the case in (2) (for the deontic reading of *deber*):

El	senador	ha	debido	leer	un	discurso.
The	senator	have-3SG-AUX-PRF	must-PART	read-INF	a	speech.
(2)						
The senator has had to read a speech.						

There are two related arguments in favour of the claim that the perfect does not modify the lexical VP, just the modal auxiliary: one is semantic, the other configurational. We may illustrate the point using a proof by contradiction of sorts. Assume, for the sake of argument, that the structure of (1) involves a cartographic-style syntax: monotonically and uniformly binary-branching and highly articulated, based on a fixed universal functional hierarchy as diagrammed in Figure 10:

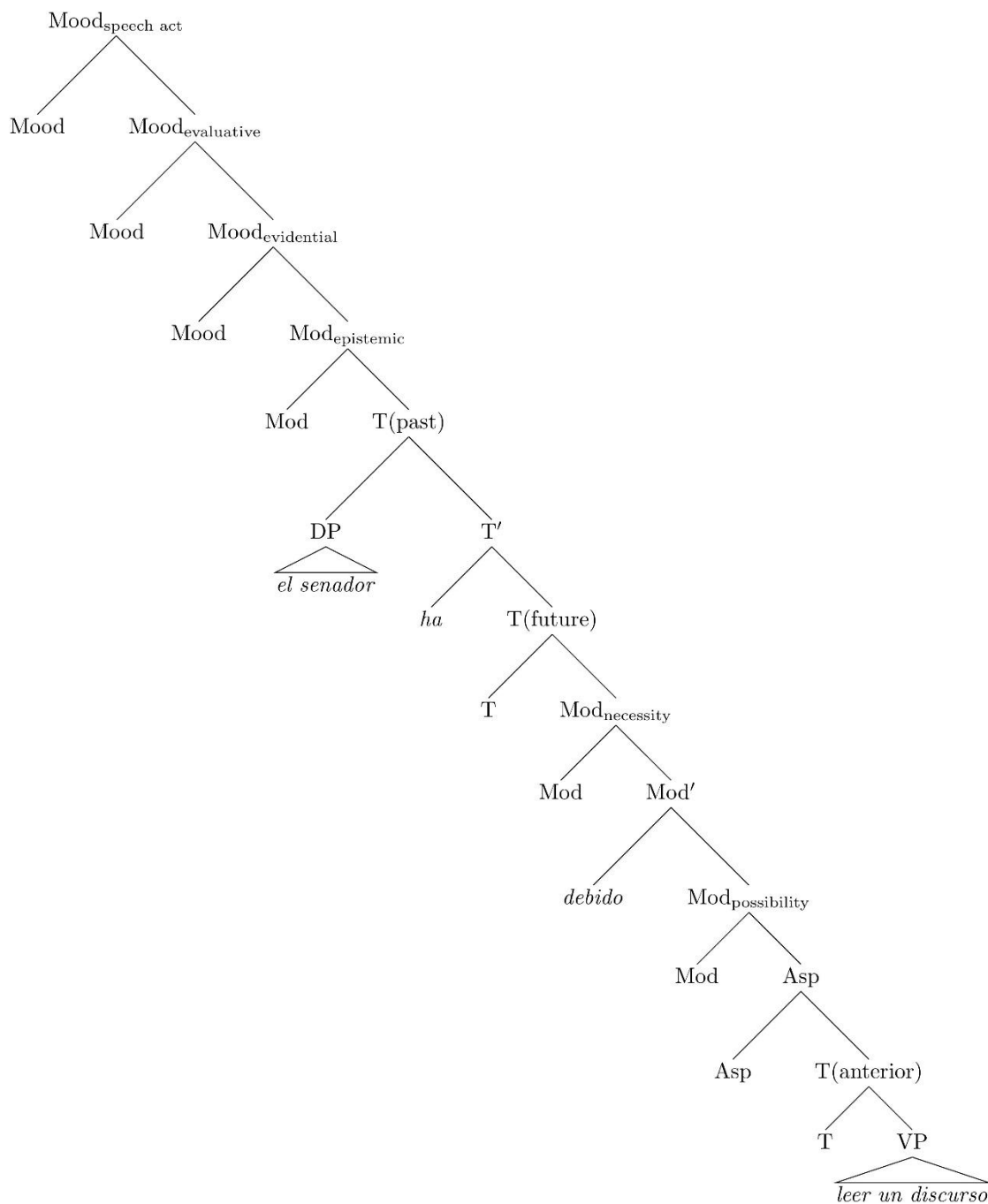


Figure 10. Cartographic clausal hierarchy.

In Figure 10, the spine that corresponds to the sequence of auxiliaries and the VP grows monotonically by introducing a single terminal node per derivational step in an order that follows an a priori fixed universal hierarchy (e.g., [9], p. 133; [10,11]). The issue arises when we evaluate the syntactic and semantic predictions that such a structure makes with respect to our Spanish cases. In a configuration such as Figure 10, given a standard syntactic definition of scope [36], we expect the perfective auxiliary *haber* to have scope over the modal *deber* and the lexical verb *leer*. If it did, then (1) should entail (3), contrary to fact:

El	senador	ha	leído	un	discurso.	
The	senator	has-3SG-AUX-PRF	read-PART	a	speech.	(3)
The senator has read a speech.						

In (3), the event denoted by the lexical VP is modified by the perfective auxiliary, and since active participles in Spanish have an anteriority interpretation (for Spanish, see [33]; for English, [37]), then we would expect the VP to be interpreted as a past event. This interpretation, furthermore, is a consequence of the assumption that the clausal structure is a monotonically growing binary-branching skeleton. Why? Because *X c-commands* all and only the terms of the category *Y* with which *X* was paired (by Merge or by Move) in the course of the derivation ([38], p. 329).

This means that under a cartographic view of phrase structure, it is unavoidable that in (8) *haber* *c-commands leer* and thus has scope over it. However, such a syntactic configuration would predict the wrong semantic interpretation: only the modal *deber* is affected by the perfect and receives an anteriority interpretation. In other words: it is only the obligation to read a discourse that is affected by the perfect, not the event of reading itself (which may or may not have taken place). As argued in [32], one of the major differences between the English and Spanish auxiliary systems lies in the fact that Spanish modals are morphologically unremarkable as verbs, in the sense that they have full inflectional paradigms (of both finite and non-finite forms). As such, Spanish modals can appear as complements to any other auxiliary since they have a full set of non-finite forms. Thus, since the modal *deber* has a participle, it may be modified by perfective *haber*: semantically, this combination is interpreted as a possibility (under the epistemic reading of *deber*) holding in the past.

Crucially, a structural description such as (9) (maybe with less functional projections, but configurationally identical) is unavoidable under currently standard minimalist assumptions: every tree is at every point binary-branching because the structure-building mechanism (*generalised transformations*, as in [15,16]; Merge, as in [39]; MERGE as in [40]) can only manipulate *two* objects at a time. Furthermore, because of the assumption of structural uniformity that is at the core of minimalist theorising, there is no room for a different kind of structural description. If we require a theory of the grammar to be descriptively adequate, then there are reasons to reject Figure 10 as an adequate structural description for (1) on semantic bases (to the extent that these depend on syntactic configuration).

Here is where a TAG treatment becomes especially attractive: a TAG allows us to define more than a single elementary tree to compose a structural description. Therefore, if we can define a syntactic object that includes (properly contains) *ha debido* and excludes *leer un discurso*, we can create a configuration where the problematic scope relation identified above does not hold.³

We can do exactly this if we assume that Spanish modals, unlike their English counterparts, are lexical predicates, and as such, may head their own elementary trees. Therefore, where English needs a single elementary tree, Spanish may require more. Summarising a lot of previous work, this is a tenable hypothesis if we consider that (i) Spanish modals, unlike English modals, are morphologically unremarkable verbs, thus (ii) they can be selected by other auxiliaries (when they are, they “absorb” the information conveyed by auxiliaries to their left [24,32]), and (iii) it has been argued that they can be thematic assigners, just like lexical verbs. We know they are auxiliaries because they can combine with meteorological verbs and because of their semantic contribution to the compositional interpretation of the structures in which they appear (epistemic/deontic/dynamic in the most widely accepted classifications). Thus, we can define the (slightly simplified) elementary trees in Figure 11:

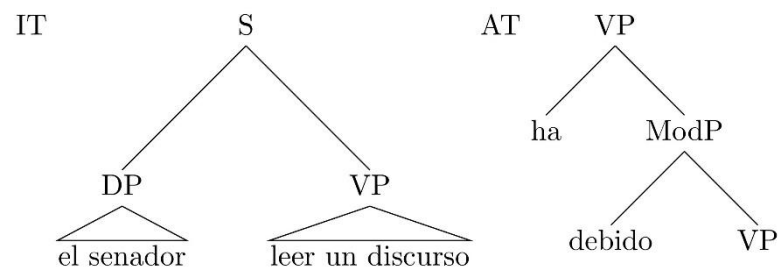


Figure 11. Elementary trees for ‘El senador ha debido leer un discurso’.

Under the assumption that Spanish modal auxiliaries are lexical predicates, it becomes possible for us to define an auxiliary tree that contains the modal and its functional modifier, the perfective auxiliary *haber*, and which excludes, naturally, the lexical verb. We can provide further evidence that the perfect only affects the deontic modal and not the lexical verb: there is no actuality entailment in (1). Therefore, (4) is not contradictory:

El	senador	ha	debido	leer	un	discurso,	pero	
The	senator	have-3SG-AUX.PERF	must-PART	read-INF	a	speech,	but	
no	lo	ha	hecho					(4)
NEG	CL-ACC	have.3SG-AUX-PERF	do-part.					

“The senator has had to read a speech, but he hasn’t done so”

The interpretation of (1) is one in which only the obligation conveyed by the deontic modal receives an anteriority interpretation due to the perfective auxiliary, but not the event denoted by the lexical verb. The same situation arises with temporal auxiliaries; consider the case (5), which combines future auxiliary <ir a + infinitive> with a dynamic modal *poder*:

El	senador	va a	poder	leer	un	discurso.	
The	senator	goes-to.3SG-AUX-FUT	be able to-INF	read-INF	a	speech.	(5)

“The senator will be able to read a speech”

In (5), it is clear that what is temporally anchored in the future is not the event of reading, which may or may not take place at all, but the senator’s capacity to engage in such an event of reading. To capture this reading, we propose, it is necessary to define a syntactic object that includes the temporal auxiliary and the modal but excludes the lexical verb, a configuration that is unavailable given the structure in Figure 10.⁴

The non-monotonicity of TAG derivations, combined with a lexicalised approach to the substantive definition of elementary trees, allows us to provide an adequate structural description for the relations between auxiliaries in Spanish chains.

3.2. German Modals: Different Elementary Trees?

When it comes to verb clusters in Continental West Germanic, variation seems to play a major role. The OV nature of these languages causes both finite and non-finite verb forms to show up in a clause-final stack with different (morpho-syntactic) properties and different reordering rules. Essentially, the analyses proposed in the early 1980s boils down to the concept of *verb raising* (VR). The standard analysis of this phenomenon is put forward in [41]: assuming a biclausal analysis of verb clusters, the embedded verb is taken to be extracted from the complement clause and Chomsky-adjoined to the matrix verb (to its left in German and to its right in Dutch):

Verb Raising:

- V₁]s V₂
- a. ... e₁][V₁ V₂]_V (German)
- b. ... e₁][V₂ V₁]_V (Dutch)

Along the lines of Evers, a derivation in Dutch might look like (6):

a. ...	dat that	Jan [PRO [Jan	een a	huis house	kopen]VP]S buy	wil (D-Structure.) wants.3SG	
b. ...	dat that	Jan [PRO [Jan	een a	huis e] VP]S house	wil want.3SG	kopen (VR) buy	(6)

“that Jan wants to buy a house”

The papers [42,43] represented a major breakthrough in the treatment of VR within the GB approach. In particular, the latter assumed that this phenomenon consisted of two separate processes: (i) syntactic reanalysis of the verb sequence resulting in clause union and (ii) inversion of the constituents in the sequence.

Moreover, the so-called *infinitious pro participio* (IPP) construction represented (and in fact still represents) a thorn in the side of scholars who specialise in West Germanic: these constructions feature a V_1 - V_3 - V_2 (Aux-V-Mod) sequence in embedded clauses, where the modal verb resists a participial morphology (for instance, *gemusst* shows up as *müssen*) and is not adjacent to the auxiliary that selects it (e.g., *hätte schreiben müssen*, “would.have write must.INF”). Over the years, different approaches have attempted to tackle the anomalous sequence V_1 - V_3 - V_2 that does not comply with the expected canonical V_3 - V_2 - V_1 order of German (we refer the reader directly to [44]).

The paper [45] concentrated on the “verb raising” nature of West Germanic verb complexes and, following the lead of [43], emphasised two specific aspects of the complexes: (i) the type of complement selected by modals and auxiliaries and (ii) the rule of extraposition (via Chomsky-adjunction) that some types of infinitives were assumed to undergo. We will see that (i) plays a major role the complements being either S-nodes or VPs while (ii) was assumed to account for the great deal of variation in West Germanic verbal complexes: in fact, if we stick to the hypothesis that modals, causatives and perception verbs pattern in the same way, coherently subcategorising for VPs and optionally undergoing raising to a higher V, then, the specific locus of adjunction and the possibility of permutations (V_3 - V_2 - V_1 / V_1 - V_3 - V_2 etc.) become the main ingredient of cross-linguistic variation. Take, for instance (7a,b) that features cluster perception-causative-V:

a. ...	dass that	Hans Hans	Peter Peter	Marie Marie	schwimmen swim	lassen let	sah (Germ.) saw.3SG-PST	
b. ...	dat that	Jan Jan	Piet Piet	Marie Marie	zag saw.3SG-PST	laten let	zwemmen (Dut.) swim	(7)

“that Hans saw Peter let Marie swim”

“that Jan saw Piet let Marie swim”

As we have seen in (6) and (7), the order of Vs in the Dutch complex—which is the mirror image of its German counterpart—is derived by assuming infinitive extraposition. On the contrary, German displays a coherent OV final ordering in this case: extraposition was assumed, e.g., for modal perfective constructions where the finite V precedes the 2-infinitive complex.

Incidentally, that the Dutch example features crossing dependencies between two categories (in abstract terms, using integers to indicate dependencies between nouns and verbs, $N_1 N_2 N_3 V_1 V_2 V_3$), the hallmark of mild-context-sensitivity. In contrast, the German example expresses the same proposition by means of center embedding ($N_1 N_2 N_3 V_3 V_2 V_1$), which is strictly context-free. The reader can see that the English translation only makes use of tail recursion ($N_1 V_1 N_2 V_2 N_3 V_3$), a formal mechanism available, in principle, in regular grammars. Given their explicitness and flexibility in the definition of elementary trees, LTAGs are particularly well-suited to represent this cross-linguistic computational variation.

Traditional assumptions on extraposed infinitives/verb raising can also be expressed in terms of a lexicalised TAG (i.e., the perspective they adopt in their paper) resorting specifically to Trees with links, “[...] where link refers to the relationship between an empty category and the Chomsky-adjointed antecedent that binds it” ([45], p. 310), see Figure 12:

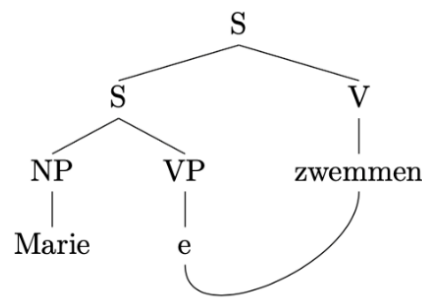


Figure 12. Elementary tree with VR.

Thus, the differences between German and Dutch might be captured with structurally identical trees, where different adjunction sites (structurally adjacent S-nodes) are assumed: this approach was defined as a string-vacuous raising hypothesis and apparently accounts for cross-linguistic variation in an elegant and straightforward way as it “reduce[s] variation in the word order of verb sequences to the choice of node at which auxiliary trees are adjoined [...]” ([45], p. 317). However, Kroch and Santorini point out that such an assumption might imply serious structural issues: for instance, a PRO subject of a modal might even end up c-commanding its antecedent (which would make it impossible to establish an appropriate binding relation). Therefore, they put forward a different analysis that discards string-vacuous verb raising and uniformly assumes that modals (and “functional” verbs) uniformly subcategorise for VP complements. Their proposal will be spelt out with a Franconian example (8a) and a standard German example (8b):

- | | | | | | | |
|--------|-------------------------------|-----|-----------|-----------|-------------|-----|
| a. ... | dass | er | singen | hat | müssen | |
| | that | he | sing | has | have-to-INF | |
| | “that he had to sing” | | | | | |
| b. ... | dass | sie | hätte | schreiben | können | (8) |
| | that | she | have-SUBJ | write | be-able-to | |
| | “that she could have written” | | | | | |

First of all, the modal *müssen* and the auxiliary *hat* are taken to head their own auxiliary trees (note the absence of extrapositions), whereas the lexical verb is represented in the initial tree with root S as in Figure 13:

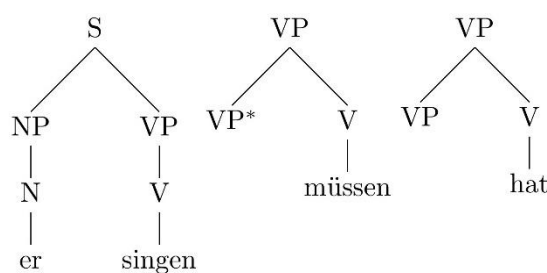


Figure 13. Elementary trees for the examples in (8).

We can now obtain (8a) by adjoining the first auxiliary tree in the VP node of the initial tree and the second auxiliary tree in the VP * node, which has become the lower VP node after the first adjoining. This is represented in Figure 14:

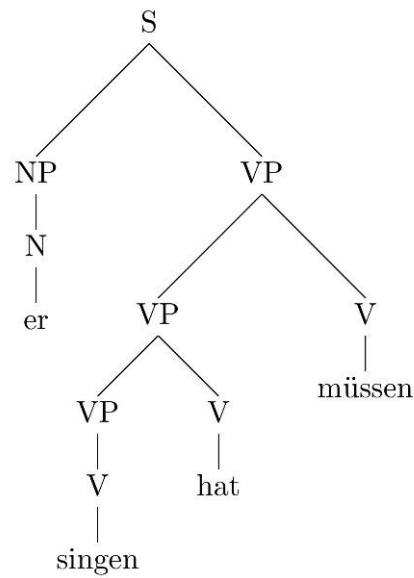


Figure 14. Derived tree after adjunction.

A certain degree of variation still remains in Kroch and Santorini's [45] model as the standard German verb complex V_1 - V_3 - V_2 (i.e., the IPP construction) requires Chomsky-adjunction of the infinitive to VP in the elementary tree. We can provide the elementary trees as in Figure 15:

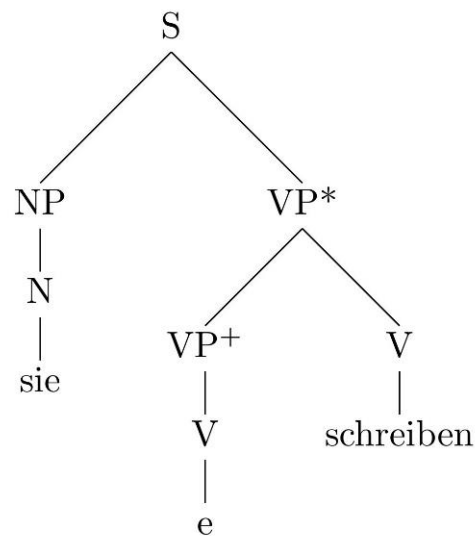


Figure 15. Elementary trees featuring V extraposition.

The two auxiliary trees are assumed to adjoin at the VP* and VP+ sites in the initial tree, where the symbols * and + are originally used in [45], p. 322, as diacritics for nodes that are targeted by adjunction. The resulting tree composition is (Figure 16):

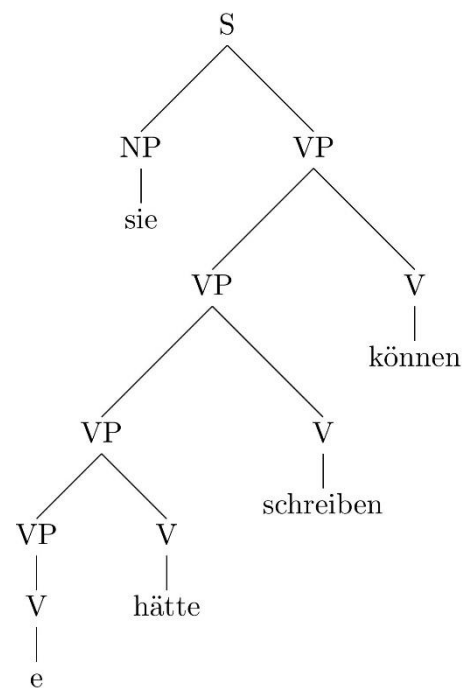


Figure 16. Derived tree with V extraposition.

Kroch and Santorini’s conclusion is that variation arises due to two different factors: (a) the choice between elementary trees with or without extraposed infinitives and (b) the choice of different nodes as sites for adjunction (the specific “height” of the target nodes within the initial tree). At any rate, discarding the string-vacuous analysis is a better solution from a theoretical perspective: moreover, the learners would not be forced to postulate the existence of trees with links ([45], p. 323). (We express our gratitude to a reviewer for pointing this out to us).

One might observe that our proposal does not rigidly constrain the possible permutations that can be obtained by adjoining at different adjunction sites: however, the degree of variation is not unconstrained. Suppose we maintain the derivational order “V extraposition > adjunction” assumed by Kroch and Santorini. In this case, the only other possibility we are left with, apart from (8), is adjoining the *hätte*-auxiliary tree first and afterwards the *können*-auxiliary tree in the same adjunction sites. What we get is the string **können schreiben hätte* (V_2 - V_3 - V_1), which is excluded in standard German and extremely rare but still possible in some northern varieties (especially Dutch). Conversely, suppose we allow ourselves the opposite derivational order, i.e., “adjunction > V extraposition”: knowing that adjunction bleeds extraposition (since once we have adjoined the auxiliary trees, there is no possible target for V extraposition), we are now left with a new initial tree (Figure 17) that can be the target of adjunction for the same auxiliary trees of (Figure 15):

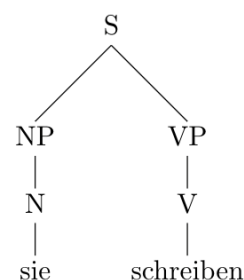


Figure 17. Elementary tree without V extraposition.

Adjoining the two auxiliary trees to the same initial tree in different orders, we obtain the strings *schreiben hätte können* (V_3 - V_1 - V_2) and *schreiben können hätte* (V_3 - V_2 - V_1): the former is rare but possible in some southern German and Austrian dialects, the latter is the strict head-final modification, found in several German, Austrian and Dutch non-standard varieties (in this case, the surface construction can actually be *infinitive-participle-auxiliary*). Thus, the derivational order and the possibility of extraposition give rise to a limited variety of combinations that makes sense of the variation found among the dialects: recall that (i) we are dealing with an initial tree and two auxiliary trees, and (ii) that adjunction always inserts auxiliary trees into initial trees (at locations determined by the identity of root-frontier labels), which limits the number of superficial strings that can be obtained.

In our discussion of the TAG approach to German, the reader may have noted that *hätte* ('have') was assigned its own elementary tree but in the Spanish discussion, perfective auxiliary *haber* ('have') was part of the elementary tree anchored by a modal: the question arises, then, whether *have* is a lexical predicate in German. Recall that a TAG is lexicalised if and only if each elementary tree is the projection of a single lexical head ([28]). As a matter of fact, this may be a point of cross-linguistic variation: auxiliaries may be lexicalised to different extents in different languages, and this impacts directly on the size of elementary trees for sequences of auxiliary verbs and verb complexes.

The argument in this section illustrates the first point mentioned above: lexicalised TAGs allow us to define a picture of cross-linguistic variation that differs from the traditional minimalist view (although it may be construed as complementary to it): note that if English modals are not lexical heads, it means that elementary trees for sequences of auxiliaries in English will be, in general, bigger than elementary trees for sequences of auxiliaries in Spanish: at most, a chain of auxiliaries in English houses four auxiliaries.

“John must have been being helped” (Modal + perfective + progressive + passive) (9)

In a representation à la Frank, there would be two elementary trees for (9): one anchored by *John*, and the other by *helped*. The limitations of English chains of auxiliaries are related to (i) the fact that modal verbs are defective, without non-finite forms, and (ii) predicates that in other languages behave as auxiliaries (such as *start*, *finish*, *continue*, *attempt*) in English fail the most fundamental tests for auxiliarihood (ability to host negation, inversion in interrogative sentences, etc. [35]). Note that the structural description assigned to English auxiliary chains by an LTAG is strictly monotonic, and in principle, compatible with cartographic assumptions, at least configurationally. However, TAGs allow us to capture aspects of cross-linguistic variation by not assuming a priori that all sequences of auxiliaries (or indeed all elementary trees) in all languages are created equal. This approach to variation may be thought of as complementary to the mainstream generative one, whereby functional categories are identified as the locus of cross-linguistic differences: in an LTAG, there is variation to be found in the definition of elementary trees depending on whether a particular expression behaves as a lexical or a functional head in a given language. Thus, there are reasons to classify Spanish modal auxiliaries as “lexical” in the relevant sense (they can modify and be modified), but not English modals. Frank [21], p. 238, formulates the issue in the following terms:

“The set of elementary trees in the grammar of a language is the only locus of cross-linguistic differences, imposing substantial restrictions on the kind of variation that we will find crosslinguistically”.

Our earlier observation about the relative size of elementary trees and the length of chains can be somewhat generalised. If a language’s auxiliary verbs (in particular, auxiliaries that take infinitival complements) are lexical predicates, then sequences of auxiliaries may be longer, but their component elementary trees will be smaller. This observation is related to Menzarath–Altmann’s law (MAL; [46]), although we must note that the MAL makes no reference to hierarchical relation; rather, it is based on quantitative properties of textual units. However, the inverse relationship between the size of composite units and the size

of elementary units (such that the bigger the composite unit, the smaller its component parts) seems to emerge from the comparison of English and Spanish auxiliary chains.

4. Locality and the Definition of Dependencies

The previous section dealt with the first of the issues we identified: the definition of elementary trees and a revision of the CETM. The second issue, how to express non-local dependencies in terms of local relations, is perhaps where the novel aspects of our contribution become more explicit.

We need to define what the nodes in the elementary trees stand for. In other words: structural descriptions specify relationships between objects; once we have defined the allowed configurations in which those objects may occur (elementary and derived trees), we can look closer at the atomic objects themselves.

Let us briefly consider some difficulties that arise with the traditional minimalist view that syntactic terminals are sets of features (as in [39,47,48]; see [49], p. 266, for a version of this critical argument).

Suppose that we have a syntactic object α , which is defined by a set of features, some of which are interpretable and some of which are uninterpretable (whether this uninterpretability is related to valuation or not is orthogonal to the present discussion). We thus define α as follows (where i = interpretable, u = uninterpretable):

$$\alpha = \{iF_1, iF_2, uF_3, uF_4\} \quad (10)$$

Now suppose that a suitable head H that can value/check the uninterpretable features of α has been introduced in the derivation. Minimalism [7,12] takes this to be the kind of situation that motivates the property of displacement, and therefore α must be copied and internally merged in a position where H and the copy of α are in a suitable relationship for the elimination of uninterpretable features (indicated with strikethrough in (11) below). Otherwise, the syntactic object containing α will collapse at the interfaces. The problem is the following: how are α and its copy related? If the internal merge of α 's copy has succeeded in creating a suitable checking configuration, then the featural definition of this copy will be:

$$\alpha_{\text{Copy}} = \{iF_1, iF_2, \cancel{uF_3}, \cancel{uF_4}\} \quad (11)$$

However, clearly, α and its copy are not defined by identical feature sets.

This means that if operations are indeed driven by the necessity to evaluate/check features [6], no two copies of a syntactic object are defined by the same feature specification. One may say that if we have a chain, then having any featural change in one link "automatically" affects all other links. That has several problems. One is that operations in standard formal systems (other than cellular automata) cannot take place simultaneously. Therefore, one link of the chain will necessarily be affected first. Then, the problem is how to determine what else is a link in that chain, given that one of the links now looks different. It can be circumvented with indices, but the problem remains. Another is that for this to work, we would need all links to be accessible: this goes against locality theory, in particular, the idea that some syntactic domains ("phases"), when completed, get transferred to the performance systems. If there is a copy in a transferred domain, it cannot be tampered with. Concretely, in what did John buy, the copy of what is in the complement position of buy is not accessible by the time we check the *wh*-feature in Spec-CP, since the VP will have been transferred.

Identity is not possible in a system that admits feature checking relations since the feature matrix of an element varies as the derivation unfolds and copies are internally merged in places in which they can evaluate and erase uninterpretable features. Any link between α and its copy should be encoded as a diacritic of sorts (e.g., a referential index that is carried throughout the derivation). An alternative could tamper with the "foot" of a chain, which may or may not be accessible. A similar argument about the problems of assuming copy-creation mechanisms in a feature-based computational system is presented

in [50], whereby taking the chain system in [39] seriously leads to a multiplication of entities in chains and the unwanted consequence of every chain having at least one member with unvalued/unchecked features: how that chain can receive an interpretation is a problem still unsolved, particularly given the strong set-theoretic commitments of more recent minimalist proposals. Some developments (e.g., [40,51]) work around a distinction between copies and repetitions, whereby a copy (produced by Internal Merge) would occupy two positions at the same time: as convincingly argued in [52], this is incompatible with the set-theoretic foundations of structure building. Requiring identity in movement chains under the copy theory of movement is not tenable under standard definitions of identity (for reasons related to set-membership as well as featural composition) was argued in [50].

There are several alternatives to this view, most of which entail abandoning the so-called single mother condition [53]: the requirement that every node in a tree have a single mother node (i.e., be dominated by only one node, distinct from itself). These include parallel merge [54] and several versions of multidominance ([50,55–57]). In classical TAG ([18,58]), filler-gap dependencies are captured by means of so-called links. Links in TAGs are relations between two nodes A and B, such that A c-commands B in an elementary tree and B dominates a null string. Links are crucially maintained after adjunction: they cannot be disrupted or created by means of tree-composition operations. In Joshi’s introduction of links [18], p. 214 no mention is made to issues of reference, but since links do not affect the weak generative capacity of grammar, the same results can be obtained with a TAG where all links have been removed [18], p. 217. This device is not exclusive to (or original to) TAGs: Peters and Ritchie [59] define the so-called *phrase linking grammars*, where links are used to relate displaced constituents to their “base generation” position in directed acyclic graphs. In classical TAGs, links relate operators to variables and remain unaffected by the adjunction of recursive structure: as such, filler-gap dependencies are always defined in elementary trees.

The main difference between a TAG and a minimalist approach to filler-gap dependencies is that there is no inter-clausal movement in TAGs (lexicalised or not). We may compare the derivation of a *wh*-interrogative such as *what do you think the senator should read?* in TAG terms and minimalist terms to illustrate this point.

First, TAGs. In a lexicalised system a la Frank, we have three elementary trees:

- i. The extended projection of the lexical verb *read*.
- ii. The extended projection of the lexical verb *think*.
- iii. The extended projection of the noun *senator*.

Let us diagram the derivation in Figure 18, which involves both substitution and adjunction:

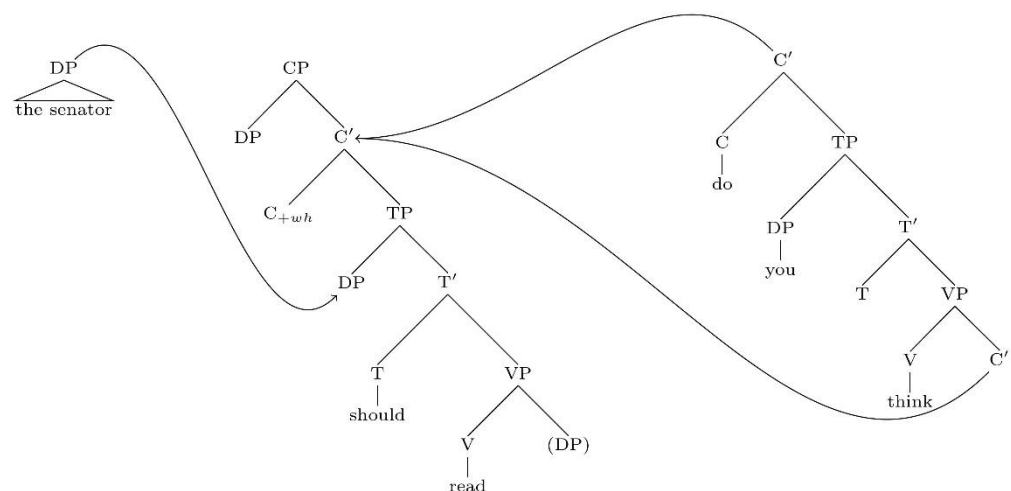


Figure 18. Derivation with Substitution and Adjunction.

The final derived tree would look like the one in Figure 19:

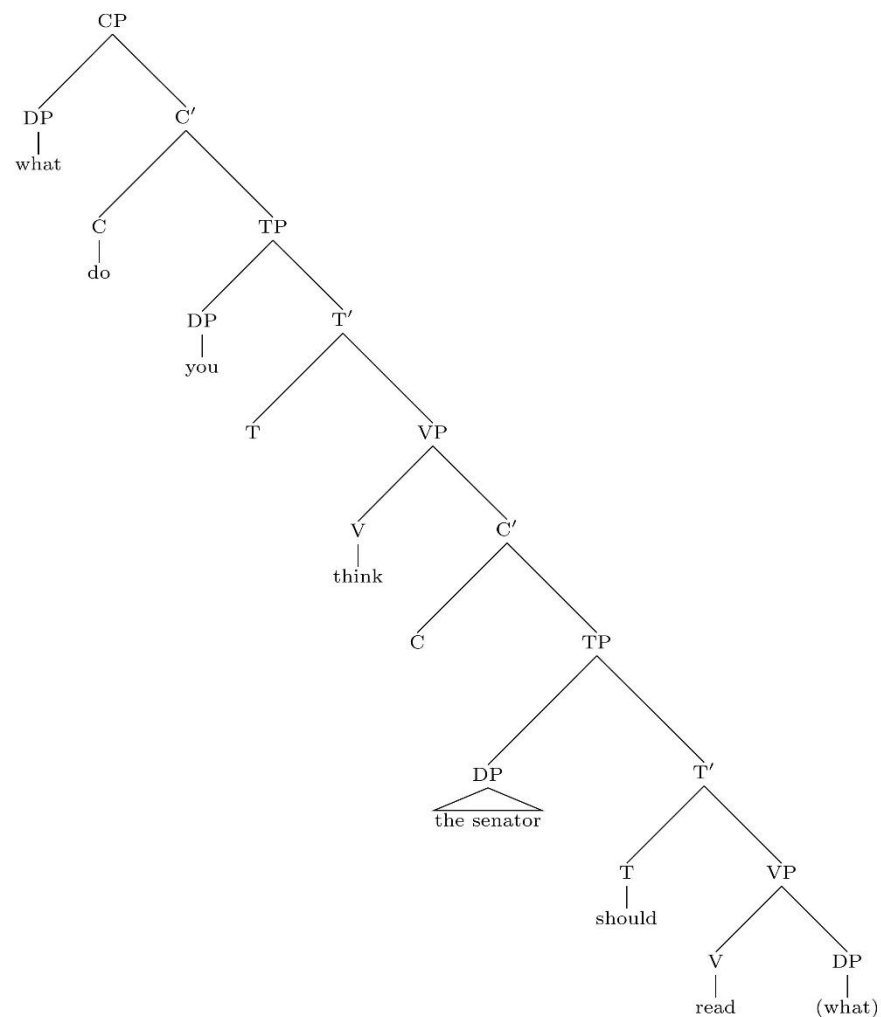


Figure 19. Derived tree for ‘what do you think the senator should read?’.

It is crucial to note that in Figure 19 there is no actual movement: the relationship between the empty string in the complement position of V *read* and the *wh*-operator in Spec-CP was defined at the level of the elementary tree of *read*, with both the operator and the variable being, as it were, “base-generated”. The adjunction of the think elementary tree between the *wh*-operator and the rest of the read elementary tree just simulates the effect of movement: as we will highlight below, the relation between *read* and *what* remains undisturbed. Note also that there is no need to propose “intermediate landing sites”, which is consistent with the characterisation of *wh*-movement as an *unbounded rule* (U-rule) in [60], whose application takes place in one fell swoop.⁵

In minimalist terms, the structure would be built from the bottom-up, adding one node at a time, and the relation between the operator and the variable would be established only after movement.

This gives us a segue to the comparison with a transformational approach to displacement. The minimalist approach is based on the idea that constituent movement is motivated by the need to eliminate uninterpretable features from syntactic representations. The property of displacement is, according to [6,12], an “imperfection” of the faculty of language, together with the existence of uninterpretable features. These two “imperfections” cancel each other out, so to speak, if displacement allows the computational system to eliminate uninterpretable features from phrase markers. Crucially for our purposes, however, the implementation of displacement in transformational generative grammar requires

a multiplication of entities: in the copy theory of movement, the displaced constituent gets copied and re-introduced (re-merged) into the structure. Furthermore, the theory of locality in minimalism, based on designated nodes in the structure (phase heads), requires additional copies of displaced constituents in an “escape hatch” position (an additional specifier) in each phase (a very similar situation to what happened in the Barriers model, as noted in [61]). Assuming that C and *v** are phase heads, a long-distance *wh*-interrogative requires (for intra-theoretical reasons) a copy of the displaced object in each Spec-C and Spec-*v** position in between the position where the *wh*-element was first merged and the final Spec-CP position, where it can check its [*uWh*] feature as in Figure 20:

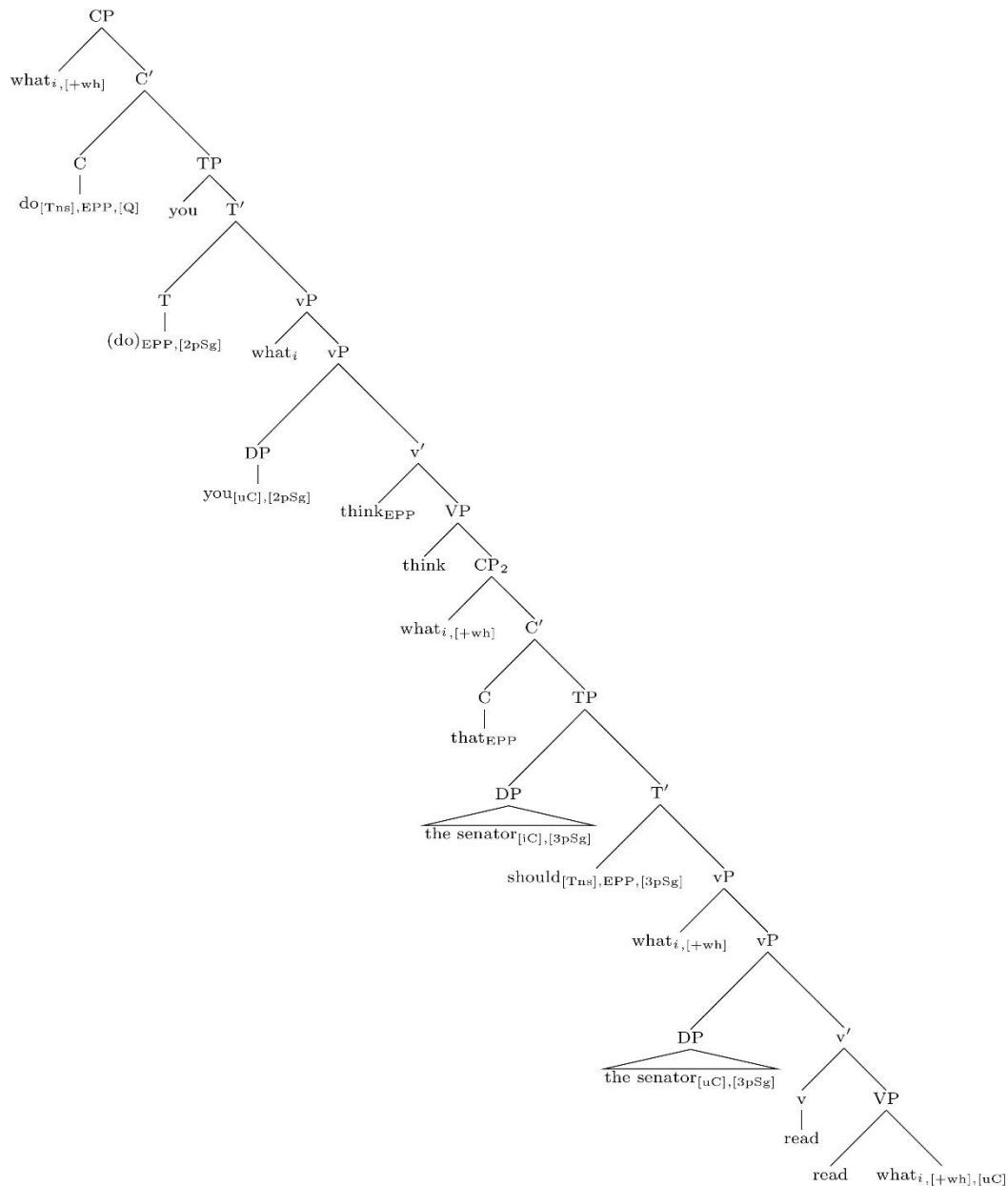


Figure 20. Minimalist derivation for ‘What do you think that the senator should read?’.

One of the fundamental questions when analysing filler-gap dependencies is whether movement is necessary at all. Non-transformational generative theories of displacement (e.g., LFG, HPSG) have been very successful since the late 1970s, and some versions of TAGs [18,28,58] are indeed non-transformational in the sense that they do not include rules for singular transformations (such as *wh*-movement), but only generalised transformations

(specifically, *substitution* and *adjunction*) and (in some cases) the concept of *links*. This is not to say that TAGs are inherently incompatible with singular transformations; as a matter of fact, the version of TAG explored in [19,21] is closer to transformational generative models than to non-transformational models (read: “without singular transformations”). The TAG formalism specifies how elementary trees are put together (and, if each elementary tree receives a local semantic interpretation, then it also defines a theory of compositionality), but it remains agnostic about the operations that apply within elementary trees; these are subject to independent empirical inquiry.

As part of our approach to the mechanisms by means of which non-local dependencies are reduced to local ones, we will briefly explore an alternative that dispenses with a rule of movement in the derivation of *wh*-dependencies (see also [32] for an application of this model to gerund fronting and the position of subjects in Spanish).

Our proposal makes use of a formal tool that is not new to TAG theorising: we define nodes in elementary trees as addresses to locations in the lexicon. In other words, elementary trees do not contain lexical tokens but addresses to lexical types. This is based on an idea by [62], which they implement in the context of a discussion about coordination in TAGs:

Each node in the tree has a unique address obtained by applying a Gorn tree addressing scheme.

This device is implemented in an account of *gapping* and *right node raising*. We may note that it is not unlike multidominance in terms of the tree language it defines, but it actually specifies how it is implemented computationally. We are extending this addressing mechanism not only for coordination schemata but also for relations within elementary trees.

Each time a terminal is introduced in an elementary tree, what is actually manipulated is an address for that terminal. This is important since an address can be called for several times in the execution of a program, in our particular case, structure building and compositional interpretation. All we need the grammar to do is recognise nodes that are assigned identical addresses. This is not requiring the computational system to do anything other than what it already does: in order for substitution or adjunction to work, identically labelled nodes in distinct elementary trees need to be identified as one. What we suggest is that this process can affect not only root-frontier nodes but also lexical nodes inside elementary trees. This entails that we can simplify the structure in Figure 19 since it contains two calls for the same address *what*. This simplification is one of the main points of the present paper. There are two ways to model this: one, topological; the other, graph-theoretical. In both cases, the formal object being defined is the same, and crucially, the structural relations between elements in the structural description are the same.

The topological view entails defining elementary trees in a workspace understood as a topological space; this space can fold and self-intersect (this proposal has parallels in the field of DNA topology). The points of self-intersection are the objects that appear in more than one structural context: the syntactic workspace folds to identify different tokens of the same type, yielding a higher-dimensional manifold (a 2-D space folds creating a 3-D manifold), where in a *wh*-interrogative like *what do you think the senator should read?* *what* in Spec-C and *what* in Compl-V have a zero distance. This is only possible in a metric topological space if *what* in Spec-C and *what* in Compl-V are identical (as per the so-called *identity property* of metric spaces, whereby $d(x, y) = 0$ iff $x = y$; see [63,64]) as desired (this approach to movement is developed in [65]).

Here we will briefly pursue a different view: the graph-theoretic account (presented in detail in [66]). The general tenets of this approach are already familiar to generative linguists, as per multidominance theories: the core idea is that instead of constructing unordered sets (a decision that forces the multiplication of nodes in instances of displacement and more generally yields problems of distinguishing copies and repetitions; see [40,51]), the grammar generates local graphs. However, the justification for structural descriptions that do not follow the single mother condition [53] in the present work is not related to properties of merge (as it is in Citko’s or Johnson’s work) but follows from the assumption

that what the syntax operates over are nodes, which are assigned addresses, and that these addresses are unique identifiers in a derivation. ⁶

Furthermore, and along the lines of [67], only substantive symbols are assigned addresses: in terms of a lexicalised grammar, we implement this as establishing that only (single-word or multi-word) expressions assigned to an indexed category of the grammar and have a semantic value (*catagorematic* elements in the logical tradition, see [68,69] for an application of this concept to categorial grammars) are assigned addresses. The set of *syncatagorematic* symbols depends on the language: for example, English infinitival *to* and copulative *be* have been proposed to be *syncatagorematic* [69], as have Spanish apparent prepositions and complementisers in verbal periphrases (e.g., *que* in *<tener que + infinitive>*, *a* in *<ir a + infinitive>*) as argued in [70].

This means that “... signs that signify nothing by themselves but serve to indicate how independently meaningful terms are combined” [68], i.e., *syncatagorematic expressions*, are not assigned addresses, *only catagorematic expressions are* (a similar perspective is adopted in [67], p. 214, within a purely mathematical framework: in that work, only “object characters” in formulae are assigned addresses).

What is the content of these addresses? A possibility, explored in depth in [66], is that the content of an address is the intensional logic translation of the corresponding expression ([71,72]). An introduction to intensional logic (IL) (specifically, the approach in [73] and related works) would greatly exceed the limits of this paper and is not strictly needed for our purposes. However, we can introduce some concepts that may clarify the issue of addresses, much research pending. Let $\{\{what\}\}$ be the uniquely identifying address of the expression *what*. We can look at how to define the IL translation of *what*, that is, the content of the address $\{\{what\}\}$: let *P* be a first-order monovalent (intransitive) predicate (in categorial terms, an IV). Thus, if we consider *what*, for any sentence containing *what* the content of the address $\{\{what\}\}$ would be:

$$\{\{what\}\} = P \forall x P\{x\} \quad (12)$$

(12) follows the proposal in [71], p. 19, and related works closely, in turn building on [73]. Karttunen proposes that the IL translation of bare *wh*-words, say *who*, is the same as the IL translation of an indefinite common NP, such as *someone*: there is an existentially quantified term that corresponds to a member of a set of entities, the *wh*-operator introduced the choice function that selects an entity from that set based on whether it satisfies the conditions imposed by the predicative structure in which that *wh* occurs (see also [74] for an extension of Karttunen’s system to lexically restricted *wh*-phrases). The denotation of (12) is the set of propositions such that, for a 1-place predicate *P*, *P*(*x*) holds. In other words: the denotation of *who runs?* is the set of propositions such that *run*(*x*) is true. A lot of details aside, what matters to us is that the meaning of a bare *wh*-word, such as *what* or *who*, can be assimilated to lexically restricted *wh*-phrases: suppose that we have the interrogative sentence *which discourse did the senator read*. The semantic form of this question, in Reinhart’s terms, would be:

$$\{P \mid (\exists \langle x, f \rangle) (CH(f) \ \& \ \text{senator}(x) \ \& \ P = \lambda(x \text{ read } f(\text{discourse})) \ \& \ \text{true}(P))\} \quad (13)$$

The interpretation of *which discourse* in this context is “choose an entity that belongs to the set of discourses (such that there is an *x*, *x* a senator, that read that discourse)”; *who*, in the same context, means “choose an entity (such that ...)”, with the added specification that the entity be animate and perhaps even also human (for inanimate entities, we have *what*). The “such that ... ” clause depends on the structure where the *wh*-word or phrase appears.

Let us go back to the analysis of the long-distance *wh*-dependency in *what do you think that the senator should read?* We want to capture the fact that, informally, the *what* in Spec-CP and the *what* in Compl-V correspond to the same element. The TAG formalism, as revised in [62], provides us with a way to do exactly this. We want to capture the fact that the address $\{\{what\}\}$ is called for in two contexts: one in which it receives an operator

interpretation (introducing a choice function; as in [74]), and another in which it satisfies the valency of the monotransitive verb *read*. At the same time, we want to capture the fact that there is only *one* syntactic element. Recall that, in the version of LTAG adopted in [32,33], arguments of a lexical head belong in the extended projection of that head. In this case, the TP and CP layers belong to the extended projection of *read*, and the nominal arguments *the senator*, and *what* also do. In the elementary tree defined by *read*, then, there are two contexts where the address of *what* is called, but not two distinct addresses. The result is that in the graph that defines the structural relations in the elementary tree of *read*, $\{\text{what}\}$ (which, recall, is the address assigned to the basic expression *what*) is called twice; this may be represented using multidominance, with the terminal node *what* having two mother nodes. We can diagram the elementary tree anchored by *read* as in Figure 21:

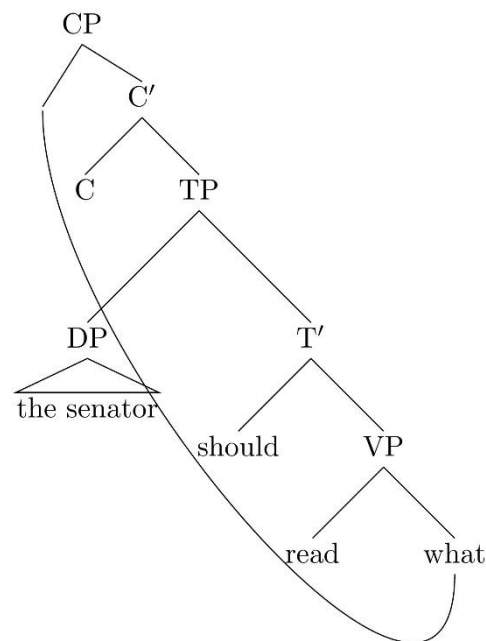


Figure 21. Multidominance structure for wh-movement.

The adjunction of the elementary tree of *think*, which includes the nominal argument *you*, does not disrupt the structural relations established at the elementary tree of *read*: the edge connecting nodes C and $\{\text{what}\}$ and the edge connecting nodes V and $\{\text{what}\}$ remain untouched, as seen in Figure 22 below (the adjoined auxiliary tree has been bolded for clarity).

In the topological interpretation, we would say that the distances between nodes remain constant after adjunction. This is important since, for any measure of distance in binary branching trees (number of nodes, number of edges, etc.), minimalism is forced to assume that the distance between filler and gap is different in *What should the senator read?*, *What do you think the senator should read?*, and *What do you think that Mary said that the senator should read?*; a consequence of this assumption is that intermediate copies are multiplied (since the theory needs to address the tension between locality in dependencies and increasing distances between fillers and gaps). Preserving structural relations after adjunction is one of the crucial features of TAGs, since all relations are established at the level of elementary trees.

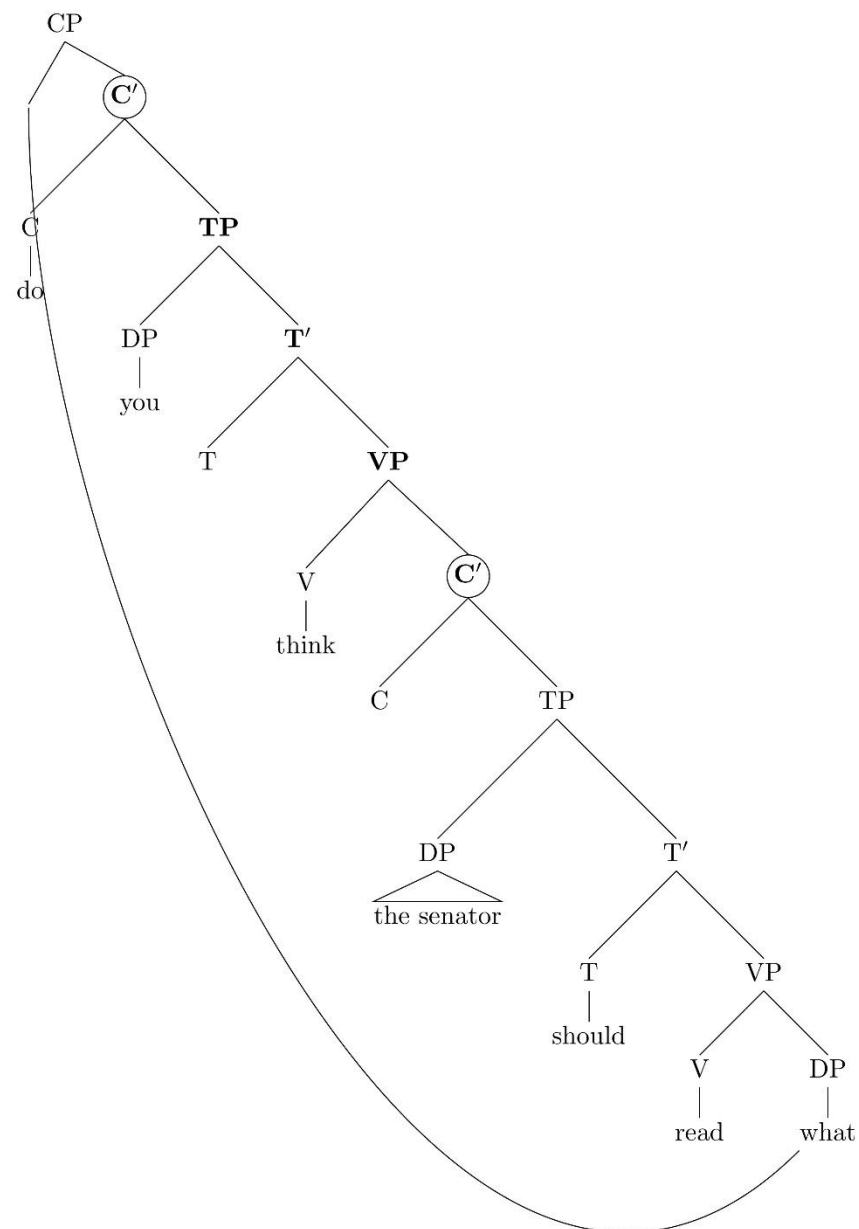


Figure 22. Derived tree for *wh*-interrogative featuring multidominance.

In traditional TAG terms [18], links are preserved under adjunction. However, whereas the definition of links in a TAG requires a multiplication of nodes such that a link is a relation between two nodes, one of which dominates an empty string, the system explored here satisfies the fundamental TAG hypothesis in a more economic way.

The main theoretical distinction between the system presented here, and the minimalist approach to displacement is that our structural descriptions are graphs connecting addresses (which correspond to *types*), whereas minimalist trees contain lexical *tokens*. An interesting consequence of this approach is that, because the same address can be called several times, there is no multiplication of nodes in instances of displacement: instead of defining *chains* of distinct terms, we define a *walk* on a graph: an ordered sequence of nodes and edges. This means that displacement rules will not replace or destroy existing syntactic relations but only create new ones on top of already existing relations: structural relations are *preserved* throughout the derivation. For example: consider the phrase marker corresponding to the string *John has read which book*. Here, *which book* receives theta-role and Case in a position internal to the VP; under standard assumptions, the direct object is

a sister to the V. The question that arises here is crucial: in order to obtain *which book has John read?*, do we need to disrupt the relation between *read* and *which book*? Under the trace theory, we actually do: *which book* is copied, replaced by a trace, and re-introduced in the derivation at a later stage. In terms of the *trace* theory of movement [75] (pp. 44–45):

“The movement of NP_i to position NP_j (where A and B are the contents of these nodes) in (30) yields (31) as a derived constituent structure.

(30) ... NP_j ... NP_i ...

| |
A B

(31) ... NP_i ... NP_i ...

| |
B *e*

In this view, NP_i and its contents are copied at position NP_j , deleting NP_j and A, and the identity element *e* is inserted as the contents of (in this case, the righthand) NP_i , deleting B under identity.”

Under standard minimalist assumptions, whereby reordering rules are reformulated in terms of Internal Merge, the steps for a *copy + re-merge* theory are the following [76,77]:

- a. Copy one of two independently merged phrases.
- b. Spell out the lower copy as trace.
- c. Merge the trace.
- d. Merge the higher copy (possibly in a separate derivation)”. ([78], p. 61)

“Independent operations of the *Copy + Merge* theory of movement:

- a. Copy.
- b. Merge.
- c. Form chain.
- d. Chain reduction”. ([79], p. 89)

Allegedly simpler systems have been proposed (e.g., [40,76] MERGE), but rarely implemented in the description of grammatical phenomena cross-linguistically.

In both the *trace* and *copy* theories of movement, there are *new* elements introduced in the syntactic derivation and thus *new* relations. In the trace version, the structural relation between *read* and *which book* is disrupted and replaced by a relation between *read* and a trace co-indexed with *which book*.

In a strict graph-theoretic interpretation of filler-gap dependencies, as implemented in our version of a lexicalised TAG, no substantive elements need to be added to representations or derivations: the work usually done by reordering transformations is done by adding edges in elementary trees. Edges in graph theory have formal status, unlike edges in diagrams of set theory-based syntax: the mechanisms to capture long-distance dependencies in our lexicalised proposal use only the tools made available to us by graph theory *plus* the proposal that nodes in structural representations are not lexical tokens but addresses to lexical types.

5. Conclusions

Generative grammar strives to provide accounts of two properties of natural languages: hierarchical structure and displacement. In contemporary minimalist analyses, both of these properties are handled by the same mechanism: Merge. Structure building, or “construal”, is the domain of External Merge (from the lexicon to the derivation), whereas structure mapping or “move” is the domain of Internal Merge (from the derivation to the derivation). However, there are two aspects that some, including ourselves, deem problematic: first, determining the size of local domains for the application of syntactic operations is left to designated nodes (so-called *phase heads*). This approach to locality is motivated by theory-internal conditions pertaining to Agree (such that phases are the domains where

feature checking and valuation operations take place), which have no semantic correlate. In this context, assigning compositional semantic interpretations to complex syntactic objects becomes an issue. Second, the way in which displacement is handled (the *copy theory of movement*) is based on strong set-theoretic commitments, which require the multiplication of nodes in reordering operations: for every element that moves, two nodes (at least) are created ([52] discusses some formal inconsistencies of the set-theoretic approach for the modelling of chain dependencies).

We would like to put forth the idea that an LTAG perspective could provide constructive solutions to these two issues. In terms of construal, elementary trees are defined in terms of a notion whose definition is independently necessary for a theory of grammar: *predication*. Having syntactic domains be the extended projection of lexical predicates also allows us to address the issue of cross-linguistic variation from a novel perspective: as we have shown with evidence from English, Spanish and German, what counts as a lexical predicate (and thus what can anchor an elementary tree) is a locus of variation. The theory of variation is thus directly related to the theory of syntactic construal and the syntax-semantics interface.

In terms of displacement, we have proposed that, if syntactic terms in derivations are defined as addresses (that stand for the intensional logic translation of categorematic basic expressions of the language), it is possible to define structural descriptions as directed graphs with a total order where nodes can be visited more than once in a path. The result is a set of elementary graphs (which are technically not trees, since in graph theory, trees do not allow for loops) where all dependencies are established at the local level. Long-distance dependencies are the result of adjoining auxiliary graphs *after* all local dependencies have been defined within an elementary graph. In this context, the role that in generative grammar is played by reordering transformations is actually reduced to *construal + graph composition* operations. Because relations in elementary graphs are preserved under adjunction, “transformations” are homomorphic mappings: they preserve relations within a selected structure and do not create new nodes or relations. Creating new edges and maximising connectivity within elementary graphs as opposed to creating new nodes (copies), which is made available to us after giving up strict set-theoretic commitments and allowing syntactic terms to be uniquely identifying addresses, has at least three advantages (see also [49]):

- The *inclusiveness condition*, in its strict interpretation, is always respected.
- If elementary trees are graphs defined in a workspace, then the workspace is never expanded ([40,76]).
- Problems pertaining to the distinction between *copies* and *repetitions* do not arise: copies are formalised as multiple visits to the same node in a strictly ordered walk through a graph, whereas repetitions are instances of distinct nodes (with the same morpho-phonological exponent) in a walk through a graph.

The approach proposed here satisfies the conditions for a minimalist theory of grammar: eliminate superfluous elements in representations as well as superfluous steps in derivations.

Author Contributions: Conceptualisation, D.G.K. and A.P.; Formal analysis, D.G.K. and A.P.; Investigation, D.G.K. and A.P. Although, the overall paper is the result of joint work, for the concerns of the Italian Academia, D.G.K. takes responsibility of Sections 1, 3.1 and 4, and the conclusions and A.P., of Sections 2 and 3.2. All authors have read and agreed to the published version of the manuscript.

Funding: No funding received.

Acknowledgments: We would like to express our gratitude towards Peter Kosta for his invitation to participate in this Special Issue, and two anonymous *Philosophies* reviewers for their useful comments. The usual disclaimers apply

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NP = Noun Phrase; VP = Verb Phrase; CP = Complementiser Phrase; PP = Prepositional Phrase; AdvP = Adverb Phrase; DP = Determiner Phrase; ACC = Accusative Case; DAT = Dative Case; COND = Conditional; INF = Infinitive; PART = Participle; PERF = Perfective aspect; AUX = Auxiliary verb; SG = Singular Number; CL = Clitic; FUT = Future tense; PST = Past Tense; SUBJ = Subjunctive mood; S = Sentence

Notes

- 1 As a matter of fact, [19] points out that elementary trees need not coincide with phases, as they can be larger than a phase; for instance, the extended projection of a verbal head might be as large as a CP [24]. Perhaps more importantly for this paper, elementary trees can be *smaller* than phases.
- 2 This system, as it is based on two binary dimensions, predicts the existence of four kinds of syntactic categories:
 - i. Lexical predicates (V, A, P, nominalisations, lexical auxiliaries).
 - ii. Lexical non-predicates (proper N, non-nominalised common N).
 - iii. Non-lexical predicates (functional auxiliaries, quantifiers).
 - iv. Non-lexical, non-predicates (Complementisers).
- 3 Strictly speaking, we also want to define a unit where there is no syntactic position between *ha* and *debido*, in the light of paradigms such as the following [77,80]:

•	¿Qué	ha	el	senador	debido	leer?
	What	has.3SG-AUX-PRF	the	senator	must-PART	read-INF?
•	¿Qué	ha	debido	leer	el	senador?
	What	has.3SG-AUX-PRF	must-PART	read-INF	the	senator?

“What has the senator had to read?”
- 4 A reviewer objects that restricting scope relations to elementary trees may be problematic, for example, in NPI licensing in multi-clausal structures, such as:
 - (i) I don’t believe that they ate anything.
 However, unlike our Spanish example, it is not derived by adjunction; it is derived by substitution:
 - (ii) [I don’t believe [S]]
 [s they ate anything].

The resulting structural description thus allows for scope relations to be read off a monotonic syntactic configuration (regardless of whether NEG originates in the embedded clause and NEG-raises or not). For the determination of scope relations it is essential to consider whether the tree composition operation that has applied is *substitution* (which yields tail-recursive structures in English clausal complementation) or *adjunction*.
- 5 Importantly, this does not mean that TAGs are incompatible with successive cyclic movement. If there are occurrences of a *wh* in intermediate positions, then we need to assume that there is a dependency to be defined within an adjoined elementary tree. English does not seem to provide evidence of intermediate landing sites in cases of unbounded *wh*-movement, but other languages do (e.g., by spelling-out copies or dummy *wh*-pronouns in intermediate locations). What this system allows us to dispense with are tokens of a syntactic object required only by formal reasons, without there being any semantic or phonological effect (this may, of course, vary across languages as it is part of the definition of elementary trees). Frank [21], p. 237, allows objects to move to the edge of an elementary tree (his system allows for singulary transformations) and then, from there, it may be further reordered by operations at the target of adjunction. Detailed discussion about aspects of successive cyclicity in LTAGs was provided in [19], e.g., scope reconstruction.
- 6 This proposal also has parallels in computer languages. In programming languages, such as Python, expressions are assigned addresses, which are accessed in the execution of a program by means of variables. It is even possible to access the memory address by means of a specific function, *id*(. . .). It is also possible to create local variables, which get created every time a function is called and erased when the function returns an output; this kind of operation would be analogous to the procedure to create copies in movement operations (since the copy is created and stored only temporarily, to be merged later in order to satisfy some featural requirement). In this case, we need two “spaces”: the “shell” (where the program is executed) and the list of objects in memory (which contains the address, type and value assigned to each variable).

References

1. Bach, E. *An Introduction to Transformational Grammars*; Holt, Rinehart & Winston: New York, NY, USA, 1964.
2. Ross, J.R. Constraints on Variables in Syntax. Ph.D. Thesis, MIT, Cambridge, MA, USA, 1967.
3. Chomsky, N.; Halle, M. *The Sound Pattern of English*; Harper & Row: New York, NY, USA, 1968.
4. Chomsky, N. *Barriers*; MIT Press: Cambridge, MA, USA, 1986.

5. Larson, R. On the double object construction. *Linguist. Inq.* **1988**, *19*, 335–391.
6. Pollock, J.-Y. Verb Movement, Universal Grammar, and the Structure of IP. *Linguist. Inq.* **1989**, *20*, 365–424.
7. Chomsky, N. Minimalist Inquiries: The Framework. In *Step by Step—Essays in Minimalist Syntax in Honor of Howard Lasnik*; Martin, R., Michaels, D., Uriagereka, J., Eds.; MIT Press: Cambridge, MA, USA, 2000; pp. 89–155.
8. Cinque, G. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*; OUP: Oxford, UK, 1999.
9. Cinque, G. ‘Restructuring’ and Functional Structure. In *Restructuring and Functional Heads. The Cartography of Syntactic Structures*; Cinque, G., Ed.; OUP: Oxford, UK, 2004; Volume 4, pp. 132–192.
10. Cinque, G.; Rizzi, L. The Cartography of Syntactic Structures. *StIL*. 2008, Volume 2. Available online: http://www.ciscl.unisi.it/doc/doc_pub/cinque-rizzi2008-The_cartography_of_Syntactic_Structures.pdf (accessed on 13 June 2020).
11. Cinque, G.; Rizzi, L. Rizzi, Functional Categories and Syntactic Theory. *Annu. Rev. Linguist.* **2016**, *2*, 139–163.
12. Chomsky, N. Derivation by Phase. In *Ken Hale: A Life in Language*; Kenstowicz, M., Ed.; MIT Press: Cambridge, MA, USA, 2001; pp. 1–52.
13. Richards, M. Deriving the Edge: What’s in a Phase? *Syntax* **2011**, *14*, 74–95. [[CrossRef](#)]
14. Boeckx, C. Phases beyond explanatory adequacy. In *Phases: Developing the Framework*; Gallego, A., Ed.; Mouton de Gruyter: Berlin, Germany, 2012; pp. 45–66.
15. Chomsky, N. The Logical Structure of Linguistic Theory. Mimeographed, MIT. 1955. Available online: http://alpha-leonis.lids.mit.edu/wordpress/wp-content/uploads/2014/07/chomsky_LSLT55.pdf (accessed on 30 July 2018).
16. Chomsky, N. Transformational Analysis. Ph.D. Thesis, University of Pennsylvania, Philadelphia, PA, USA, 1955.
17. Joshi, A.; Levy, L.; Takahashi, M. Tree adjunct grammars. *J. Comput. Syst. Sci.* **1975**, *10*, 136–163. [[CrossRef](#)]
18. Joshi, A.K. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions. In *Natural Language Parsing*; Dowty, D., Karttunen, L., Zwicky, A., Eds.; CUP: Cambridge, MA, USA, 1985; pp. 206–250.
19. Frank, R. *Phrase Structure Composition and Syntactic Dependencies*; MIT Press: Cambridge, MA, USA, 2002.
20. Stabler, E. The epicenter of linguistic behavior. In *Language down the Garden Path*; Sanz, M., Laka, I., Tanenhaus, M., Eds.; OUP: Oxford, UK, 2013; pp. 316–323.
21. Frank, R. Tree adjoining grammar. In *The Cambridge Handbook of Generative Syntax*; den Dikken, M., Ed.; CUP: Cambridge, MA, USA, 2013; pp. 226–261.
22. Frank, R. Phase theory and Tree Adjoining Grammar. *Lingua* **2006**, *116*, 145–202. [[CrossRef](#)]
23. Grimshaw, J. Locality and Extended Projection. In *Lexical Specification and Insertion*; Coopmans, P., Everaert, M., Grimshaw, J., Eds.; John Benjamins: Amsterdam, The Netherlands, 2000; pp. 115–133.
24. Abney, S.P. The English Noun Phrase in Its Sentential Aspect. Ph.D. Thesis, MIT, Cambridge, MA, USA, 1987.
25. Frank, R.; Kulick, S.; Vijay-Shanker, K. Monotonic c-command: A new perspective on tree adjoining grammars. *Grammars* **2002**, *3*, 151–173. [[CrossRef](#)]
26. Gazdar, G.; Klein, E.; Pullum, G.K.; Sag, I.A. *Generalised Phrase Structure Grammar*; Blackwell: Oxford, UK, 1985.
27. Bresnan, J. *Lexical Functional Syntax*, 1st ed.; Wiley Blackwell: Oxford, UK, 2001.
28. XTAG Group. *A Lexicalized TAG for English*; Technical Report; University of Pennsylvania: Pennsylvania, PA, USA, 2001; Available online: https://repository.upenn.edu/cgi/viewcontent.cgi?article=1020&context=ircs_reports (accessed on 13 June 2020).
29. Joshi, A.K.; Schabes, Y. Tree-Adjoining Grammars and Lexicalized Grammars. *Tech. Rep. (CIS)*. 1991, p. 445. Available online: http://repository.upenn.edu/cis_reports/445 (accessed on 22 April 2019).
30. Hegarty, M. *Deriving Clausal Structure in Tree Adjoining Grammar*; University of Pennsylvania: Pennsylvania, PA, USA, 1993.
31. Rambow, O. Mobile Heads and Strict Lexicalization. Master’s Thesis, University of Pennsylvania, Pennsylvania, PA, USA, 1993.
32. Krivochen, D.G.; García Fernández, L. On the position of subjects in Spanish periphrases: Subjecthood left and right. *Boreal. Int. J. Hisp. Linguist.* **2019**, *8*, 1–33. [[CrossRef](#)]
33. Krivochen, D.G.; García Fernández, L. Variability in syntactic-semantic cycles: Evidence from auxiliary chains. In *Interface-Driven Phenomena in Spanish: Essays in Honor of Javier Gutiérrez-Rexach*; González-Rivera, M., Sessarego, S., Eds.; Routledge: London, UK, 2020; pp. 145–168.
34. Bravo, A.; García Fernández, L.; Krivochen, D.G. On Auxiliary Chains: Auxiliaries at the Syntax-Semantics Interface. *Borealis* **2015**, *4*, 71–101. [[CrossRef](#)]
35. Huddleston, R.; Pullum, G. *The Cambridge Grammar of the English Language*; CUP: Cambridge, MA, USA, 2002.
36. May, R. *Logical Form: Its Structure and Derivation*; MIT Press: Cambridge, MA, USA, 1985.
37. Fabb, N. Three squibs on auxiliaries. In *MIT Working Papers in Linguistics, Volume 5. Papers in Grammatical Theory*; Haik, I.Y., Massam, D., Eds.; Department of Linguistics and Philosophy: Cambridge, MA, USA, 1983; pp. 104–120.
38. Epstein, S.D. Un-Principled Syntax and the Derivation of Syntactic Relations. In *Working Minimalism*; Epstein, S., Hornstein, N., Eds.; MIT Press: Cambridge, MA, USA, 1999; pp. 317–345.
39. Chomsky, N. *The Minimalist Program*; MIT Press: Cambridge, MA, USA, 1995.
40. Chomsky, N. The UCLA Lectures. University of Arizona, Tucson, AZ, USA. 2020. Available online: <https://ling.auf.net/lingbuzz/005485> (accessed on 11 July 2020).
41. Evers, A. *The Transformational Cycle in Dutch and Germany*; Indiana University Linguistics Club: Indiana, IN, USA, 1975.
42. Den Besten, H.; Edmonson, J. The verbal complex in continental West Germanic. In *On the Formal Syntax of the Westgermania*; Abraham, W., Ed.; John Benjamins: Amsterdam, The Netherlands, 1983; pp. 155–216.

43. Haegeman, L.; van Riemsdijk, H. Verb Projection Raising, Scope, and the Typology of Rules Affecting Verbs. *Linguist. Inq.* **1986**, *17*, 417–466.
44. Wurmbrand, S. Verb clusters, verb raising, and restructuring. In *The Blackwell Companion to Syntax*; Everaert, M., van Riemsdijk, H., Eds.; Blackwell: Oxford, UK, 2017.
45. Kroch, A.; Santorini, B. The derived constituent structure of the West-Germanic verb-raising construction. In *Principles and Parameters in Comparative Grammar*; Freidin, R., Ed.; MIT Press: Cambridge, MA, USA, 1991; pp. 269–338.
46. Altmann, G. Prolegomena to Menzerath's law. *Glottometrika* **1980**, *2*, 1–10.
47. Adger, D.; Svenonius, P. Features in minimalist syntax. In *The Oxford handbook of linguistic Minimalism*; Boeckx, C., Ed.; OUP: Oxford, UK, 2011; pp. 27–51.
48. Panagiotidis, P. Towards a (Minimalist) Theory of Features. University of Cyprus, Nicosia, Cyprus. 2021. Available online: <https://ling.auf.net/lingbuzz/005615> (accessed on 15 February 2021).
49. Krivochen, D.G. Copies and Tokens: Displacement Revisited. *Studia Linguist.* **2015**, *70*, 250–296. [CrossRef]
50. Gärtner, H.-M. *Generalized Transformations and Beyond. Reflections on Minimalist Syntax*; Akademie Verlag: Berlin, Germany, 2002.
51. Collins, C.; Groat, E. Distinguishing Copies and Repetitions. 2018. Available online: <http://ling.auf.net/lingbuzz/003809> (accessed on 7 May 2021).
52. Gärtner, H.-M. Copies from 'Standard Set Theory'. 2020. Available online: <https://ling.auf.net/lingbuzz/005942> (accessed on 10 September 2020).
53. Sampson, G. The Single Mother Condition. *J. Linguist.* **1975**, *11*, 1–11. [CrossRef]
54. Citko, B. On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge. *Linguist. Inq.* **2005**, *36*, 475–496. [CrossRef]
55. Levine, R.D. Right node (non-)raising. *Linguist. Inq.* **1985**, *16*, 492–497.
56. McCawley, J.D. Parentheticals and Discontinuous Constituent Structure. *Linguist. Inq.* **1982**, *13*, 91–106.
57. Johnson, K. Toward a Multidominant Theory of Movement. Lectures Presented at ACTL, University College. June 2016. Available online: https://people.umass.edu/kbj/homepage/Content/Multi_Movement.pdf (accessed on 16 August 2020).
58. Kroch, A.; Joshi, A.K. *The Linguistic Relevance of Tree Adjoining Grammar*; Technical report MS-CS-85-16; Department of Computer and Information Sciences, University of Pennsylvania: Philadelphia, PA, USA, 1985.
59. Peter, S.; Ritchie, R. *Phrase-Linking Grammar*. University of Texas at Austin: Austin, TX, USA, 1981.
60. Postal, P. On some rules that are not successive-cyclic. *Linguist. Inq.* **1972**, *3*, 211–222.
61. Boeckx, C.; Grohmann, K.K. Barriers and Phases: Forward to the Past? Presented in *Tools in Linguistic Theory 2004 (TiLT)* Budapest (16–18 May 2004). Available online: http://people.umass.edu/roeper/711-05/Boexce%20barriers+phases%20tilt_bg_ho.pdf (accessed on 26 April 2018).
62. Sarkar, A.; Joshi, A.K. *Handling Coordination in a Tree Adjoining Grammar*; Technical Report; University of Pennsylvania: Philadelphia, PA, USA, 1997.
63. Sutherland, W. *Introduction to Metric and Topological Spaces*, 2nd ed.; OUP: Oxford, UK, 2009.
64. Searcoid, M. *Metric Spaces*; Springer: Berlin/Heidelberg, Germany, 2007.
65. Krivochen, D.G. Towards a Theory of Syntactic Workspaces: Neighbourhoods and Distances in a Lexicalised Grammar. 2021. Available online: <https://ling.auf.net/lingbuzz/005689> (accessed on 6 May 2021).
66. Krivochen, D.G. Syntax on the Edge: A Graph-Theoretic Analysis of English (and Spanish) Sentence Structure. Under Review. 2018. Available online: <https://ling.auf.net/lingbuzz/003842> (accessed on 6 May 2021).
67. Gorn, S. Handling the growth by definition of mechanical languages. In Proceedings of the Spring Joint Computer Conference, Atlantic City, NJ, USA, 18–20 April 1967; pp. 213–224. [CrossRef]
68. MacFarlane, J. Logical constants. In *The Stanford Encyclopedia of Philosophy*. 2017. Available online: <http://plato.stanford.edu/archives/fall2015/entries/logical-constants/> (accessed on 18 March 2021).
69. Schmerling, S.F. *Sound and Grammar: Towards a Neo-Sapirian Theory of Language*; Brill: Leiden, The Netherlands, 2018.
70. García Fernández, L.; Krivochen, D.G.; Martín Gómez, F. Los elementos intermedios en las perífrasis verbales. *Lingüística Española Actual* **2021**.
71. Karttunen, L. Syntax and semantics of questions. *Linguist. Philos.* **1977**, *1*, 3–44. [CrossRef]
72. Hamblin, C.L. Questions in Montague English. *Found. Lang.* **1973**, *10*, 41–53.
73. Montague, R. The proper treatment of quantification in ordinary English. In *Approaches to Natural Language*; Hintikka, J., Moravcsik, J., Suppes, P., Eds.; Reidel: Dordrecht, The Netherlands, 1973; pp. 221–242.
74. Reinhart, T. Wh-in situ in the framework of the minimalist program. *Nat. Lang. Semant.* **1998**, *6*, 29–56. [CrossRef]
75. Fiengo, R. On Trace Theory. *Linguist. Inq.* **1977**, *8*, 35–61.
76. Chomsky, N. Some Puzzling Foundational Issues: The Reading Program. *Catalan J. Linguist.* **2019**, 263–285. [CrossRef]
77. Torrego, E. On Inversion in Spanish and Some of Its Effects. *Linguist. Inq.* **1984**, *15*, 103–129.
78. Uriagereka, J. Multiple Spell-Out. In *Derivations: Exploring the Dynamics of Syntax*; Routledge: London, UK, 2002; pp. 45–65.
79. Nunes, J. *Linearization of Chains and Sideways Movement*; MIT Press: Cambridge, MA, USA, 2004.
80. García Fernández, L.; Krivochen, D.G. Dependencias no locales y cadenas de verbos auxiliares. *Verba* **2019**, *46*, 207–244. [CrossRef]