John Benjamins Publishing Company

This is a contribution from Evolutionary Linguistic Theory 3:2
© 2021. John Benjamins Publishing Company

This electronic file may not be altered in any way. The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's/s' institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu.

Please see our rights policy on *https://benjamins.com/content/customers/rights*
For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: *www.copyright.com*).

Please contact *rights@benjamins.nl* or consult our website: *www.benjamins.com*

Mixed computation

Grammar up and down the Chomsky hierarchy

Diego Gabriel Krivochen
University of Oxford

Proof-theoretic models of grammar are based on the view that an explicit characterization of a language comes in the form of the recursive enumeration of strings in that language. That recursive enumeration is carried out by a procedure which strongly generates a set of structural descriptions Σ and weakly generates a set of strings S; a grammar is thus a function that pairs an element of Σ with elements of S. Structural descriptions are obtained by means of Context-Free phrase structure rules or via recursive combinatorics and structure is assumed to be *uniform*: binary branching trees all the way down. In this work we will analyse natural language constructions for which such a rigid conception of phrase structure is descriptively inadequate and propose a solution for the problem of phrase structure grammars assigning too much or too little structure to natural language strings: we propose that the grammar can oscillate between levels of computational complexity in local domains, which correspond to elementary trees in a lexicalised Tree Adjoining Grammar.

Keywords: syntax, derivations, mixed computation, tree adjoining grammars, compositionality

1. Introduction

Contemporary generative grammar (since at least Chomsky, 1986; Kayne, 1984) operates under the assumption that syntactic structure is universally based on a single kind of template: single-rooted, binary-branching, and endocentric structures. The structural variety that characterised earlier versions of the theory, where flat structures were combined with binary branching (Ross, 1967; Emonds, 1976), was replaced by a single template (X-bar theory and, later, Merge) that converged in a uniform Context-Free (CF) description for the base component of a generative grammar, with some additional power made available through transformations and operations over features (Move and Agree).

https://doi.org/10.1075/elt.00034.kri
Evolutionary Linguistic Theory 3:2 (2021), pp. 215–244. ISSN 2589-1588 | E-ISSN 2589-1596
© John Benjamins Publishing Company

A generative grammar in the transformational tradition consists of a set of context-free rules (which at most can refer to the last line of a derivation) plus a set of rules that may access the 'derivational history' of a sequence to various effects (move, copy, delete, and substitute). That architecture entails assumptions about the strong generative capacity of natural language grammars. Empirically, the current mainstream generative grammar (MGG) view implies *structural uniformity*: the computational complexity of linguistic dependencies is invariable. Fukui & Narita (2014: 20) express this view clearly:

> considerations of binding, quantifier scope, coordination, and various other phenomena seem to lend support to the **universal binary branching** hypothesis. [...] it is likely that theories of labeling and linearization play major roles in this binarity restriction. [our highlighting]

Crucially, the existence of empirical phenomena that are indeed amenable to a binary-branching analysis does *not* preclude the existence of phenomena for which binary-branching is inadequate. However, binary branching as a model of structural uniformity does imply rejecting *a priori* the possibility that other configurations are available.

Let us analyse some consequences of *structural uniformity* in the context of the generative power of grammars: strong generative capacity refers to the structural descriptions enumerated by a grammar, and weak generative capacity to the actual strings it can produce (Chomsky, 1965: 60). Consider now this quotation from Stabler (2013: 318):

> computational consensus was identified by Joshi (1985) in his hypothesis that human languages are both strongly and weakly mildly context sensitive (MCS)

Stabler identifies theories that verify this consensus: from strictly CFGs, Tree Adjoining Grammars (TAGs; Joshi, 1985), Combinatory Categorial Grammars (Steedman, 2019) to several kinds of Minimalist Grammars (Stabler, 1997). What exactly does this entail for the development of descriptively adequate theories of natural language grammars? Stabler's position is clear:

> Joshi's original definition of MCS [Mildly Context-Sensitive] grammars was partly informal, so there are now various precise versions of his claim. One is that human languages are defined by tree adjoining grammars (TAGs) or closely related grammars, and another [...] is that human language are definable by the more expressive (set local) multi-component TAGs or closely related grammars.

From the perspective of descriptively adequate grammars, it is legitimate to ask whether recognising that a specific construction in a specific language displays, e.g., restricted CS dependencies means that *all* constructions in that language

© 2021. John Benjamins Publishing Company
All rights reserved

must be assigned a structural description of that computational power. In other words, whether structure uniformity determines that, once an expression in L displays a certain computational complexity, all expressions in L (or universally) automatically do. In Section 2 we provide counterexamples to a structurally uniform approach to syntax. Section 3 develops a general approach to these issues, and Section 4 extends the framework.

2. Empirical problems

We require of a descriptively adequate theory of grammar that it be able to assign a structural description that represents semantic dependencies between elements in a string. This is essential, since it may be here that one of the crucial differences between natural and formal languages reside. In a formal language such as first order logic, the so-called *unique readability theorem* holds: well-formed formulae are defined in such a way that there is only one way to construct each formula given an alphabet of constants and Boolean connectives and set of formation rules, and only one way to read it (Epstein, 2011: 8–9). It may be apparent to the reader that natural language does not behave like that: humans use strings of words in ways that do not always obey strict compositionality nor are unambiguous (irony, plays on words, jokes…). But we do not need to consider *how language is used* to reject unique readability as a condition on natural language sentences: it seems to be a property of *how the grammatical system works*. Consider the following strings:

(1) Wakanda is a big small country

(2) Gandalf is an old old man

(3) That's fake fake news

To see why (1–3) pose a problem for theories of syntax based on structural uniformity, let us consider what kind of structural description they would assign to our strings. There has been remarkably little attention paid to iteration in MGG, although several analyses of the syntax of adjective stacking are available (Alexiadou, 2014). Adjectives have been claimed to be adjuncts or specifiers to NP/DP (and therefore full phrases; Jackendoff, 1977; Svenonius, 1994) or heads (Abney, 1987; Cinque, 1994), always with the focus set on the ordering restrictions between different semantic kinds of adjectives as well as their relative position with respect to the noun they modify (pre-nominal vs. post-nominal distribution).

© 2021. John Benjamins Publishing Company
All rights reserved

In Cartographic approaches, each adjective must head its own functional projection within the NP/DP; each modifier is placed in a unique head position (Cinque, 2010; Scott, 2002; Bortolotto, 2016) or as specifiers of functional phrases FP (Aljović, 2010). Cartography provides a rich set of functional projections where adjectives of different semantic classes are located. These functional categories are arranged in a strictly binary-branching {Head, XP} fashion, following Kayne (1994) and Chomsky (2013). An approximation to the functional hierarchy for adjectives is given in (4) (see Scott, 2002: 114 for a refined version):

(4) $[_{DP}$ D° $[Adj_{poss}$ $[Adj_{card}$ $[Adj_{ord}$ $[Adj_{qual}$ $[Adj_{size}$ $[Adj_{shape}$ $[Adj_{color}$ $[Adj_{nation}$ $[_{NP}$ N°]]]]]]]]]]

Scott (2002: 112) proposes that '*an adjective generated in the hierarchy relates directly to the semantic reading it receives*', and considers only readings of adjective iteration where the adjectives contribute distinct meanings: in *a good good typist*, he distinguishes the first *good* as pertaining to 'morality' and the second as pertaining to 'manner'; therefore, he assumed a stacked structure of the kind $[_{Subj.CommentP}$ *good* $[_{MannerP}$ *good* $[_{NP}$ *typist*]]]; similarly, in *an old old Etonian*, '*this fragment can only receive the interpretation an old (= in age) old (= former) Etonian*'. Scott's position is representative of the cartographic view, heavily influenced by the Chomsky-Kayne approach to phrase structure, which is committed to structural uniformity. It is, however, surprising that readings like 'a very good typist' or 'a very old Etonian' are not even considered.[1]

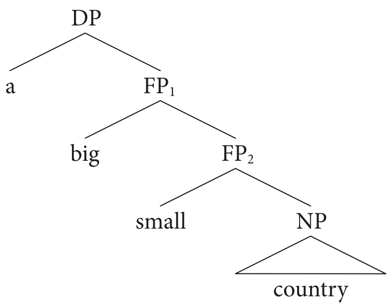
Going back to our sentences, consider first (1). It only seems to allow for a reading in which one adjective has scope over the other, such that the corresponding interpretation is roughly 'a country that is big for a small country'. We are not concerned with fine-grained issues pertaining to the semantics of adjectives in this paper, but rather with the fact that a PSG would assign (1) a structural description in which, as pointed out above, one adjective c-commands the other, thus having scope over it (Ladusaw, 1980; May, 1985). Using Cinque's (1994) FPs as proxies for whatever labels are assumed to be in play, the structural representation would look like (5):

1. Cinque (2010: 118) similarly neglects the possibility of intensive reduplication with adjective iteration:

> If both [adjectives] are prenominal, with some degree of cumbersomeness (*la vecchia vecchia bicicletta di Gianni* 'Gianni's old old bicycle'), the first *vecchia* is **necessarily** interpreted as POSS-modifying and the second *vecchia* as N-modifying, just as in English. [our highlighting]

© 2021. John Benjamins Publishing Company
All rights reserved

(5)

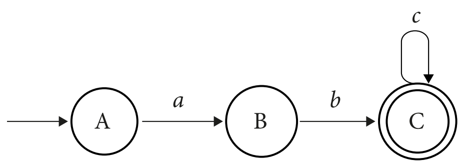


What kind of computational device can generate the structure in (5)? To an extent, the answer depends on the status of adjectives in X-bar theoretic terms: heads or specifiers. The least powerful grammars in the Chomsky Hierarchy, finite-state (FS) grammars, allow for rules of the form A → *a*B, which means that it is possible to produce the tree in (5) by using only FS rules if adjectives are FP heads, since:

> a given finite-state language L can be generated either by a psg [Phrase Structure Grammar] containing only left-linear rules: Z → aY, Z → a, or by a psg containing only right-linear rules: Z → Ya, Z → a, and a psg containing either only left-linear rules or only right-linear rules will generate a finite state language
>
> (Greibach, 1965: 44)

In other words, the following are strongly equivalent:

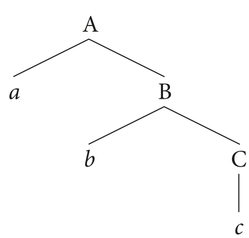
(6) A → *a*B
B → *b*C
C → *c*



$\delta(q_1, A, a) = (q_2, B)$

$\delta(q_2, B, b) = (q_3, C)$

$\delta(q_3, C, c) = (q_3, C)$

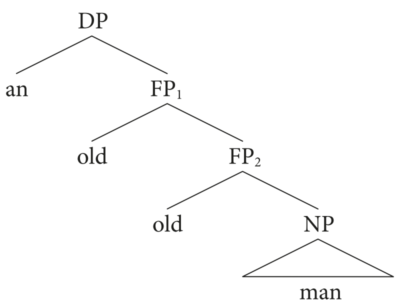


A consistently left-branching or consistently right-branching PSG that only allows for production rules involving a terminal and a non-terminal generates an FS lan-

© 2021. John Benjamins Publishing Company
All rights reserved

guage (see also Uriagereka, 2012: 53).[2] If, however, adjectives are taken to be *specifiers* of FP (thus, maximal projections), then the grammar must contain also rules of the form A → BC, with A, B, C nonterminals. Such a grammar is strictly context-free. In either case, however, there is a point to be made: the semantic scope of *big* must be *small country*, in order to get the proper semantic interpretation. That means that, under more-or-less strong compositional assumptions (Bach, 1976; Jacobson, 2012) the semantic contribution of *big* must apply to the output of a syntactic rule generating an expression *small country*. A *big small country* is *not* a country that is small and big.

The second case contrasts with the first in an interesting respect: the only possible interpretation for this case of adjective iteration is *intensive reduplication*; that is, 'a very old man'. Schmerling (2018: 3) observes that '*the semantic value of an NP with multiple occurrences of an adjective, say* old, *does not increase as the number of instances of* old *increases*.', and attributes this to the finite state syntax of total reduplication in these instances: in *old*$_1$ *old*$_2$ *man* it would be inaccurate to say that *old*$_1$ takes as its input the semantic value of *old*$_2$ *man*; rather, *man* is modified by the semantic value of the sequence {*old*$_1$ *old*$_2$, …*old*$_n$}. The direct consequence of this reasoning is that a structural description like (5), as in (7), would be flawed insofar as it would be unable to adequately represent the relations between the adjectives and the semantic representation assigned to (2):

(7)


2. A similar argument was used in Reich (1969: 835), who enriches FS devices with a set of Boolean connectives; he replaces rewrite grammars with circuit-like transition graphs which instantiate strictly right-branching or strictly left-branching trees (but no symmetric bifurcation). Reich, however, is also committed to structural uniformity:

> English (and, I suspect, all languages) can, in fact, be described by a finite state device, namely a network of relationships, where each relationship itself a finite state device.
> (Reich, 1969: 834)

As observed in Shieber (1985), the view that natural languages are regular entails that they are sets of finite strings, with embedding and cross-serial dependencies requiring fixed bounds.

© 2021. John Benjamins Publishing Company
All rights reserved

In (7), one instance of *old* takes a constituent *old man* as its complement and takes scope over it, which does not adequately represent the semantics of the NP. This is unavoidable in a system set up as in Cartography or MGG, given structural uniformity. This is not a completely novel point: the inadequacy of phrase structure representations for strings like (2) has been recognised since at least Chomsky & Miller (1963):

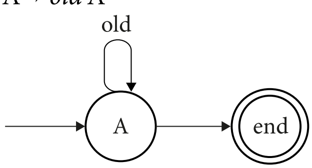
> a constituent-structure grammar necessarily imposes too rich an analysis on sentences because of features inherent in the way P-markers are defined for such sentences. (Chomsky & Miller, 1963: 297–298)

Following up on Chomsky & Miller (1963) and Postal (1964); Uriagereka (2008), and Lasnik (2011) acknowledge the problem imposing 'too much structure' on structural descriptions for strings if one moves up uniformly in the Chomsky Hierarchy ('FSGs are inadequate for some strings, then we proceed to CSGs; these also have limitations for some strings, thus we go further up...'):

> In a manner of speaking, what we really want to do is move down the [Chomsky] hierarchy. Finite-state Markov processes give flat objects, as they impose no structure. But that is not quite the answer either. (Lasnik, 2011: 361)

While it may be too strong to claim that FS representations 'impose no structure' (since an FS language may be generated by uniformly right- or left-branching phrase structure rules), it seems accurate to want to assign an FS representation to instances of intensive reduplication in English, most of all considering that, although increasing the number of iterations does not modify the semantic value of the adjective beyond perhaps intensifying the degree to which a property holds for an entity (such that *old old old old* also means 'very old'[3]) it is indeed possible to multiply the occurrences of an adjective beyond two. In this context, what would an adequate structural description for (2) look like? We would represent the adjective iteration in (2) as a loop in an FSA:

(8) A → *old* A



3. The meaning of intensive reduplication is reminiscent of the "rhetorical accent" identified in Newman (1946) and could have also been obtained by means of vowel lengthening. We thank Susan F. Schmerling for calling this reference to our attention.

© 2021. John Benjamins Publishing Company
All rights reserved

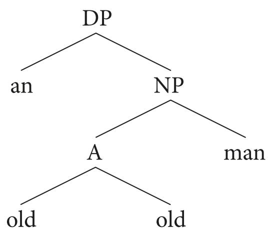
Take (8) to be a representation of a fragment of an FS recogniser: given a string, we ask whether that string belongs to the language defined by a regular grammar. Then, rule (8) means 'if you are in state A and read input *old*, proceed to state A'. But this representation would be adequate only for the iterative part: whereas it is true that there is no hierarchy to be found among the adjectives (there is strict *parataxis* according to any syntactic or semantic test, and this iteration is closed under Kleene star; see Karlsson, 2010), there *is* a hierarchical relation between the adjectives and the noun, such that the sequence of iterated adjectives modifies the noun. And here is where a pure FS structure is insufficient (as noted by Lasnik): if the sequence of adjectives has the properties of a non-terminal symbol (if it is an AP), then either as an NP specifier or as an adjunct there is derivational line that contains a sequence of non-terminal symbols: AP and NP or AP and N′. In either case, a structure that is beyond FS power.

However, we cannot say that *all sequences of adjectives* or even all instances of repeated adjectives have this property: it would be inadequate to assign a 'flat' structure to (1) (we will come back to (3) below), and a theory of grammar needs to be able to account for this. In this sense, Chomsky & Miller's (1963: 298) proposed (but rejected) solution to the problem of 'too much structure' for adjectives in predicative position would also be inadequate:

(9) Predicate → Adj^n and Adj ($n \geq 1$)

The reason is not (only), as Chomsky & Miller claim, that there are '*many difficulties involved in formulating this notion so that descriptive adequacy may be maintained*', but rather that the combination of (9) with the assumption of structural uniformity creates empirical problems. In what pertains to the expressive power of the grammar, the intersection of a CF language with a FS language is always a CF language. However, for purposes of grammatical description, this may be taken to mean that the base component of a transformational grammar is *at most* CF, not that the structural description assigned to every single sub-string is strictly CF (since, as we have seen, that adds unnecessary additional structure). What would an alternative structural description look like? We propose a diagram like (10):

(10)



© 2021. John Benjamins Publishing Company
All rights reserved

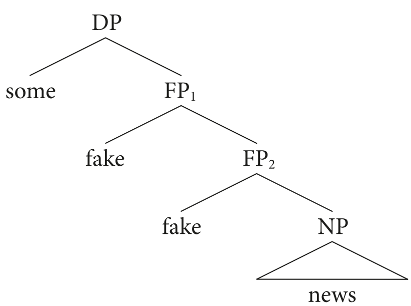
Our point is not notational: however it is drawn, an empirically adequate structural representation for (2) simply cannot assign the sequence of reduplicated *old* internal hierarchical structure. What we claim, echoing points made in Lasnik (2011) and Uriagereka (2008) is that *an old old man* is the result of combining a CF structure with a FS one in a grammatically relevant unit (an NP/DP), and that for purposes of grammatical analysis, it must be possible to express the distinction between these two sub-structures.

We now turn to (3), *that's fake fake news*. This example is particularly interesting because there are *two* possible readings, which correspond to two different segmentations:

(11) a. [fake [fake news]] (i.e., truthful news, or at least not fake news)
b. [fake fake [news]] (i.e., very fake news, obviously fake news)

This is a crucial case, since it makes it evident that a descriptively adequate grammar should be able to assign two distinct structural representations to (3) if it is to capture the semantics in any way. In other words, not all iterations are structurally identical: this is a crucial point that is not emphasised in Lasnik (2011) or Uriagereka (2008). Consider (12), a Cinquean approach to adjective stacking:

(12)



This representation illustrates an important point: *fake* in FP_1 has scope over a constituent *fake news*, whose label is FP_2. If the semantic representation is built in tandem with the syntactic structure (as per direct compositionality), then this structure corresponds to segmentation (11a), and therefore with the meaning 'not fake news' (i.e., the semantic value of *fake* applied to the semantic value of *fake news*). However, because of precisely these reasons, (12) is inadequate as a structural description for the reading (11b), which involves intensive reduplication. As argued above, intensive iteration should receive an FS treatment. The expressive power of a FSA can characterise iterative patterns of adjectives, adverbs, and other categories whose semantic value is intensified in iteration, as in (13):

© 2021. John Benjamins Publishing Company
All rights reserved

(13) a. The coffee was *very very* hot
b. It's been a long *lonely, lonely, lonely, lonely, lonely* time
(Led Zeppelin, *Rock and Roll*, 1971)

Total reduplication, understood here as the iteration of a complete lexical form, is amenable to finite-state modelling. For instance, Dolatian & Heinz (2019) propose a 2-way FSA which scans an input, letter by letter, and produces an output associated to each transition between states. These states allow for loops in such a way that the FSA can produce a reduplicated output. In our case, however, we are not interested in the reduplication of strings of letters *per se*, but rather on the iteration of basic expressions under the same non-terminal (Adj⁀Adj⁀Adj…): in this sense, the number of states needed is in principle only the number of indexed categories in the lexicon. The 'explosion' of states that occurs in FS analyses of partial reduplication in morpho-phonology (Chandlee, 2014; Dolatian & Heinz, 2019) need not happen in syntax if intensive reduplication is adequately restricted. This means, among other things, that there is no need to be able to recognise infinite sequences of iterated categories, since these do not occur in the input: we want the grammar to be able to assign strongly adequate structural descriptions (Joshi, 1985: 208) to natural language strings, and in this evaluation, semantics play a crucial role. Claiming that a language that allows for iteration is necessarily context-free is faced with the difficulty of distinguishing instances of iteration in which each occurrence of an expression has scope over whatever appears at its right from instances of iteration in which there is no hierarchical relation between the iterated expressions.

3. Towards a general solution

The picture that emerges from the previous section can be summarised as follows: if the goal of grammatical theory is to assign adequate structural descriptions to natural language strings, and if adequacy is defined in terms of semantic interpretation as well as structural relations (constituency, dependencies, etc.), then the axiom of structural uniformity that is prevalent in generative grammar conspires against the descriptive adequacy of the grammar. Assuming that all syntactic structure is created equal leads to the assignment of too much structure to certain constructions, among which we have identified some cases of adjective stacking and total reduplication (iteration) in which semantic interpretation requires that either there be no hierarchical dependency between the iterated elements, or that there be, depending on specific properties of the terminals involved. Identifying what exactly in *fake* in (3) allows for both interpretations but not *old* in (2) is out-

© 2021. John Benjamins Publishing Company
All rights reserved

side the scope of the present paper,[4] but clearly it is a problem that needs to be addressed as part of an explanatory theory of the lexicon-syntax interface.

We have identified the computational complexity of these non-scopal iterated adjective constructions in English as FS, in the sense that a formal system with no memory and only capable of recognising finite expressions can assign these substrings 'flat' structural descriptions (by virtue of not imposing extra structure in the form of non-terminals nodes) which, we argue, capture their syntactic and semantic properties. However, a *uniformly* FS approach to syntax would not be descriptively adequate (as observed by Lasnik and Uriagereka). What to do? Our proposal is that *structural uniformity* must be abandoned if descriptive adequacy is a goal of grammatical theory. This entails also abandoning the idea that there is a set of mechanisms for the production and assignment of structure which are completely independent from semantics.

At this point it is legitimate to ask whether, in natural languages, evaluation in terms of levels in the CH should not proceed in terms of *constructions* in specific languages rather than as universal generalisations. In other words: given a well-formed derived expression, each component expression which is also a well-formed expression is assigned the computationally simplest structural description which captures the semantic dependencies between its component parts. We require of the competing grammars to generate the same stringsets: we can compare a FS grammar and a CG grammar that generate the string *old old man* in terms of how the structural descriptions they provide capture semantic dependencies, but not grammars that differ in their weak generative capacity.

The simplicity requirement serves the purpose to avoid extra structure. Assigning a context-free structural description to an expression for which a finite-state description suffices has consequences at both syntactic and semantic levels, particularly under direct compositionality assumptions: since additional structure is inserted in the form of non-terminal nodes, the possible targets for rules of the grammar multiply. In (14), neither example allows for an intensive reading, and whereas that is fine for *fake fake* (since there is an alternative structure available), *old old* becomes hard to interpret:

4. A reviewer suggests it is related to gradeability, such that only non-gradable adjectives would allow for the recursive reading. This is possibly part of the reason, but gradeability alone does not account for the contrast: wider construal factors need to be considered. For example, whereas in *old old man* the scopal interpretation is not available, in *old old friend* it is (referring to a friend who is old in age and who I have known for a long time); here it depends on some property of the modified N. Furthermore, we can have a non-gradable adjective, like *dead*, only admitting an intensive reading in reduplication: *a dead dead fish*, as pointed out to us by Susan Schmerling (p.c.) 'needs to stink'.

© 2021. John Benjamins Publishing Company
All rights reserved

(14) a. Fake is what the fake news was
b. ?Old is what the old man was

Similarly, reduplicated adjectives cannot be independently modified (*An [old] [very old] [old] man), which suggests that we are dealing with a syntactic unit that cannot be tampered with. The crux of the issue is to define a way in which this may be accomplished.

As a first approximation, assume that the grammar defines two processes: *chunking* and *composition* (*substitution/adjunction*). If chunking is sensitive to semantics, then we need to be able to capture the fact that in (2) we have essentially two chunks which display different kinds of structural relations:

(15) a. A man
b. old old

The structure assigned to *a man* needs to capture the fact that the quantifier has scope over the noun. Because there is hierarchical structure between the two basic expressions that make up the derived expression (15a), using a term with long provenance in grammar, we will refer to this as *hypotaxis*. But the relation between both instances of *old* is of a different kind: it is strictly *paratactic*. The distinction between *hypotaxis* and *parataxis* is necessary, in this context, to account for relations of (a) modification (b) selection, and (c) iteration. These are theory-independent, so far as we can see.

The relations of *modification* and *selection* involve hierarchy, such that the modifier is hierarchically higher than the modified and the selector, higher than the selected. This interpretation of the relations is close to that of Dependency Grammars (Osborne, 2014) and Arc Pair Grammar and its successors (Postal, 2010). In semantic terms, if A is a modifier and B a modified, or A a predicate and B its argument, for ⟦A⟧ and ⟦B⟧ the semantic values of A and B, then the relations are defined as follows:

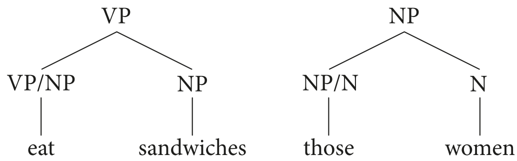
(16) a. Modification: ⟦A⟧(⟦B⟧)
b. Selection: ⟦A⟧(⟦B⟧)

The crucial point here is that the structural condition that the predicate be higher that its argument is common to both relations. The difference pertains to the definition of the categories involved: following Dowty (2003: 37), a modifier may be classified as an *adjunct* in a structure (A/A)/A, where A is an indexed category; the format for *argumental* dependencies is (A/B)/B, where A and B are *distinct*

© 2021. John Benjamins Publishing Company
All rights reserved

indexed categories.[5] The system outlined by Dowty assigns the following analysis trees to modifier-modified and predicate-argument relations:

(17)



We can get to the rule from the category definitions by means of *functional abstraction* (Dowty, 2012: 41): if we have an expression of category (A/…B)/B (i.e., a sequence of categories with B as its rightmost element), then removing B will result in an expression of category A/B. In rule format, what we have is

(18) a. VP → (VP/NP)/NP
b. NP → (NP/N)/N

Without a rule of *quantifying-in*, a CG is equivalent to a CF PSG (Lewis, 1970: 20), which means that we are within the computational ballpark we are interested in, at the level of local syntactic objects. Let us consider the case of *fake fake news* in the sense 'fake news that are not really fake news': here, we would say that the first *fake* affects *fake news*. Now we can be more explicit: the semantic value of *fake* applies to the semantic value of *fake news*:

(19) ⟦fake⟧(⟦fake news⟧)

Thus, they need to be introduced in the derivation sequentially, under (some version of) the direct compositionality hypothesis: syntactic objects are interpreted as they are introduced. We will see how this is operationalised shortly.

The case of iteration is different: here there is no hierarchy between the iterated elements. This is a structural scheme that does appear elsewhere in the grammar, as we will see shortly. Furthermore, the iterated elements do not change the category definition of the target of iteration. In other words: if *old* is an expression of category C (e.g., AP, NP/NP, etc.), then *old old* is also an expression of category C. The semantic interpretation rule has the same format, of course, but now it is the semantic value of the predicate which gets intensified that has scope over the argument:

(20) ⟦old old⟧(⟦man⟧)

As argued above, the best way to represent the lack of hierarchical dependency between the instances of *old* is to see them as a loop on a single state.

5. See Joshi & Kulick (1997) for a TAG approach to CG analysis trees.

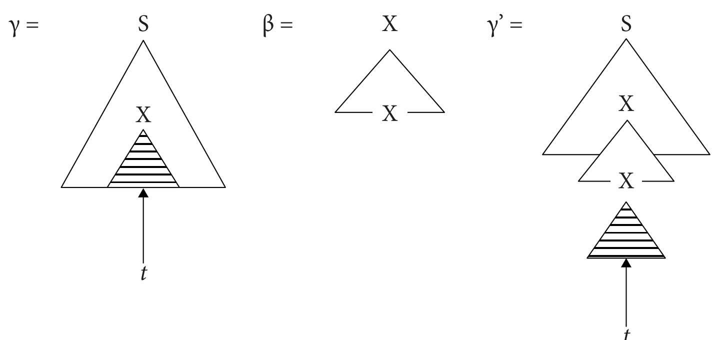
© 2021. John Benjamins Publishing Company
All rights reserved

Summarising: an adequate structural description for a string like *an old old man* combines two kinds of dependencies: iteration and modification. The former imposes no hierarchy on the expressions it contains (and is thus strictly regular), the latter does (and pushes the computational power of the grammar up; in the case under consideration here, to CF power). If we require of a grammar to assign no more structure than strictly needed to represent semantic dependencies, then there are two distinct sub-structures in *an old old man*, as in (15).

Locally, (2) displays both context-free and finite-state dependencies. But there must be a way to put both chunks together; otherwise, it would be impossible to build a compositional interpretation. We mentioned above that the grammar contains two mechanisms, *chunking* and *substitution/adjunction*. It is by means of *chunking* that expressions can be segmented into computationally uniform sub-units. What we need now is to characterise an operation that can insert a chunk in a designated position within another chunk. This kind of operation is common in syntactic theory, from Chomsky's *generalised transformations* to Joshi's *adjunction*. Here we explore a version of the latter. The grammar contains two sets of elementary trees: *initial trees* and *auxiliary trees*. *Initial trees* are the target for *adjunction* of *auxiliary trees*, which yields a *derived tree*. Adjunction

> […] composes an auxiliary tree β with a tree γ. Let γ be a tree with a node labelled X and let β be an auxiliary tree with the root labelled X also. (Note that γ must have, by definition, a node – and only one – labelled X on the frontier)
> (Joshi, 1985: 209)

Diagrammatically,



Joshi (1985: 209)

Our system has the following properties:

(21) a. It is *lexicalised*
b. It rejects *structural uniformity*
c. It rejects the autonomy of syntax

© 2021. John Benjamins Publishing Company
All rights reserved

The first property, *lexicalisation*, means that elementary trees are defined around a single lexical head (Joshi & Schabes, 1991; XTAG, 2001; Frank, 2013). Crucially, in addition to restricting the size of the elementary units in the grammar, (21a) can be interpreted in a stronger sense, as in Frank's 'Elementary TAG hypothesis':[6]

> Every syntactic dependency is expressed locally within a single elementary tree (Frank, 2013: 233)

The second, *rejection of structural uniformity*, means that these local units need not be generated by the same system of rules or display the same kind of structural and semantic dependencies: if we consider two local units α and β, α may display only *paratactic* dependencies and β only *hypotactic* dependencies. This represents a departure from some versions of TAG, where each elementary tree can be built via Merge (Frank, 2013: 240–241) and therefore the only dependencies possible are established in structurally uniform trees (binary-branching all the way down, endocentric, single-rooted). We do require, however, that dependencies within a single structural unit be computationally uniform.

Lastly, the *rejection of autonomy of syntax* entails: (i) that syntactic rules operate over semantic material (McCawley, 1971: 285), and (ii) that the semantic properties of the lexical head of a local structural unit determines the rules that can apply to that unit. A theory of syntax is, in this view, a theory of (compositional) semantics.

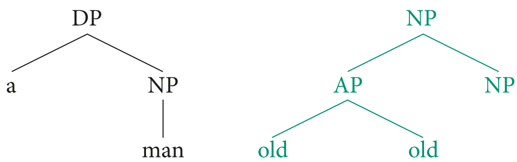
In a lexicalised TAG, the only way to separate elementary trees is to locate a single lexical head since structure is uniform all throughout (in elementary trees and derived trees). However, if we require that structural dependencies be uniform only within elementary trees, then we have a second criterion to define elementary trees: not only as the '*extended projection of a single lexical head*' (Frank, 2013: 239), but also as computationally uniform local units. This entails a novel approach to cyclicity effects, which (i) does not require designated terminal or nonterminal nodes (e.g., barriers or phase heads) to be endmarkers for probing operations, and (ii) does not define cycles *a priori* within a structural description, but depends on the presence of lexical heads in specific structures and the establishment of computationally uniform dependencies. The requirement that *local* syntactic units be computationally uniform is precisely what allows us to get the segmentation in (15): *a man* is the extended projection of the lexical item *man*, and is a local *hypotactic* unit. The iterated adjective *old old* also contains a

6. The elementary TAG hypothesis has as a direct consequence that dependencies are maintained after adjunction, by the definition of elementary trees and adjunction. MGG approaches need additional machinery to account for the same effects (see, e.g., Uriagereka, 2008: 205, fn. 13).

© 2021. John Benjamins Publishing Company
All rights reserved

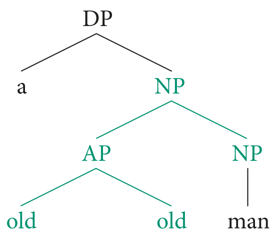
single lexical head, and only *paratactic* dependencies. The structural description assigned to *an old old man* in our terms, then, should contain the following elementary trees (nothing hinges on the particular choice of labels):

(22) Initial Tree (IT) Auxiliary Tree (AT)



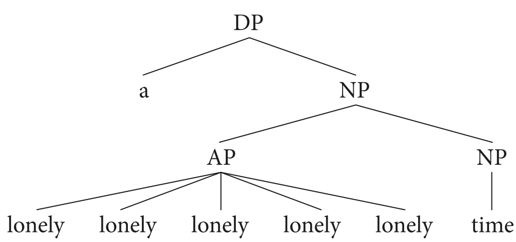
After *adjunction*, the derived tree is (23):

(23)



Evidently, the number of iterated elements can be greater than two, as in the line from *Rock and Roll* cited above:

(24)



A direct consequence of our proposal is that the two readings for *fake fake news* should receive two distinct structural descriptions, which is an advantage of the present proposal. In a way, the framework advanced in this paper follows rather closely Dowty's (2007: 30) criteria:

> **Compositional transparency**: the degree to which the compositional semantic interpretation of natural language is readily apparent [...] from its syntactic structure.
> **Syntactic economy**: the degree to which the syntactic structures of natural language are no more complicated than they need to be to produce compositionally the semantic interpretation that they have.

© 2021. John Benjamins Publishing Company
All rights reserved

The specific way in which *syntactic economy* is interpreted here is precisely defined in relation to the CH: in this way, binary branching is not always the 'simplest' possible construal for the analysis of local natural language substrings. The simplest kind of syntactic construal is a loop on a single state:

(25) A → *a*A (e.g., AP → *lonely* AP)

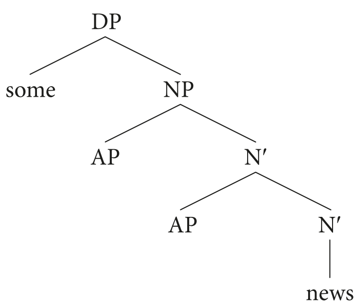
This produces a sequence of *a*'s in which there is no hypotaxis by virtue of their being dominated by the same mother label. This is the pattern for reduplicative iteration as in *old old*.

When modification (as in (16a)) is required, then additional structure is needed to capture *hypotaxis*, and consequently complexity increases with respect to the *paratactic* case. If we want to capture the non-iterative reading for *fake fake news*, we need to assign the string a structural description in which *fake* modifies *fake news*. A context-free grammar to this effect must contain, minimally, the following rules (intended for illustrative purposes only):

(26) i. DP → D NP
ii. NP → AP N′
iii. AP → A
iv. N′ → AP N′
v. N′ → N

Which generate the structure (27):

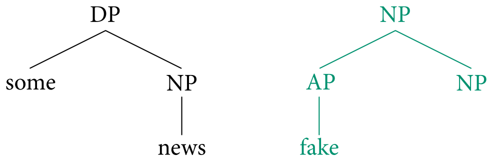
(27)



Importantly, the grammar in (26) is *not* lexicalised, because rules (i), (ii), and (iv) are not lexically anchored (XTAG, 2001: 5–6). The lexicalised elementary trees for *some fake fake news* would then be as in (28):

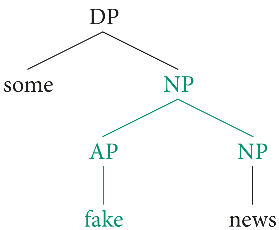
© 2021. John Benjamins Publishing Company
All rights reserved

(28) Initial Tree (IT) Auxiliary Tree (AT)



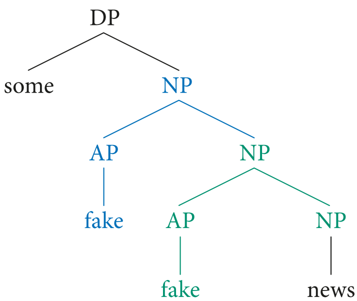
A directly compositional approach, given these elementary trees, delivers the correct reading if adjunction applies stepwise. Let us illustrate the derivation. The interpretation of the initial tree is the extended projection of its lexical 'anchor': *news*. Simplifying, let us refer to that as ⟦news⟧. The first application of adjunction produces the derived tree (29):

(29)



At this point, the interpretation is the semantic value of *fake* applied to the semantic value of *news*: ⟦fake⟧(⟦news⟧). The AT gets adjoined again, yielding (30):

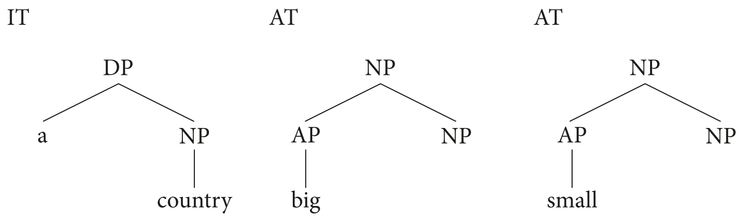
(30)



The interpretation of (30) is what we indicated in (19): ⟦fake⟧(⟦fake news⟧). The derivational system advanced here provides us with a way to link syntax and semantics in a straightforward manner.

The case of *small big country* is actually exactly parallel to the scopal interpretation of *fake fake news*, with the only difference that we have *two* ATs instead of one:

© 2021. John Benjamins Publishing Company
All rights reserved

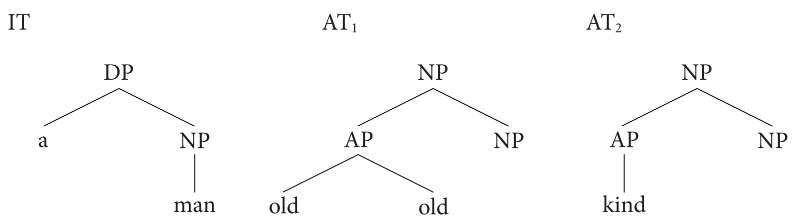
(31)


The order in which the auxiliary trees get adjoined gives two distinct semantic interpretations (which correspond to two distinct derivations): *big small country* (a country that is big for a small country) and *small big country* (a country that is small for a big country).

It is possible also to combine *paratactic* and *hypotactic* dependencies within a single NP, a possibility that has not been explored in the syntactic literature so far as we know. Consider, for instance, (32):

(32) A kind, old old man

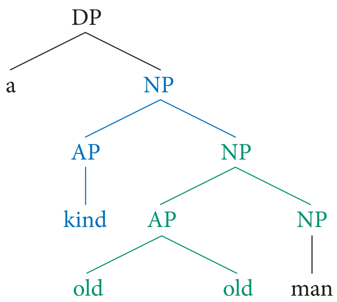
The interpretation of (32) goes along the lines of 'a very old man who was also kind': *kind* has scope over *old old man*, and *old old* has scope over *man*, with neither instance of *old* having scope over the other. The elementary trees are those in (33):

(33)


Adjunction proceeds as follows: first, AT_1 is adjoined, yielding *an old old man* (exactly as in (23)). At this point, the semantic interpretation is ⟦old old⟧(⟦man⟧), as above. Then, AT_2 is adjoined above AT_1. The semantic value of *kind* now has scope over the semantic value of the target of adjunction: ⟦kind⟧(⟦old old⟧(⟦man⟧)). The derived tree is:

© 2021. John Benjamins Publishing Company
All rights reserved

(34)



The extra structure introduced by adjunction of *kind* seems to be justified; if we attempt the same test as in (14), we see that *kind* may be clefted, just like *fake* in the non-iterative reading:

(35) Kind is what the old old man was

To summarise, a directly compositional lexicalised TAG approach, in addition to the rejection of the axiom of structural uniformity seems to pay off in the empirical analysis of simple cases, yet challenging for many current models of syntactic structure. The strictly CF dependencies that MGG models attempt to capture are indeed accounted for, in addition to iterative patterns that receive a simpler structural analysis in terms of FS loops.

4. Iteration and coordination

The system sketched in the previous section provides adequate structural descriptions for English sentences containing adjectival iteration, to which a CFG assigns too much structure which is not justified in semantic terms or syntactic terms. In the present view, iteration is best handled by FS rules. Can we extend the 'mixed computation' mechanisms proposed above to other phenomena in English? In what remains of this paper, we will consider the case of reduplicative coordination and symmetric conjunction.

Consider Examples (36) and (37):[7]

(36) The Trump Twitter Archive shows the Republican whining about his predecessor's golfing *over and over and over and over and over* again.
Steve Bonen, 'The problem with Trump's defense of his many golf outings', MSNBC, 13 July 2020

7. We owe (36) and (37) to Susan F. Schmerling, as well as much discussion about their analysis.

© 2021. John Benjamins Publishing Company
All rights reserved

(37) She would *wait and wait and wait and wait* / For her steady date
Robert Byrd, ‘Over and Over’, 1958

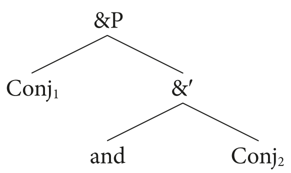
Superficially, sentences like (36) and (37) seem to contain coordinated structures. As observed by De Vos (2005), these structures (which he dubs *repetitive pseudo-coordinations*, ReCo) may yield serial and repetitive readings, but the fact that they are compatible with states (e.g., *wait*) suggests that repetition of an event is not a necessary condition. Even with dynamic verbs, repetitive readings are not guaranteed:

(38) Caesar’s legions *marched and marched* for days (De Vos’ Example (4))

In (38) there is only one event of marching, which extends for a long period of time. De Vos claims, correctly in our opinion, that a plural subject licenses (but not necessarily coerces) an *iterative* reading, whereas a singular subject allows for an *intensive* reading.

What kind of structural description is adequate for examples like (36–38)? In a structurally uniform system coordination always binary:

(39)



Analyses of this form are defended in Kayne (1994); Progovac (1998); Chomsky (2013), among many others. Borsley (2005) provides an overview of arguments against a structure like (39) and observes that in fact such an analysis is *not* widely assumed outside MGG. There are two main problems with (39): (i) structural uniformity (binary-branching and obligatory hypotaxis) and (ii) projection of the coordinating conjunction to phrasal level (which means that if we coordinate NPs, the result will not be an NP, but a &P).

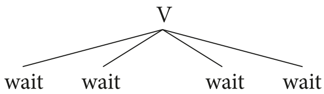
The proposal we want to make with respect to (36–39) is that they are *not* instances of coordination at all (cf. De Vos, 2005: 4); rather, the iterated expression constitutes a syntactic unit for all subsequent intents and purposes, just like *old old old*.... Furthermore, at least in the cases treated here, a very similar semantic effect arises: the semantic value of the iterated expression intensifies. Thus, (36) is not interpreted as literally referring to five events of whining, but rather expresses that whining took place frequently (in an unspecified number); in (37) there are not four events of waiting but only one, which extends for a long time.[8] These

8. In this case, *wait* may receive rhetorical accent: *wa·it.*

© 2021. John Benjamins Publishing Company
All rights reserved

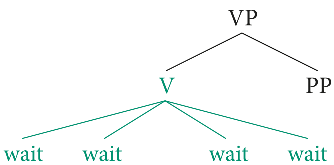
cases of iteration, like adjective iteration, are not restricted to a certain number of elements, nor is there anything *inside* the iteration that requires memory storage (as also pointed out in Uriagereka, 2008). Thus, we may assign (36–38) the same *flat* (FS) structure we assigned to iterated adjectives:

(40)



Let us comment on two features of this representation: (i) the lack of any *and* in the structure and (ii) the use of the label V (as opposed to VP). Recall that, in our view, there is no coordination in De Vos' ReCo: the word *and* makes no contribution to the syntax or the semantics of the construction. Just like *and* in cases like *Which dresses is she going to go/up/take and ruin now?* (Ross, 1967), the presence of the word *and* does not mean that we are in the presence of coordination. Each of the leaves in (40) could be *and wait*, but in this context its grammatical properties would not change. Which takes us to the second point: labelling (40) as V. This entails that a structure with ReCo would require *substitution* of V in the frontier of an initial tree by the root of the tree in (40).

(41)



Note that the PP is *outside* the iterated structure. This makes two predictions: (i) there is no structural position available for an internal argument inside the auxiliary tree (40), and (ii) movement of an NP internal argument in a parallel structure with a transitive verb does not violate the Coordinate Structure Constraint (CSC). Note that, if ReCo was an instance of coordination involving Vs or VPs, then NP movement would require chopping the NP from one conjunct. We can rule out the CSC violation with examples like (42a–b):

(42) a. Six web series that we can *watch and watch and watch…*
(https://economictimes.indiatimes.com/magazines/panache/2016-all-oer-again-six-web-series-that-we-can-watch-and-watch-and-watch/articleshow/55838344.cms. Relativisation, A′-movement)

b. It was *said, and said, and said, and said*, and that's why I say it
(Tom Mould, *Choctaw Prophecy: A Legacy for the Future*, p. 29. Passivisation, A-movement)

© 2021. John Benjamins Publishing Company
All rights reserved

Note that in (42b) there is only one auxiliary *be*, yet all instances of *say* appear in participial form; passive *be* determines that every expression dominated by V (be it a single orthographical word or an iterated orthographical word) will bear participial morphology. This entails accepting the existence of multi-word expressions in the grammar. Finding examples with *wh*-interrogatives is trickier, probably because of the intensive value of iteration: the exchange in (43a) sounds more natural than (43b):

(43) a. – What did she say?
– She *said and said and said* that I should clean my room
b. – #What did she *say and say and say*?
– She said that I should clean my room

In (43a) there is only one direct object: if the iterated *said* were VPs, additional mechanisms would need to be invoked (e.g., RNR) which would in turn predict a more articulate internal structure to the iteration than what we find. However, it is possible to have FS ReCo of VPs: here, the direct object *must* appear in each reduplicated term (as in (44a), suggested by a reviewer):

(44) a. We [$_{VP}$ ate fish] and [$_{VP}$ ate fish] and [$_{VP}$ ate fish] and [$_{VP}$ ate fish] until I couldn't stand it anymore
b. *We [$_{V}$ ate] and [$_{VP}$ ate fish] and [$_{V}$ ate] and [$_{V}$ ate] until…

It seems that the finite-state approach to iteration sketched in the previous sections can be extended to ReCo fruitfully. In the next section we will address structural issues that emerge for *true* coordination under structural uniformity.

4.1 Symmetric coordination and flat structures

Schmerling (1975) observes that there are cases of natural language conjunction in which the following equivalence holds:

(45) $p \wedge q \equiv q \wedge p$

These are instances of *symmetric coordination*, illustrated in (46):[9]

(46) a. Paris is the capital of France, and Rome is the capital of Italy.
b. Rome is the capital of Italy, and Paris is the capital of France.

This is the conjunction available in propositional logic, but as is frequently pointed out in introductory textbooks in logic, ∧ and the word *and* are not

9. (46), (47), and (48) are taken from Schmerling (1975: 211)

© 2021. John Benjamins Publishing Company
All rights reserved

equivalent, since natural language *and* is not always commutative (Allwood et al., 1977: 33, ff.). Commutativity is *not* a property of the coordination in (47):

(47) a. Harry stood up and objected to the proposal.
b. Harry objected to the proposal and stood up.

In (47) there is a temporal and causal ordering of the events, such that the situation described by (47a) is not the same as that described by (47b). In (47) the permutations, while semantically different, are both grammatical. However, it is easy to come up with examples in which not even this weaker condition holds:

(48) a. Smile and the world smiles with you
b. *The world smiles with you and smile

Coordinations of this sort are *asymmetric*. The question arises whether *symmetric* and *asymmetric* coordination should receive different structural analyses. Our view is that the semantic differences between *symmetric* and *asymmetric* conjunction need to be accounted for in the structural descriptions assigned to them. In particular, we will look at *symmetric* conjunction.

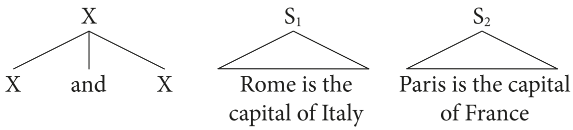
As before, one of the crucial issues is *iteration*:

(49) Rome is the capital of Italy, Paris is the capital of France, and Berlin is the capital of Germany

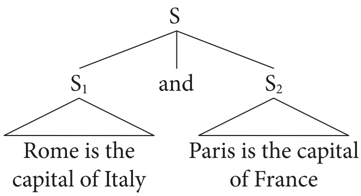
The treatment of coordinated structures with more than two conjuncts is problematic: Ross (1967); McCawley (1998) and others would assign (49) a ternary branching structure; such a structure would be unavailable in MGG. Borsley (2005) argues that in a coordination like *Tom ate a hamburger, and Alice drank a Martini and Jane a beer* neither the first and the second nor the second and the third conjuncts form a constituent excluding the other conjunct; his arguments extend to conjunctions of *n* terms. Our argument is that symmetric coordination of *n* terms behaves like iteration in the sense that it needs to be assigned a *paratactic* structure in which there is no hierarchy between the conjuncts; asymmetric coordination does require there to be hierarchy between the conjuncts, in the form of a *hypotactic* structural description. Symmetric conjunction defines an *unordered* set of terms which presents the property of *commutativity*; asymmetric conjunction defines an *ordered* set of terms, which do not commute. The mixed computation approach thus extends beyond iteration. Our proposal answers the question posed in Abeillé (2003: 6) whether the structure of coordinations is hierarchical or flat in a novel way: it depends on the semantic relations between the conjuncts. Thus, if we consider the semantic interpretation assigned to (46), under direct compositionality, the equivalence between $p \wedge q$ and $q \wedge p$ only holds if the terms of the coordination are introduced together (as in a MCTAG).

© 2021. John Benjamins Publishing Company
All rights reserved

Following Sarkar & Joshi (1997: 612–613), the elementary trees in a symmetric coordination of likes would be:

(50)


The derived tree involves only *substitution* of X by S,[10] since there are no common terms to the conjuncts:

(51)


Note that there is no hierarchical relation between S_1 and S_2, the structure is *paratactic*. Then, more conjuncts can be added to the symmetric structure (as in (49)) simply by adding more branches. Importantly, (51) differs from the structure assigned to iteration in that the latter may be modelled as a finite-state loop, but the former requires CF power, since we have a non-terminal dominating a string of non-terminals and a terminal: the rule format of CF grammars in Greibach-normal form. However, it also means that a template like (39) would assign too much structure to a symmetric conjunction.

5. Some conclusions

In this paper we examined some empirical consequences of assuming structural uniformity in natural language syntax. Structural uniformity assigns too much structure in some cases and too little in others; we illustrated the problem of 'too much structure' with English attributive adjective iteration. The phenomenon, however, is much more general as is our proposed solution. We argued that a descriptively adequate theory of grammar must take into consideration how syntactic configuration compositionally specifies semantic interpretations; an ade-

10. Sarkar & Joshi propose an operation *conjoin* to deal with coordination with structure sharing (RNR, ATB movement, etc.); this operation can be formulated in terms of *adjunction* under an appropriate definition of the elementary trees (Sarkar & Joshi, 1997: 613). These additional mechanisms are not necessary for symmetric coordination of likes without structure sharing, so far as we can see.

© 2021. John Benjamins Publishing Company
All rights reserved

quate theory of natural language syntax cannot be independent of semantics. In proposing that syntactic structure is not uniform, we are also forced to make explicit the kinds of dependencies that we find within local domains; FLT proves an invaluable tool to characterise these domains, which we identified with elementary trees in a LTAG. An exploration of the consequences of our view, which we call 'mixed computation', leads to interesting proposals concerning the nature of local domains in syntax: syntactic *cycles* are defined as chunks of structure with a single lexical anchor and uniform dependencies which allow to be targeted by tree composition operations (*adjunction* and *substitution*) and rules of semantic interpretation (in the sense of Dowty et al., 1980). *Grosso modo*, change the dependencies, and you change the cycle; change the lexical anchor, and you change the cycle (by virtue of changing the elementary tree). The dependencies include hypotaxis and parataxis, as indicated above, but within hypotaxis we need to be able to refine the system: the descriptive adequacy of structural descriptions featuring center embedding is not the same as tail recursion or crossing dependencies.

Let us close by briefly considering a further example of the empirical advantages of mixed computation over structural uniformity for comparative purposes. Above we considered distinct grammars that generated the same stringsets, concluding that if a FS grammar is available for *old old man*, it must be preferred over a CF grammar that generates the same string. Cross-linguistically, we need to consider each language separately, and compare construction by construction. For example,

(52) a. Jan Piet Marie zag helpen zwemmen (Dutch German)
Jan Piet Marie saw help swim
'Jan saw Piet help Mary swim'
b. Merve Ömer'in Esra'nın yüzmesine yardım ettiğini gördü (Turkish)
Merve Ömer Esra swim help give saw
'Merve saw Ömer help Esra swim'
c. John saw Peter help Mary swim (English)

If we consider the relations between NPs and the VPs to which they correspond, we obtain the following abstract formats:

(53) a. NP_1 NP_2 NP_3 VP_1 VP_2 VP_3 (crossing dependencies between two sets of elements)
b. NP_1 NP_2 NP_3 VP_3 VP_2 VP_1 (center embedding)
c. NP_1 VP_1 NP_2 VP_2 NP_3 VP_3 (tail recursion)

The differences in structure between the sentential complementation patterns of different languages suggest that a universal template may not be the best aid for

© 2021. John Benjamins Publishing Company
All rights reserved

grammatical analysis:[11] structural uniformity requires that all constructions in all languages follow the same format. However, applying that reasoning to the comparative analysis of (52) we would either assign too much or too little structure to some cases, depending on what the underlying universal format for phrase structure is. After all, if the English example can be generated using only *substitution*, why assume anything more complex? If English and Turkish are both well within CF territory, do we really need to have CS machinery available in the grammatical descriptions for these languages and make full use of those formal tools? We do not aim at settling the issue here, but want to pose it as a fundamentally empirical question in what pertains to the relation between FLT and grammar. Mixed computation provides the grammarian with an interesting tool to probe for unexpected sources of cross-linguistic variation.

References

Abeillé, A. (2003). A lexicalist and construction-based approach to coordinations. In S. Müller (ed.), *Proceedings of the HPSG03 Conference* (pp. 5–25). Stanford: CSLI.

Abney, S. (1987). The English noun phrase in its sentential aspect. PhD thesis, MIT.

Alexiadou, A. (2014). *The syntax of adjectives*. In A. Carnie, Y. Sato & D. Siddiqi. (eds.), *Routledge Handbook of Syntax* (pp. 89–107). London: Routledge.

Aljović, N. (2010). Syntactic positions of attributive adjectives. In P. Cabredo Hofherr & O. Matushansky (eds.), *Adjectives: Formal analyses on syntax and semantics* (pp. 29–52). Amsterdam: John Benjamins. https://doi.org/10.1075/la.153.01alj

Allwood, J., Andersson, L-G. & Dahl, Ö. (1977). *Logic in Linguistics*. Oxford: OUP. https://doi.org/10.1017/CBO9780511621017

Bach, E. (1976). An Extension of Classical Transformational Grammar. In *Problems of Linguistic Metatheory* (pp. 183–224). Michigan State University.

Borsley, R. D. (2005). Against CoordP. *Lingua 115*. 461–482. https://doi.org/10.1016/j.lingua.2003.09.011

Bortolotto, L. (2016). The Syntax of Relational Adjectives in Romance. PhD thesis, University of Venice.

Chandlee, J. (2014). Strictly Local Phonological Processes. PhD thesis, University of Delaware.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

Chomsky, N. (1986). *Barriers*. Cambridge, Mass.: MIT Press.

11. Proposals such as Uriagereka (2008) and Lasnik (2011) restrict the possibility of defining syntactic units of different computational complexities to very specific cases (generally, iteration and reduplication; Uriagereka, 2008: 206 adds small clauses to the set of FS structures, based on MGG-internal issues pertaining to phrasal projection that would not hold given the definition of *modification* in (16a)). Our perspective is much more general: mixed computation is the norm, not the exception, in grammatical analysis (Krivochen, 2015; Krivochen & García Fernández, 2019).

© 2021. John Benjamins Publishing Company
All rights reserved

Chomsky, N. (2013). Problems of Projection. *Lingua 130*, 33–49. https://doi.org/10.1016/j.lingua.2012.12.003
Chomsky, N. & Miller, G. (1963). Introduction to the Formal Analysis of Natural Languages. In R. Luce, R. Bush & E. Galanter (eds.), *Handbook of Mathematical Psychology* (pp. 269–321). New York: Wiley & Sons.
Cinque, G. (1994). On the Evidence for Partial N-Movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi & R. Zanuttini (eds.), *Paths towards Universal Grammar* (pp. 85–110). Washington, D.C.: Georgetown University Press.
Cinque, G. (2010). *The Syntax of Adjectives. A Comparative Study*. Cambridge, Mass.: MIT Press. https://doi.org/10.7551/mitpress/9780262014168.001.0001
De Vos, M. (2005). The syntax of pseudo-coordination in English and Afrikaans. PhD thesis, Leiden University.
Dolatian, H. & Heinz, J. (2019). Learning reduplication with 2-way finite-state transducers. *Proceedings of the 14th International Conference on Grammatical Inference*, 67–80.
Dowty, D. (2003). The Dual Analysis of Adjuncts/Complements in Categorial Grammar. In E. Lang, C. Maienborn & C. Fabricius-Hansen (eds.), *Modifying Adjuncts* (pp. 33–66). Berlin: de Gruyter.
Dowty, D. (2012). Compositionality as an empirical problem. In C. Baker & P. Jacobson (eds.), *Direct Compositionality* (pp. 23–101). Oxford: OUP.
Dowty, D., Wall, R. & Peters, S. (1980). *Introduction to Montague semantics*. Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-9065-4
Emonds, J. (1976). *A Transformational Approach to English Syntax*. New York: Academic Press.
Epstein, R. (2011). *Classical mathematical logic*. Princeton: Princeton University Press. https://doi.org/10.1515/9781400841554
Frank, R. (2013). Tree adjoining grammar. In M. den Dikken (ed.) *The Cambridge Handbook of Generative Syntax* (pp. 226–261). Cambridge: CUP. https://doi.org/10.1017/CBO9780511804571.011
Fukui, N. & Narita, H. (2014). Merge, labelling, and projection. In A. Carnie, Y. Sato & D. Siddiqi (eds.) *The Routledge Handbook of Syntax* (pp. 3–23). London: Routledge.
Greibach, S. (1965). A New Normal-Form Theorem for Context-Free Phrase Structure Grammars. *Journal of the ACM 12*(1), 42–52. https://doi.org/10.1145/321250.321254
Jackendoff, R. (1977). *X-bar syntax: A study of phrase structure*. Cambridge, Mass.: MIT Press.
Jacobson, P. (2012). Direct Compositionality. In W. Hinzen, E. Machery & M. Werning (eds.) *The Oxford Handbook of Compositionality* (pp. 109–129). Oxford: OUP.
Joshi, A. (1985). Tree adjoining grammars. In D. Dowty, L. Karttunen & A. Zwicky (eds.) *Natural Language Parsing* (pp. 206–250). Cambridge, Mass.: CUP. https://doi.org/10.1017/CBO9780511597855.007
Joshi, A. & Kulick, S. (1997). Partial Proof Trees as Building Blocks for a Categorial Grammar. *Linguistics and Philosophy 20*(6). 637–667. https://doi.org/10.1023/A:1005311532280
Joshi, A. & Schabes, Y. (1991). Tree-Adjoining Grammars and Lexicalized Grammars. Technical Reports (CIS). Paper 445. http://repository.upenn.edu/cis_reports/445
Karlsson, F. (2010). Syntactic recursion and iteration. In H. van der Hulst (ed.) *Recursion and Human Language* (pp. 43–67). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110219258.43
Kayne, R. (1984). *Connectedness and Binary Branching*. Dordrecht: Foris. https://doi.org/10.1515/9783111682228
Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

© 2021. John Benjamins Publishing Company
All rights reserved

Krivochen, D. G. (2015). On Phrase Structure building and Labeling algorithms: towards a non-uniform theory of syntactic structures. *The Linguistic Review 32*(3). 515–572.
Krivochen, D. G. & L. García Fernández. (2019). On the position of subjects in Spanish periphrases. *Borealis*, *8*(1), 1–33. https://doi.org/10.7557/1.8.1.4687
Ladusaw, W. (1980). *Polarity sensitivity as inherent scope relations*. Bloomington: Indiana University Linguistics Club.
Lasnik, H. (2011). What Kind of Computing Device is the Human Language Faculty? In A-M. Di Sciullo & C. Boeckx (eds.) *The Biolinguistic Enterprise* (pp. 354–365). Oxford: OUP.
May, R. (1985). *Logical Form: Its Structure and Derivation*. Cambridge, Mass.: MIT Press.
McCawley, J. (1971). Interpretative semantics meets Frankenstein. *Foundations of Language 7*, 285–296.
McCawley, J. (1998). *The Syntactic Phenomena of English*. Chicago: University of Chicago Press.
Newman, S. (1946). On the stress system of English. *Word 2*. 171–187. https://doi.org/10.1080/00437956.1946.11659290
Osborne, T. (2014). Dependency grammar. In A. Carnie, Y. Sato & D. Siddiqi. (eds.) *Routledge Handbook of Syntax* (pp. 604–626). London: Routledge.
Postal, P. (1964). *Constituent Structure*. Bloomington, Indiana: University of Bloomington.
Postal, P. (2010). *Edge-Based Clausal Syntax*. Cambridge, Mass.: MIT Press. https://doi.org/10.7551/mitpress/9780262014816.001.0001
Progovac, L. (1998). Structure for coordination. *Glot International 3*(7), 3–6.
Reich, P. (1969). The finiteness of natural languages. *Language 45*(4), 831–843. https://doi.org/10.2307/412337
Ross, J. R. (1967). Constraints on Variables in Syntax. PhD Thesis, MIT.
Sarkar, A. & A. Joshi. (1997). Handling coordination in a tree adjoining grammar. Technical report, University of Pennsylvania.
Schmerling, S. (1975). Asymmetric Conjunction and rules of Conversation. In P. Cole & J. Morgan (eds.), *Syntax and Semantics 3: Speech Acts*, 211–231. New York: Academic Press. https://doi.org/10.1163/9789004368811_009
Schmerling, S. (2018). Rhetorical meaning. *Linguistic Frontiers 1*(1). 1–8. https://doi.org/10.2478/lf-2018-0001
Scott, G-J. (2002). Stacked adjectival modification and the structure of nominal phrases. In G. Cinque (ed.) *Functional structure in DP and IP*. (pp. 91–120). Oxford: OUP.
Shieber, S. (1985). Evidence against the Context-Freeness of Natural Language. *Linguistics and Philosophy 8*(3), 333–343. https://doi.org/10.1007/BF00630917
Stabler, E. (1997). Derivational Minimalism. In C. Retoré (ed.), *Logical Aspects of Computational Linguistics* (pp. 68–95). New York: Springer. https://doi.org/10.1007/BFb0052152
Stabler, E. (2013). The epicenter of linguistic behavior. In M. Sanz, I. Laka & M. Tanenhaus, (eds.), *Language Down the Garden Path*. Oxford: OUP. 316–323. https://doi.org/10.1093/acprof:oso/9780199677139.003.0018
Steedman, M. (2019). Combinatory Categorial Grammar. In A. Kertész, E. Moravcsik & C. Rákosi (eds.), *Current Approaches to Syntax: A Comparative Handbook* (pp. 389–420). Berlin: de Gruyter. https://doi.org/10.1515/9783110540253-014
Svenonius, P. (1994). On the structural location of the attributive adjective. *Proceedings of the West Coast Conference on Formal Linguistics 12*, 439–454.

© 2021. John Benjamins Publishing Company
All rights reserved

Uriagereka, J. (2008). *Syntactic anchors: on semantic structuring*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511481482
Uriagereka, J. (2012). *Spell-Out and the Minimalist Program*. Oxford: OUP.
XTAG group. (2001). A lexicalized TAG for English. Technical report, University of Pennsylvania. https://repository.upenn.edu/cgi/viewcontent.cgi?article=1020&context=ircs_reports

Address for correspondence

Diego Gabriel Krivochen
Department of Linguistics, Philology and Phonetics
University of Oxford
Centre for Linguistics and Philology
Walton Street
Oxford OX1 2HG
England
diego.krivochen@ling-phil.ox.ac.uk

Publication history

Date received: 10 March 2021
Date accepted: 30 June 2021

© 2021. John Benjamins Publishing Company
All rights reserved