DE GRUYTER MOUTON Folia Linguistica 2024; 58(3): 699–727

Diego Gabriel Krivochen*

Constituents, arrays, and trees: two (more) models of grammatical description

https://doi.org/10.1515/flin-2024-2025
Received July 25, 2023; accepted May 20, 2024; published online June 19, 2024

Abstract: Generative syntax was built on the foundations of Immediate Constituent (IC) analysis, and IC methods and heuristics were an important tool in the early days of the generative enterprise. However, developments in the theory entailed a departure from some fundamental IC assumptions: we will argue that structural descriptions in contemporary generative grammar (transformational and non-transformational) define not constituents, but strictly ordered sequences closer to *arrays*. We therefore define and characterise IC approaches to syntax as opposed to what we will call *Array-Based* (AB) approaches. IC grammars define distributional generalisations, and proper containment and *is-a* relations between indexed distributionally defined categories. AB grammars, in contrast, define strictly ordered sequences of categories. In this paper we introduce and define the fundamental properties of IC grammar, and the changes in the generative theory that introduced arrays in phrase structure. We argue that it is crucial to distinguish between IC and AB grammars when evaluating the empirical adequacy of structural descriptions used in current syntactic theorising, as structures in AB and IC grammars represent different relations between expressions and may be better suited for different purposes.

Keywords: immediate constituents; arrays; segmentation; substitution; hierarchy

1 Introduction

Hockett (1954) and Schmerling (1983) distinguish two kinds of grammatical theories: Item-and-Arrangement (IA) and Item-and-Process (IP). The former include all versions of post-Bloomfieldian American structuralism and generative grammar (transformational or not): IA models are based on a series of assumptions about the nature of the annotated descriptions for natural language sentences ('structural descriptions' or 'phrase markers' in generative grammar), in particular: (i) that the categories of the grammar are structured in 'levels of description' or 'levels of linguistic analysis' (e.g., Bloomfield

***Corresponding author: Diego Gabriel Krivochen**, Faculty of Linguistics, Philology and Phonetics, University of Oxford, Oxford, UK, E-mail: diego.krivochen@ling-phil.ox.ac.uk. https://orcid.org/0000-0001-6145-1057

1926; Harris 1946, 1951, 1970 [1964]; Wells 1947) such as *morpheme*, *word*, *phrase*, *sentence*, (ii) that categories assigned to grammatical objects are defined in terms of the contexts in which they may appear, and (iii) that basic and derived expressions are related by means of combinatoric or compositional operations. This gives rise to a distinction between lexical and phrasal categories at the core of the formalism: the former cannot be further decomposed into smaller constituent parts, whereas the latter can. The goal is to provide maximally accurate explicit transcriptions of 'utterances' at various levels of analysis in terms of abstract categories (cf. Chomsky 1951: 2).

The kinds of structural descriptions that can be obtained with IA grammars were originally focused on capturing and generalising distributional constraints on expressions (Harris 1946: 161, 1957: 284). Expressions with the same distributional constraints are assigned to the same category: a category is a set of distributional regularities.[1] These categories constitute the vocabulary of non-terminal symbols of the grammar, with expressions being the terminal vocabulary (see Hopcroft and Ullman 1969: Chapter 2 for discussion). Syntactic operations (e.g., Harrisian expansions applied to irreducible structural units called *kernels*), then, refer to and operate over these non-terminal symbols (with different degrees of generality at different points in the history of the theory), which are variables ranging over sequences of expressions. We will see examples of such expressions shortly.

These assumptions, to which we return in detail in Section 2, guided the early days of generative grammar (Chomsky 1955: Chapters VII, VIII; Postal 1964: 7). However, towards the mid-80s changes were made to the generative theory that, we contend, pushed generative grammar away from constituency-based analyses and towards the mathematical definition of arrays: instead of distinct, distributionally defined, levels of organisation of linguistic material (e.g. lexical vs. phrasal), the grammar now generates strictly ordered sequences of categories following a universal hierarchy. We propose that instead of IC-based, much of contemporary syntactic theory is *array-based* (AB). These changes are examined in detail in Section 3. We argue that the ordered sequences that underpin much contemporary generative work do not implement constituency-based analyses, and thus need to be carefully distinguished from IC grammars. The predictions made by IC theories are not necessarily valid when considering arrays, and vice-versa: IC and AB models differ not only formally but also in terms of their methods of data analysis and their aims. We contend that the use of common terminology

1 Such a system has the consequence that every expression with a distribution has a category. This is problematic insofar as natural languages seem to contain expressions assigned to no categories and which receive no model-theoretic interpretation, but every well-formed sequence in which they appear does have a meaning: these are called *syncategorematic expressions* (see Krivochen 2023: 81–82; Schmerling 2018: 151–154 and references therein). Examples of syncategorematic expressions in natural language include English infinitival *to* and Spanish DOM *a*.

and symbology has obscured the distinction and given rise to debates about syntactic structure where parties talk at cross purposes. A careful examination of the theoretical and methodological foundations of IC and AB constitutes an important aspect of syntactic metatheory (alongside, for instance, the *procedural-declarative* distinction, cf. Pullum 2019), can inform the linguist's choice of analytical framework in specific cases, and provides a framework within which to evaluate arguments for or against specific analyses (e.g., nominal constructions as Determiner Phrases or Noun Phrases, sentences as exocentric Ss or Tense Phrases, etc.). This paper focuses on the fundamentals of IC analysis and examines the transition between IC and AB in syntactic theory.[2]

2 Immediate constituent analyses

Some formal preliminaries. In formal language theory, which constituted the basis for the mathematical work by Chomsky and others in the 50s and early 60s, a grammar G is a pair

1) $G = (\Sigma, \delta)$

where Σ is the alphabet of the grammar and δ is a set of rules that operate over the symbols in the alphabet. The alphabet contains two kinds of symbols: terminals and non-terminals. δ is a set of transition functions (or 'productions', or 'rewrite rules') which take the form of mappings from (read: with input) well-formed formulae to (read: with output) well-formed formulae (Post 1943: 197). Transitions are notated as rules: $X \rightarrow Y$ (rewrite X as Y), where X and Y are variables over (sequences of) expressions in the alphabet.[3] These rules were initially defined as mappings from

2 It is important to note that in this work we focus on the grammatical consequences of IC/AB, without saying anything about the 'cognitive reality' of the grammar. For our purposes, syntactic theory deals with expressions and allowed relations, without necessarily making any claim about human knowledge of language. In this sense, we follow Postal (2010: 3):

> […] *I understand grammatical study to be concerned with the characterization of NL* [Natural Languages], *not with the characterization of knowledge of NL nor with any mechanisms that yield such knowledge.*

3 The idea can be traced back to Harris (1946, 1951), in terms of substitutability: for example, Harris (1946: 166) uses notation like BC = A to indicate that a sequence of morphemes of classes B and C can be substituted by a morpheme of class A and BC = AC to indicate that A can substitute for B only when followed by C, which is equivalent to a context-sensitive rule. Instead of 'rewriting' rules as mappings from strings to strings, there is a system of equations that involve the substitution of a morpheme or sequence of morphemes. Harris (1968: 30) considers the system of rewrite rules in Chomsky (1957) a 'codification' of the expansions of word classes in the distributional analysis; Chomsky (1970: 211), after presenting the context-free X-bar schema, observes that

strings to strings, such that a rewriting rule takes as its input a string S_n containing X and outputs a string S_{n+1} containing Y instead of X (specific conditions are imposed over X and/or Y. For instance, Y cannot be identical to X, nor can Y be the null element Ø). This allows us to define the relation *follows from*:

> *A string β* *follows from* *a string α if* $\alpha = Z^\frown X_i^\frown W$ *and* $\beta = Z^\frown Y_i^\frown W$ [where ⁀ is linear concatenation] […]
>
> *A* *derivation* *of the string* S_t *is a sequence* $D = (S_1, \ldots, S_t)$ *where* $S_1 \in \Sigma$ *and for each* $i < t$, S_{i+1} *follows from* S_i (Chomsky 1956: 117. Underlined in the original)

We can further specify -simplifying matters slightly- that in D, S_1 is a designated initial symbol, which we call the 'axiom'[4] (Mateescu and Salomaa 1997a: 39), and that S_t is a sequence of terminal symbols: a derivation of a terminal string is the ordered set of strings whose first member is a designated non-terminal symbol and whose last member is that terminal string, with the order given by *follows from* and its transitive closure. Let us give an example:

2) $\Sigma = \{S, A, B, a, b, c\}$
$S \rightarrow a\text{A}$
$A \rightarrow b\text{B}$
$B \rightarrow c$

This grammar generates the following sequence of strings:

> *A structure of the sort just outlined* [the X-bar schema] *is reminiscent of the system of phrase structure analysis developed by Harris in the 1940's. In Harris' system, statements applying to categories represented in the form* X^n *(n a numeral) applied also to categories represented in the form* X^m *(m < n).*

In *From morpheme to utterance* (1946), Harris defines a system with 'unidirectional substitutability' which delivers recursive hierarchical structure. Let's see how it works. Suppose that the heuristic of segmentation and substitution tells us that *book* and *books* are constituents. Then, *books* = *book* + *-s*. If we assign these two expressions to the same category, however, we would have a problem: *books* cannot replace *book* in the context __ + *-s*. We want to say that they are both Ns, but at the same time that there are restrictions: their distribution is not identical. Harris solves this issue by effectively creating different levels within a category: take the base form *book* to be N^1. Then, *books* will be N^1 + -*s*: call this N^2. If we add a determiner, we get *the books*: *the* + N^2. Call this N^3. The general statement is, then, that a symbol X is assigned a superscript 1 the first time we encounter it. Then, every other instance of X, should it substitute for X in every context, will also be X^1. If we encounter an X that cannot be substituted for all the preceding X^1 we add one to the value in the superscript: we get an X^2.

4 Classical generative grammar used S, which stands for *sentence*. Henceforth, and following Chomsky's convention, if we use S with a subscript, then we are referring to a string; if it has no subscript, it is the designated initial symbol.

3) S_1: S (the initial symbol, or ‘axiom’)
S_2: *a*A
S_3: *ab*B
S_4: *abc*

The sequence D = (S_1, S_2, S_3, S_4) is a derivation of the terminal string *abc*: there is no S_n such that S_n *follows from* S_4 (in which case we say that the derivation is *terminated*, Chomsky 1956: 117). Some (e.g., Oehrle 2000: 276) call the terminal string *abc* the *closure* of the relation *follows from* (strictly speaking, Oehrle considers the inverse relation, which he calls *directly derives*). The symbols S, A, B belong to the set of nonterminal symbols V_N, and *a*, *b*, *c* belong to the set of terminal symbols V_T. When we approach the issue from the perspective of natural language grammars, some of the symbols of the grammar will stand for words and morphemes (grammatically significant parts of words). These words and morphemes constitute the terminal vocabulary of the grammar: they cannot appear on the left-hand side of production rules. Other symbols stand for sequences (of sequences) of terminals (that is, phrases) and constitute the nonterminal vocabulary (the designated root node S is one of these, standing for the ‘*longest phrase*’, Chomsky 1959: 140).

Rewrites as allows us to define a relation from symbols to sequences: the two-place relation *is-a*. Specifically, if a basic condition over phrase structure grammars is that for any S_n, S_m, if S_m *follows from* S_n by replacing a single symbol A of S_n by a non-null sequence α (Chomsky 1959: 143), then the mapping from S_n to S_m defines the relation *is-a*(α, A) (Chomsky 1955: 175; Postal 1964: 7). For example, let S_n = NP, VP, and S_m = Det, N, VP. Then, the production rule that maps S_n to S_m (in our example, NP → Det⁀N, the concatenation of Det and N) defines the two-place asymmetric relation *is-a*(Det⁀N, NP). If our grammar contains a rule of ‘lexical insertion’, whereby Det → {the, a, …} and N → {ball, man, …} (Chomsky 1964: 225–226), then the transition rules of the grammar define (either directly or transitively) the two-place relation *is-a*(*the⁀man*, NP) (Chomsky 1964: 214). In plain English: the grammar specifies that *the⁀man* is an NP. This, in turn, entails that whatever distributional constraints apply to expressions of category NP will apply to the sequence of terminal symbols *the man*. Importantly, this property of pure constituent structure grammars is independent of endocentricity. The relation *is-a* is insensitive to categorial labels, such that

is-a(*the⁀old⁀man⁀arrived*, S) is a well-formed statement in an IC grammar despite there being no terminal symbol of category S in the alphabet of the grammar to which can correspond a phrasal constituent.[5]

We can further define *follows from** to be the transitive closure of *follows from*. Then, we can say that, in Chomsky's definition of *derivation*, S_t *follows from** S_1.

5 Pullum and Miller (2022: 8) observe that

Derivations can be diagrammed by means of trees, called ‘phrase markers’ (P-markers) in linguistics (Chomsky 1955: 180–183, 1959: 144; Postal 1964: 7) and sometimes ‘derivation trees’ in the FLT literature (Hopcroft and Ullman 1969: 18; Mateescu and Salomaa 1997b: 222–223; but also Oehrle 2000: 277), although the term has a different meaning in syntactic theory (specifically within Tree Adjoining Grammars, e.g. Kallmeyer and Joshi 2003: 4). Let us see the FLT definition of *derivation tree*:

> *Let G = (V_N, V_T, P, S) be a cfg* [Context Free Grammar; P stands for ‘productions’]. *A tree is a derivation tree for G if:*
> 1. *Every node has a label, which is a symbol of V_N.*
> 2. *The label of the root* [the unique node in a tree that is not dominated by any other node] *is S.*
> 3. *If a node n has at least one descendant other than itself, and has label A, then A must be in V_N.* (Hopcroft and Ullman 1969: 19)

In grammars of natural languages, the alphabet may contain symbols such the following:

V_N = {S, NP, VP, PP, AP, AdvP, N′, V′, CP, TP, …}

V_T = {the, some, a, funny, dark, obnoxious, table, man, guitar, before, run, kick, want, that, if, sadly …}

That is: lexical items are traditionally considered terminal symbols (thus, basic expressions of the grammar).[6] Sequences of (sequences of) terminal symbols, in NL

> *Chomsky’s early work* [e.g., Chomsky 1956, 1959] *assumed only constituent structure* [see also Postal 1964: Chapter 2]*; X-bar grammars, which incorporate a representation of the ‘head’ relation, represent a kind of hybrid* [between constituent structure and dependency grammar].

Of particular interest in the analysis of structure without headedness may be proposals such as Abels and Neeleman’s (2012) analysis of Universal 20, which makes use of label-less trees: these encode linear order and some notion of constituency, but no prominent notion of ‘head’.

6 We leave aside details of lexical insertion, and assume that lexical items are terminal nodes. However, if terminal nodes in phrase markers stand for bundles of features, then we need a further rule (maybe post-syntactic) that replaces these bundles of features with lexical items. In Chomsky’s (1970: 185) words:

> *the context-free grammar generates phrase markers, with a dummy symbol as* [each] *one of the terminal elements. A general principle of lexical insertion permits lexical entries to replace the dummy symbol in ways determined by their feature content.*

Distributed Morphology and related theories have proposed that syntactic terminals are either bundles of abstract syntactic and semantic features or roots, in both cases acquiring phonological exponents after the syntactic computation (Embick 2015). The relations between recent proposals like grammars, are indexed by members of the set of non-terminal symbols: *the funny man* (a sequence of members of V_T) is categorised as an NP (a member of V_N); we will see how this is accomplished in IC terms shortly.

Early generative grammar, since its inception until approximately the mid-1980s, followed the IC tradition quite closely, for the most part. Chomsky (2006: 172) says

> *The concept of 'phrase structure grammar' was explicitly designed to express the richest system that could reasonably be expected to result from the application of Harris-type procedures to a corpus.* (see also Chomsky 1959: 144, 1964: 214–216)

The crucial insight embodied in the use of 'labels' in IC generative grammar is that they specified distributional properties of the strings they dominated, because categories, or 'morpheme classes' (in the terms of Wells 1947) are defined '*on the basis of the environments in which* [morphemes] *occur*' (Op. Cit.: 81). This means that to assign an expression (a -possibly unary- string of symbols) to a category, it is necessary to determine the distribution of that expression: expressions that have identical distributions will be assigned to the same category. In this context, a category -at any level of linguistic analysis- is essentially a collection of distributional regularities. This process can be extended from individual morphemes to sequences of morphemes (Wells 1947 speaks of 'sequence classes'). The fact that sequence classes are defined distributionally makes it possible to have sequences with the same distribution but different internal structures: Wells gives the example of *they*, which can substitute for a coordination of NPs: *Tom and Harry came* ~ *they came*. In cases like this, Wells calls the 'longer' sequence an *expansion*, and the other, a *model* (such that, in our example, *Tom and Harry* would be an *expansion* of *they*, and *they* would be a *model* of *Tom and Harry*). The aim of an IC approach, in Wells' view, was to analyse every utterance in terms of segments which are expansions and models of other segments. In an example like

4) *The king of England opened Parliament* (Wells 1947: 84)

Wells distinguishes twelve constituents: (i) *the,* (ii) *King,* (iii) *of,* (iv) *England,* (v) *the King,* (vi) *of England,* (vii) *opened,* (viii) *open,* (ix) *-ed,* (x) *Parliament,* (xi) *opened Parliament*, and (xii) *the King of England*. It is interesting to note that Wells does not recognise the whole sentence as a constituent, since the notion is reserved to proper subparts of a sentence (or 'utterance'); it is equally interesting (with some historical perspective) that the morpheme *-ed*, which indicates past tense, is also recognised as a constituent by itself. In the presentation of the goals of an IC analysis, Wells

DM and the early Standard Theory approach to lexical insertion are explored in e.g. Marantz (1997).

We also leave aside multi-word basic expressions, which were identified as early as Jespersen (1985 [1937]), featured in some IC analyses, and have been extensively analysed mostly in non-transformational models.

anticipates a good part of early generative practice in attempting to rule out segmentations that would classify segments such as *King of England opened* as constituents on formal grounds alone (semantic reasons are left for later in his discussion and are not relevant in the present context): each part of the process of segmentation and substitution is based on the idea that a sequence is divided into the segment that we are interested in classifying (the *focus*) and the *environment* in which it occurs. It is the *focus* that needs to be substitutable: *King of England opened* is not substitutable for a *model* such as *John* or *worked*.

The structuralist approach to the definition of categories in terms of segmentation and substitution was greatly developed (and semi-formalised) by Zellig Harris. Harris' *segmentation + substitution* heuristic is summarised thus:

> *we take a form A in an environment C __ D and then substitute another form B in the place of A. If, after such substitution, we still have an expression which occurs in the language concerned, i.e. if not only CAD but also CBD occurs, we say that A and B are members of the same substitution-class, or that both A and B fill the position C __ D, or the like* (Harris 1946: 163)

> *we determine by means of substitution what is the status of the given stretch in respect to the utterance (or to the succession of utterances in the speech): e.g. given the stretch* ***gentlemanly****, we determine that it is a case of A*[djective] *from the fact that it is replaceable by* ***fine****,* ***narrow-minded****, etc. in* ***He's a – fellow****, etc.* (Harris 1951: 279; see also Chomsky 1955: 110)

As observed in Postal (1964: 25–28), Harris' substitution rules are not all of the same level of formal complexity: both context-free and context-sensitive rules are admitted. And, insofar as expansions are reversible, a sequence such as *really quite old* can be substituted for *old* (in more modern terms, the 'head' of the AP). If looked at from the perspective of Chomsky normal grammars, such a relation between an expansion and its corresponding kernel would entail deletion, but we must bear in mind that Harrisian processes are based on distributional equations, not on unidirectional input-output mappings. The aim of IC analyses is to define an exhaustive partition of a sentence into segments that are assigned to distributionally defined classes. These distributional classes, because they are defined based on a segmentation of a string into substrings, depend on an (underlying) relation of contiguity: to use Harris' terms, a 'stretch' (which is individuated by substitution) is a sequence of linearly contiguous 'morphemes' (see e.g. the classes proposed in Harris 1946: §4.1). These can be grouped and analysed at a higher level of description. The analysis of sentences is conducted in terms of classes of distributionally related expressions, not individual expressions; furthermore, substitution classes are defined to include sequences of morphemes as well as individual morphemes (see also Hockett 1954: 215). By defining a set of diagnostics for 'co-occurrents' (morphemes that have sufficiently overlapping sets of possible syntactic contexts where they occur), it is possible to formulate general statements about classes of co-occurrents, as well as operations that combine certain sequences of classes. For example, the sequence of classes T (article) N (noun) P (preposition) N V (verb) corresponds to a description of the utterance the_T $fear_N$ of_P war_N

grew$_V$; in turn, this sequence would be derived by the *substitution* of N in T N V (*the fear grew*) by the sequence N P N (*fear of war*) (Harris 1957).[7]

In Post-style rewriting systems, as adopted by Chomsky, rules operate over members of the alphabet and ultimately produce terminal strings. In defining the derivation of a sentence, it is possible to also define the categories to which terminal sub-strings belong by following the mapping between strings. Therefore, if we have a symbol NP that dominates any other node or set thereof in a derivation tree, it means that whatever set of nodes (of V_N or V_T) is properly contained in the NP can appear within a string whose distributional (and semantic, but these have a very marginal role in IC heuristics) properties are that of an N. In other words: if *a student of linguistics* is an NP, and *of linguistics* is a PP properly contained in the NP, then it means that *of linguistics* is a substring within a string that behaves, distributionally, like any other string **assigned to the category** NP (or **dominated by** NP in a derivation tree, see below). The PP is a **constituent of** the NP. Given the phrase structure rules (PSRs) in (5a), we illustrate the derivation of the terminal string *a student of linguistics* in (5b):

5) a. NP → Det N′
N′ → N (PP)
PP → P NP
N → {student, book, linguistics, …}
Det → {a, the, ∅, …}
P → {of, in, under, …}
b. S_1: NP

7 The reader may have noted that in this case Harris works with a notion of *substitution* that is very much akin to a generalised transformation; indeed, Harris (1957, 1970 [1964]) presents a theory of irreducible, basic sentential types ('kernels') which can be mapped into more complex, derived sentences by means of transformations, which included replacements of symbols by sequences (e.g., replacing N by N P N) as well as to what later became known as *reordering* or *movement rules* (Ross 1967): for example, Harris formulates a relation between sentence forms N_1 V N_2 P N_3 (call it *prepositional indirect object construction*) and N_1 V N_3 N_2 (call it *double object construction*), and says

> *In setting up this relation I have called it 'transformation', since it is a transformation of members of a set into other members of the set, preserving some important properties: It is the rearranging of the words of a sentence from one form into another in which the difference among the sentences of a form, as to acceptability, or as to like properties, is preserved* (Harris 1970 [1964]: 474)

However, Harrisian transformations often involve combining 'elementary transformations', as well as defining 'operators' that act on distributional classes (e.g., *have-en* is an operator that acts on V to yield *have* V-*en*). Furthermore, Harris' work evolved into a theory of string combinations (e.g., Harris 1962; somewhat akin to what would later be known as String Adjunct Grammars; Joshi et al. 1972: §6 for explicit reference to Harris' approach). A fuller treatment of Harris' theory of transformations is unfortunately outside the scope of this paper.

S_2: Det N′
S_3: Det N PP
S_4: Det N P NP
S_5: Det N P Det N′
S_6: Det N P Det N
S_7: a student of Ø linguistics (by lexical insertion, see fn. 6)

If we think about what a phrasal label means in an IC phrase structure grammar, it is just a way to say that a set of substrings behave in a distributionally coherent way: a label defines a set of Harrisian 'co-occurrents'. That is: every string dominated by a node NP, for example, behaves for all relevant intents and purposes 'like an N' (the N being a Wellsian *model* of the NP). A constituent, then, is a sequence of symbols that behaves like a unit for distributional purposes and for the rules of the grammar, which in turn pertain to distribution and configuration. The usual 'constituency tests' (*wh*-movement, pronominalisation, topicalisation, clefting, coordination, etc.; see e.g. Barrie 2022: §3.2; McCawley 1998: Chapter 3; Müller 2020: Chapter 1) apply to sequences of terminals symbols that belong to an indexed category of the grammar, indicated with a nonterminal symbol.[8]

As mentioned above, derivations can be diagrammed as trees (see Chomsky 1959; McCawley 1968; Postal 1964 for discussion). Early IC analyses like Wells' or Harris' did not make use of trees, but since tree diagrams have become ubiquitous in syntactic theorising (transformational or not), something must be said about them. In terms of relations between symbols (nodes) in trees, we can semi-formally define the notion 'constituent' as follows:

8 It is important to note that some of the tests derived from the transformational tradition are category-sensitive, applying to constituents that have a specific categorial label: for example, fronting works for lexical verbs, but not auxiliaries:

i) She said she would read all day, and read she has all day (V fronting)
ii) *She said she can read all day, and can she read all day (Aux fronting)

Furthermore, the relation between certain rules of the grammar and constituency is one of a criterial definition: for example, for the case of topicalisation, Müller (2018) takes it as a 'strong indicator of constituency status', not as an infallible test. That is, if rule R can apply to sequence *s*, then *s* may be a constituent; the fact that R applies is not a sufficient condition, however. Nor is the fact that R does not apply to a specific *s* an unequivocal indicator that *s* is not a constituent (since, as we saw, rules may be category-sensitive). Saying that 'if *s* is affected by R, then *s* is a constituent' is very different from saying that 'a constituent is any *s* that is affected by R': the material implication is reversed. Also relevant in this context is the fact that Harrisian heuristics for constituency are based on distribution, not reorderings: reordering rules (e.g., topicalisation and other frontings) apply to sequences where constituency has already been determined (kernels and combinations of kernels via substitution).

> Let s be a terminal string of arbitrary length: $s = a_1 \frown a_2 \frown a_3 \frown \dots \frown a_n$, where $\{a_1, a_2, a_3, \dots, a_n\} \in V_T$. Let each a_i be a leaf in a phrase structure tree T. Then, s is a **constituent** in T iff s is *exhaustively transitively dominated* by a single symbol $\alpha \in V_N$ in T.

What is **exhaustively transitively dominated**? Let us specify this:

I. We refer to *all* symbols in *s*, and
II. All the parent nodes of all symbols in *s* (that is: the set of all symbols for which the relation **is-a** is defined for members of *s*), and
III. There is no symbol that transitively excludes a symbol in *s* if we follow dominance relations

The representation of hierarchical constituency in terms of exhaustive dominance in phrase structure trees was central in early generative grammar: Postal (1964: 7), for instance, says

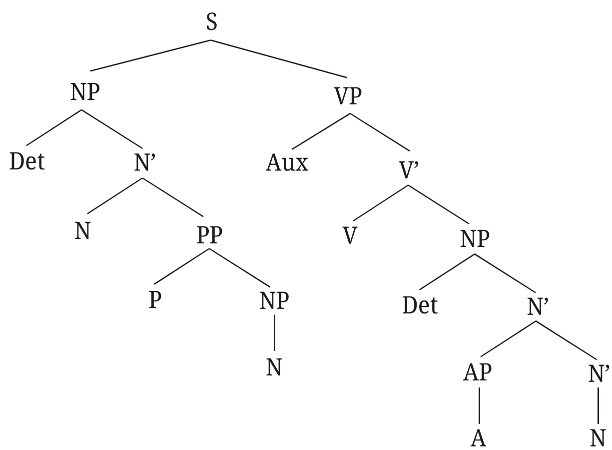
> The fundamental notion in the P-marker is that certain strings of elements are related to certain other single elements by the relation 'is a (member of the category)'

Let us exemplify this idea. Take a sentence like

6) *Some students of linguistics will pass the difficult test*

We may assign to it a phrase marker as in (7) (based on Chomsky 1964, 1970: 210; McCawley 1998; Ross 1967, and others):[9]

7)



9 Variants of this structure are of course possible. For example, we have Chomsky-adjoined AP to the non-lexical, non-phrasal constituent N', but had we followed Jackendoff's (1977) triple-bar system, Det would be in Spec-N''', and AP in Spec-N''. The relevant aspects of the geometry of the tree, however, would remain the same under these alternative analyses.

A phrase marker, graph-theoretically, is a finite set of nodes and edges with the basic relations *dominates*, *precedes*, and *bears-the-label* defined over (pairs of) nodes (McCawley 1968). A rigorous translation between sequences of strings and derivation trees is provided in McCawley (1968: 245), where each symbol in a line of a derivation is made to correspond to a (non-null, possibly unary) sequence of symbols in the line below (cf. Chomsky 1959: 144). As we have emphasised, there are two kinds of information conveyed by a phrase structure tree like (7): first, any set of terminal symbols that is exhaustively dominated by a non-terminal symbol can be referred to by a rule of the grammar. This much may be required of any proof-theoretic formal system. However, since syntactic operations in natural language grammars may be category-sensitive, for linguistic purposes it may also be necessary to indicate what the *head* of a constituent is, should there be one: in this way, having a set of terminal symbols being exhaustively dominated by NP tells us that (i) that terminal sequence is a constituent of the grammar, and (ii) under an *endocentric* approach to IC analysis, the distribution, internal constitution, semantic interpretation, and selectional properties of that constituent will be determined by its *head* (Lyons 1968: 233; Müller 2020.[10] The mechanism behind this is sometimes called *Percolation*, see Zwicky 1985 for discussion from the syntax-morphology interface). The American structuralist and -particularly- generative literatures have emphasised the necessity to consider endocentricity as an important property of grammatical constituents: distributionally, '*the head is the class which can always substitute for the sequence*' (Harris 1946: 180). *Segmentation + substitution*, in contrast to much generative work especially in the 1980s and 1990s, allows for both endocentric and exocentric constructions: Harris gives the examples of T + A = N as exocentric (e.g., the_T $poor_A$ behaves distributionally like an NP) and A + N = N as endocentric (e.g., $smart_A$ $students_N$ behaves distributionally like an NP).

We mentioned above that some syntactic rules are category-sensitive. For instance, *gapping* applies only to sequences containing verbs and possibly auxiliaries:

8) a. *On Wednesdays, Mary practices the piano, and on Tuesdays __ the clarinet*
   b. *Max wanted to order Cheerios and Ann __ Wheaties* (Ross 2012)

The specific labels we use for elements of V_N are not really important for our argument, insofar as they are simply indexed symbols that correspond to distributional categories (Bruening 2020; Hunter 2021): what is important is that, for example, all constituents within the NP are dependents of the N and do not project a phrase on top of N. What a representation like (7) tells us is that, for example, the string *students of linguistics* is a

10 *The head of a constituent/phrase is the element which determines the most important properties of the constituent/phrase. At the same time, the head also determines the composition of the phrase. That is, the head requires certain other elements to be present in the phrase* (Müller 2020: 28). See also Bruening (2009: 30).

constituent (of category N), which shares distribution with other strings, like *professors of linguistics*, *citizens of London*, *books*, etc. in being possible in the context *some*___, since there is a non-terminal symbol that exhaustively dominate that string (namely, N′). In terms of linguistic significance of the category labels chosen, as argued extensively by Bruening (2009, 2020), it is N that can be subcategorised for, and it is properties of N that may be specifically selected by a predicate (cf. also Pullum and Miller 2022). Similarly, the same procedure allows us to determine that *Some students of linguistics will* is not a constituent, because even though there is a symbol that dominates that string (namely, S), this symbol does not *exhaustively* dominate the relevant string: there are symbols excluded from the string that are included in dominance relations (namely, *pass the difficult test*). It also tells us, as required by a descriptively adequate theory of natural language, that a string like *difficult test* is of category N, and not A: the distribution of *difficult test* and its interpretation is that of a (modified) noun, not an adjective. Informally, in IC terms, a *difficult test* is a kind of *test*, not a kind of *difficult*; it can appear in contexts where Ns can appear, not As (e.g., complement to a determiner: *a difficult test*, but **a difficult*; complement to prepositions: *they are studying for difficult tests*, but **they are studying for difficult*; subject position: *difficult tests can frighten students*, but **difficult can frighten students*, etc.). And finally, but not less importantly, (7) tells us that the string *some students of linguistics will pass the difficult test* is a well-formed sentence, a possible 'utterance' (in the Harrisian sense) of English.

There is a further point that we would like to call the reader's attention to: in a tree representation like (7), without further restrictions about the identity of labels or branching stipulations, it is possible to define, for example, a term that includes only the subject and excludes everything else, or a term that includes what traditionally would be called the 'predicate' of the clause (e.g. Chomsky 1965: 71; Rosenbaum 1965: 5), and exclude the subject. The fact that we can say that a sentence contains a subject and a predicate, and that the predicate is neither part of (a constituent of) the subject nor is the subject part of the predicate is a property of a very specific kind of IC analysis, what we will call *pure* IC. Pure ICs are not committed to *a priori* restrictions on the geometry of structural descriptions, because at their core there is the Harrisian mechanism of segmentation + substitution, and there is no principled reason why segmentation should proceed two-by-two or yield exclusively binary segmentations. This allows for segmentations such as (9):

9) Mary bought $[_{NP}$ $[_{NP}$ milk], $[_{NP}$ some apples], and $[_{NP}$ a loaf of bread]]

Coordination is not limited to two terms, since any number of elements may be coordinated; in this case, the terms are all of category NP. Given the fact that there is no asymmetric relation between these NPs, it being simply an enumeration of items, the most natural way to represent this is by having a node NP dominating three other NPs (see e.g. Dalrymple et al. 2019: 216; McCawley 1998: 290; Pollard and Sag 1994). The analysis in (9) captures two properties of the ternary coordination: (i) there is no hierarchical relation between the coordinands (it is a case of *symmetric coordination*), and (ii) a coordination of NPs distributionally behaves like an N (and not, say, like a conjunction; see Borsley 2005 for critical discussion about the grammatical usefulness of structurally uniform, endocentric phrase markers for coordination). The subcategorisation properties of *buy*$_V$ specify [_ +NP], not [_+ConjP]. In this case, the coordination exemplifies an *endocentric* structure. Although intuitive (and, in our opinion, correct), these two properties have not always been maintained in contemporary syntactic models.

The crucial insight of IC 'labels', under 'headedness' analyses, is that they specify distributional properties of the strings they exhaustively dominate (directly or transitively). Therefore, if we have a symbol NP that dominates any other symbol or set thereof, it means that whatever set of symbols (of V_N or V_T) is properly contained in the NP can appear within a string whose distribution and semantic properties are that of an N. Crucially, the Harrisian structuralist approach allows for 'exocentric' constructions at all levels of grammatical analysis: derived constituents whose categorial specification does not coincide with that of any of its component parts. In morphology it is usual to talk of 'exocentric compounds', but in syntax exocentricity is a more contentious property: not all approaches have a clear way of identifying and dealing with exocentric constructions. Perhaps the easiest illustration of this is the rule S → NP VP. What this says is that the concatenation of an NP and a VP does not behave like either; it is a Sentence, which is intuitive enough (see Chomsky 2020: 25 for a recent perspective).

3 Towards array-based grammars

The developments of generative grammar during the late 70s and early 80s brought two theoretical elements to the forefront: (i) the distinction between functional and lexical categories at the core of grammar, and (ii) the idea that all phrasal categories are necessarily *endocentric*. As a refinement of the first point, categories started to be defined in terms of feature matrices: as part of the argument in Chomsky (1970), cross-categorial properties (e.g., both Vs and Ns can take complements) motivated the abandonment of the distributional definition of categories (a point seldom made with the explicitness of Marantz 1997: 214[11]) in favour of a system based on features which only *at first* corresponded to categories sharing distribution (as Rauh 2010: 94

11 Marantz says:

> *Crucial for Chomsky are the consequences* ***of giving up the distributional definition of grammatical categories****. If both Ns and Vs can have complements, and have the head/complement relation interpreted semantically in the same way, then N and V must be distinguished by some internal property, i.e., some feature.* (our highlighting)

observes, the [± N], [± V] system was introduced in Chomsky 1975, not in *Remarks* …). In the same paper the endocentric X-bar system was introduced, whereby phrasal constituents are generated by context-free PSRs of the following form:

10) XP → YP, X′
X′ → X, ZP

where XP, YP, X′, and ZP belong to V_N, and X, Y, Z range over N, V, A (Chomsky 1970: 210). In this system, nominal, verbal, and adjectival phrasal categories are *endocentric*. The system was quickly extended to prepositions and adverbs (Jackendoff 1977). Chomsky (1986: 3) includes the functional categories recognised thus far (Inflection (I) and Complementiser (C)) in the range of the variable X, to which Abney (1987) adds Determiner (D). The expansion of the system of functional categories was coupled with the assumption that selection entailed complementation: for example, I rigidly selects V, thus the projection of V is a complement of the projection of I. The identification of phrasal labels was no longer determined by syntagmatic relations/properties (of the kind reviewed in Section 1). Sentences became endocentric phrases of category IP (Inflection Phrase) with head I, and the S′ of old was now a CP (Complementiser Phrase), both of which phrasal categories satisfy the X-bar template in (10): they have a head, a specifier, and a complement. The clausal structure proposed in Chomsky (1986: 3) (also Stowell 1981: 67) has drastic consequences for the theory of syntactic structure in terms of IC analysis:[12]

11) $[_{CP}\ [_{C'}\ C\ [_{IP}\ NP\ [_{I'}\ I\ [_{VP}\ V\ldots]]]$

In terms of the changes with respect to the previous approach, note that now there is a (non-lexical) category whose phrasal projection properly contains both the subject and the predicate. Every matrix sentence contains a CP layer on top of the IP for purposes of clause typing (Chomsky 1995: 289): a relatively complete sketch of clausal structure in GB for a monotransitive construction would look like (12), below (assuming as well the VP-internal subject hypothesis, the idea that subjects are base-generated within the VP, where they are thematically marked, and move to Spec-IP; Koopman and Sportiche 1991; Zagona 1982):

12) $[_{CP}\ C\ [_{IP}\ NP_{Subj}\ [_{I'}\ I\ [_{VP}\ t_{Subj}\ [_{V'}\ V\ NP_{Obj}]]]]]$

It is important to emphasise that this move already implied a departure from pure IC analyses: the claim that sentences are IPs is not based on distributional information

12 The structure in (11) is not exclusive of transformational generative grammar: Lexical Functional Grammar assumes the same clause structure for 'configurational' languages (languages where structural positions map reliably to grammatical functions). See e.g. Westcoat (2005), Bresnan et al. (2016), and Dalrymple et al. (2019).

and certainly not on segmentation: inflection is a grammatical category that surfaces (when it does) as an affix or an auxiliary. We will look at the consequences of this claim in detail.

The extension of the X-bar schema to functional categories made it possible to propose functional categories in order to accommodate word order phenomena (assuming X-bar to deliver planar trees, an assumption derived from the classical definition of PSRs as mappings between strings) and assign a phrasal category to each of those. In Chomsky (1986), I contained two binary features: [± Tense] and [± Agreement]. Linguists such as Pollock (1989) and Belletti (1990) argued that each of these features should project its own phrasal category:[13] a Tense Phrase and an Agreement Phrase. In the context of a theory of Case assignment where structural Case (Nominative, Accusative) was assigned to NPs in a Spec-Head relation with specific functional projections, the idea was that AgrP were in charge of checking structural Case features. Two AgrP were proposed, one for Nominative Case (AgrS) and one for Accusative Case (AgrO). The clausal skeleton was thus expanded from (12) to something along the lines of (13):

13) [AgrSP [TP [AgrOP [VP]]]

In turn, Larson's (1988) work on the double object construction (which, as Larson 2014: 2–5 acknowledges, represent an attempt to combine phrase structure with categorial grammar) expanded the VP into two layers, or 'shells'. Let us consider the structure of a prepositional indirect object construction such as *Mary gave a book to John*. The first 'shell' has as its specifier the external argument (EA) of a (di)transitive clause, a functional affixal null head, and another VP as its complement, this time with a lexical head; the second shell has as its specifier the direct object (OBJ) and the indirect object (OBL) as its complement:

14) [VP_1 NP_{EA} [V_1 *e* [VP_2 NP_{OBJ} [V_2' V [PP_{OBL}]]] (see also Chomsky 1995: 180 and much related work).

The lexical verb V_2 raises to V_1 (Larson 1988: 343) to deliver the right word order: again, the planarity of tree structures comes to the foreground. It is customary now to call the high, functional VP layer *v*P[14] ('little V') and leave the label VP for the lower, lexical layer. We thus expand (13) accordingly:

15) [AgrSP [TP [AgrOP [*v*P [VP]]]]

13 The extreme form of this approach is the so-called 'One Feature, One Head' hypothesis (Baunaz and Lander 2018; Caha 2020 and references therein).

14 Alternative nomenclatures include VoiceP (Kratzer 1996) and EventP (Harley 1995).

We need to look at these structures more closely. If we apply the definition of constituency that we introduced in the previous section, it follows that everything that is dominated by AgrS, for example, is a constituent of category AgrS. And because AgrS was the highest functional head in the clausal skeleton, the maximal projection AgrSP effectively dominated the whole (declarative) sentence: verbs like *seem* would not select for sentential complements, but for AgrSP complements. The motivation for AgrP is *not* distributional: what predicate *selects* for Agr (i.e., does Agr feature in the subcategorisation frame of any predicate)? Nor is it semantic, for that matter: what is the semantic contribution of Agr? Does AgrP denote a proposition? Can it be assigned a truth value? The answer seems to be a straightforward 'no': their uninterpretability was one of the reasons why Agr projections were eliminated in early Minimalism (Chomsky 1995: Chapter 4). As far as Larsonian shells are concerned, Davies and Dubinsky (2001: 5) observe that

> *Larson's approach is developed further in Pesetsky (1995), where VPs having nonlexical heads* [. . .] *are done away with and replaced by lexical projections (sometimes having phonologically null heads).* ***The essence of these proposals was to divorce phrase structure from traditional constituency tests,***[15] *in favor of representations in which thematic hierarchies map directly to a syntactic analog* [. . .] (our highlighting)

It would not be fair or true to say that these developments were unmotivated, however. The analysis of cross-linguistic evidence, interpreted through the lens of the hypothesis that functional categories projected phrases and that syntactic objects could move (either overtly or covertly) led to the multiplication of functional heads and movement operations such that, given a universal hierarchy of projections,

15 Indeed, Pesetsky's (1995) 'cascades' also represent a drastic departure from IC assumptions, but more so in terms of the configuration than the labels used. For example, he assigns a 'cascade' structural description like (ii) to the VP in (i):

i) give books to these people on each other's birthdays

ii) [$_{\text{VP}}$ give [$_{\text{PP}}$ [$_{\text{DP}}$ books] [$_{\text{P}'}$ to [$_{\text{PP}}$ [$_{\text{DP}}$ these people] [$_{\text{P}'}$ on [$_{\text{DP}}$ each other's birthdays]]]]]] (Pesetsky 1995: 174)

The structure is strictly binary branching, with no distinction between arguments and adjuncts. Note that the string *to these people* is not even a constituent. The departure from IC here is not related to the proliferation of functional projections: it is the revised structure of PPs that makes (ii) unsuitable as an IC analysis. In contrast to *cascades*, Pesetsky presents *layered* derivations: these do have *to NP* as a constituent, but in his terms do not provide the right structures for coordination and binding. The analysis of ditransitives in Collins (2021: 102) also fails to recognise *to NP* as a constituent:

iii) [$_{\text{TP}}$ [$_{\text{DP}}$ John] [$_{\text{T}'}$ T [$_{v\text{P}}$ <John> [$_{v'}$ [*v*+*give*$_{\text{V}}$] [$_{\text{AgrOP}}$ [$_{\text{DP}}$ the car] [$_{\text{AgrO}'}$ AgrO [$_{\text{VoiceP}}$ VP [$_{\text{Voice}'}$ to [$_{\text{ApplP}}$ [$_{\text{DP}}$ Mary] [$_{\text{Appl}'}$ Appl<[$_{\text{VP}}$ give [$_{\text{DP}}$ the car]]>]]]]]]]]]

different languages may either order the heads differently (see Laka 1990 for Negation) or move syntactic objects at different derivational points (Chomsky 1995; Pollock 1989). Let us briefly review the arguments.

A basic condition on the movement of heads is that a head can only move to the head position that immediately c-commands it[16] (the Head Movement Constraint HMC, Chomsky 1986: 71; Travis 1984: 131). The HMC determines that head movement is strictly local, and that a head cannot move 'jumping over' an intermediate head: X cannot move to Y if Z c-commands X and is c-commanded by Y. If we go back to the clausal structure in (11), we see that V cannot move to C unless it moves to I first: otherwise, movement from V-to-C would violate the HMC. The GB proposal was that languages with rich inflectional morphology (such as Spanish or Italian) allowed the lexical verb to move out of the VP to I (later, Agr), from where it could keep moving towards C. Languages like English do not have Subject-Verb inversion: only auxiliaries may invert (and when there is a sequence of auxiliaries, only the first one may invert). Thus, instead of (16a) we have (16b):

16) a. **What wants John?*
b. *What does John want?*

Because English lexical verbs cannot move to I (due to its inflectional poverty, English I is 'weak', and cannot host lexical verbs; Chomsky 1995: 135), a dummy auxiliary *do*-support must be inserted as a 'last resort' to have something to move to C.

In this context, consider the following paradigm, from Pollock (1989: 367):

17) a. **John likes not Mary.*
b. *Jean (n') aime pas Marie.* (= a)
J. NEG loves NEG M.
c. **Likes he Mary?*
d. *Aime-t-il Marie?* (= c)
Love-EUPH-Prn.3SG M.
e. **John kisses often Mary.*
f. *Jean embrasse souvent Marie.* (= e)
J. kiss.3SG often M.
g. *John often kisses Mary.*
h. **Jean souvent embrasse Marie.* (= g)

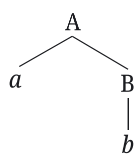
The basic idea is that, given a clausal hierarchy T(ense) > Agr(eement) > VP, as assumed by Pollock, in French V moves to Agr, whereas in English Agr lowers to V. Furthermore, negation heads its own projection, NegP, which dominates Agr and is

16 The terminology used in these works is *properly governs*, but -essentially- it boils down to c-command and a locality condition.

dominated by T. This, in Pollock's view, yields the correct word order given planar X-bar trees. And since -in Chomsky's view- head movement does not affect meaning, there is no prediction that raised verbs are interpreted in any way differently from non-raised verbs (Chomsky 2001: 37–38).

These developments led to the so-called 'expansion' of the 'functional skeleton' of the clause. Lexical categories were dominated by functional categories, such that I/T > V and, analogously, D > N (Abney 1987). Expanding I into AgrS, T, and AgrO results in a more articulate clausal skeleton, determined by Universal Grammar. The core properties that are captured by a system like this are *order* and *hierarchy*, and in such a way that certain hierarchical relations completely determine word order. The classical perspective, derived from phrase markers being diagrams of mappings between strings, was based on tree planarity. However, planarity is but one way to 'linearise' terminals in phrase structure trees. The proposal in Kayne (1994) is based on translating structural relations in X-bar compliant trees onto precedence. Given a tree diagram, *a* c-commands *b* if the first branching node that dominates *a* also dominates *b* and neither *a* nor *b* dominate the other (Reinhart 1976). The term *asymmetric* c-command defines a relation in which *a* c-commands *b* but *b* does not c-command *a*. In (18),

18)

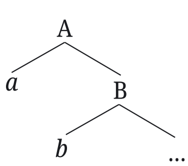


a c-commands B and *b*, because A, the first branching node that dominates *a*, also dominates B (and, because B dominates *b*, A dominates *b* transitively). If the categorial specification of A is identical to the categorial specification of B, then we would say that we have a two-segment category. Note that *a* and B c-command each other, but *b* does *not* c-command *a*: we say in this case that *a asymmetrically* c-commands *b*. This relation is particularly relevant. Kayne (1994: 3) states that phrase structure '*always completely determines linear order*', and it does so by translating asymmetric c-command relations between terminals and nonterminal nodes into linear precedence Kayne (1994: 3) relations: if *a* asymmetrically c-commands *b*, then *a* precedes *b*. Linear order cannot vary independently of structure. This principle is used to restrict the class of allowed trees as much as to linearise structure: any point of mutual c-command between terminals or phrases means trouble for linearisation, and thus must be either dissolved by moving one of the terms or not generated in the first place (Kayne 2022). Specifically, Kayne's theory has led to the claim that '*Arguments invariably* [even if they appear to surface 'in situ'] *raise at least once, in all languages*' (Kayne 2022: 2), and, for example, that V and O are never sister nodes derivation-finally, and suggests that objects may never be directly merged with lexical verbs '*as* *head and complement*' (Op. Cit.: 3). Cases of apparent VP fronting (e.g., *read his paper, I never will*) must be dealt with via multiple instances of movement, without there ever being a constituent [V [Object]].

This approach resulted in the emergence of theories that aimed at defining the specific order of functional heads as defined by UG such that a fixed universal hierarchy could be formulated. In general, the clause is divided in three or four domains, each of which is given its own sequence of functional heads: (i) the VP domain, related to event structure, (ii) the TP domain, related to Tense, Aspect, Modality, and Agreement, (iii) the CP domain, related to illocutionary force, sentential operators, discourse properties, and (iv) the NP/DP. A more underspecified approach is defended in Borer (2005), where a sequence of aspectual functional heads introduce all arguments, as opposed to having them be syntactic dependents of lexical categories. An (evidently non exhaustive) overview of almost 40 years of syntactic research within generative grammar reveals that the exploration of functional categories has given rise to a number of more or less highly articulated 'functional skeletons' for clause structure, of which we can provide some examples.[17] In (19) below, > is a strict ordering, such that

A > B

corresponds to the tree notation



where a phrase of category A takes a phrase of category B as its complement (we abbreviate internal structure of A and B, omitting e.g., specifiers).

19) a. Chomsky (1986): CP > IP > VP (Speas and Fukui 1986, among many others; also Bresnan et al. 2016; Dalrymple et al. 2019)

b. Pollock (1989), Belletti (1990), and Chomsky and Lasnik (1993): AgrSP > TP > AgrOP > VP

17 Whereas most of these examples are taken from the cartographic literature, it is important to emphasise that this paper is neither a critique of nor concerned exclusively with cartography. Contemporary 'orthodox' Minimalist structures, as well as Exoskeletal models, share the same formal properties as cartographic structures (in terms of defining arrays), for all current intents and purposes. There are also theories that remain agnostic about these arrays, and are compatible with either strict IC or arrays (for instance, Tree Adjoining Grammars).

c. Löbel (1991): KP > DP > NP
d. Cinque (1994) DP > Poss > Cardinal > Ordinal > Speaker-oriented > Subject-oriented > NP
e. Rizzi (1997): Force > Topic* > Focus > Topic* > Finiteness > IP (* = Kleene star)
f. Chomsky (2000): CP > TP > *v*P > VP (see also Adger 2003; Hornstein et al. 2005; Radford 2009, among many others)
g. Chomsky (2020): <φ, φ> > TP > *v*P > VP
h. Déchaine and Wiltschko (2002): DP > φP > NP
i. Scott (2002): DP > Adj$_{\text{poss}}$ > Adj$_{\text{card}}$ > Adj$_{\text{ord}}$ > Adj$_{\text{qual}}$ > Adj$_{\text{size}}$ > Adj$_{\text{shape}}$ > Adj$_{\text{color}}$ > Adj$_{\text{nation}}$ > NP
k. Ticio (2010): FocP > DP > AgrP > *n*P > NP
l. Adger (2003): DP > *n*P > NP
m. Cardinaletti (2004): ForceP > SubjP > AgrSP > TP > VP
o. Cinque (2004) and Rizzi and Cinque (2016): MoodP$_{\text{speech act}}$ > MoodP$_{\text{evaluative}}$ > MoodP$_{\text{evidential}}$ > ModP$_{\text{epistemic}}$ > TP$_{\text{(Past)}}$ > TP$_{\text{(Future)}}$ > MoodP$_{\text{irrealis}}$ > ModP$_{\text{alethic}}$ > AspP$_{\text{habitual}}$ > AspP$_{\text{repetitive(I)}}$ > AspP$_{\text{frequentative(I)}}$ > ModP$_{\text{volitional}}$ AspP$_{\text{celerative(I)}}$ > TP$_{\text{(Anterior)}}$ > AspP$_{\text{terminative}}$ > AspP$_{\text{continuative}}$ > AspP$_{\text{retrospective}}$ AspP$_{\text{proximative}}$ > AspP$_{\text{durative}}$ > AspP$_{\text{generic/progressive}}$ > AspP$_{\text{prospective}}$ > ModP$_{\text{obligation}}$ ModP$_{\text{permission/ability}}$ > AspP$_{\text{Completive}}$ > VoiceP > AspP$_{\text{celerative(II)}}$ > AspP$_{\text{repetitive(II)}}$ > AspP$_{\text{frequentative(II)}}$ > **verb**
p. Cinque (2005): Agr$_{\text{W}}$P > DemP > Agr$_{\text{X}}$P > NumP > Agr$_{\text{Y}}$P > AP > NP
q. Svenonius (2008): Dem > Art > Num > unit > Pl/sort > Adj > *n* > N
r. Caha (2020): ComitativeP > InstrumentalP > DativeP > GenitiveP > AccusativeP > NominativeP > root

The crucial point for this paper is the following: having a sequence C > T > V for clause structure (where each of these is a shorthand for a more articulate set of functional projections) forces us to give up the classical heuristic for constituency, and indeed immediate constituency, in a way. Why? Because it is not clear what it means to say that in *John read a book*, with a structural description along the lines of [$_{\text{CP}}$ Ø [$_{\text{TP}}$ John [$_{\text{T'}}$ T [$_{v\text{P}}$ ~~John~~ [$_{v'}$ read [$_{\text{VP}}$ ~~read~~ [$_{\text{DP}}$ a book]]]]]]] (following the order of projections in Chomsky 2000), the sequence *read a book*, which is exhaustively dominated by a segment of T, *is-a* term of category T (a phrasal category TP). Does it have the distribution of Tense? Is it Tense that features in subcategorisation frames of verbs taking clausal arguments? Are immediate constituents in TP and their properties subcategorised by T?

It is important to note that the status of tense morphemes as constituents was indeed observed e.g. by Wells (1947) and Harris (1946, 1951); however, this does not mean that a category Tense 'projects' a phrasal constituent which *properly contains* a lexical verb and its arguments. The structuralist IC approach allowed for the recognition of lexical constituents without the necessity that there be a phrasal constituent corresponding to each. The extension of the X-bar schema to functional categories, and later developments in the theory made it impossible to recognise a segment without also assigning to it not just a category (as in Harrisian IC), but also the role of phrasal head. There is, in this sense, a fundamental difference between (20a) and (20b):

20) a. $[_{S}$ $[_{NP}$ the man] $[_{T/Aux}$ has] $[_{VP}$ refused $[_{NP}$ the present]]]
b. $[_{TP}$ $[_{NP}$ the man] $[_{T}$ has $[_{VP}$ refused $[_{NP}$ the present]]]]

Even if there may be agreement about what category each segment belongs to (in some cases), the implications of each structural description for grammatical analysis are vastly different. In (21a) there is a sentence that properly contains an NP subject, an auxiliary, and a VP (e.g. Chomsky 1965, 1970; Emonds 1970 also generates T morphemes as immediate constituents of S); a variant of this structure considers that the auxiliary is an immediate constituent of the VP (Ross 1967), not of S. An antecedent of the Rossian analysis includes the PSRs in Chomsky (1964: 224):

21) Sentence → NP + VP (*Adv*)

$$\text{VP} \rightarrow \text{Aux} \left\{ \begin{matrix} be \left\{ \begin{matrix} Pred \\ Adv_1 \end{matrix} \right\} \\ VP_1 \end{matrix} \right\}$$

The node Aux, in turn, would dominate Tense, modals, perfective, progressive, and passive auxiliaries: if we consider a transitive sentence, at the level of phrase structure there is a node that includes only the auxiliary and excludes the V and its direct object. These two form a phrasal unit that excludes the auxiliary and which is of category V.

Let us try to read (20b) with IC glasses. Due to being dominated by all segments of T, the VP would be an immediate constituent of a phrasal category T, and so would the subject (by virtue of being the Specifier of TP). Recall that in IC terms, syntactic categories are proxies for distributional properties (which also provides the heuristics for the definition of new categories should these be required), and under endocentricity and percolation the properties of phrasal constituents are determined by the properties of their head. Then, the distribution of the sequence *the man has refused the present* would be determined by (note: *not* identical to) the distribution of the terminal *has*, since it bears the label TP: whatever head can select for *has* (e.g., a modal), should also select for *the man has refused the present*, since their categorial specification is the same (the difference being only one of lexical vs. phrasal status). This is the same reasoning that applies when saying that verbs select for NPs, regardless of their internal structure: they may or may not contain Num(eral) heads, Part(itive) heads, D heads, but these are all dominated by segments of an NP (Bruening 2009, 2020).[18] While clearly not what the structural description intends, it seems to be the only possible IC interpretation. If it is not inconsistent, then, (20b) must be taken to represent a kind of information different from constituency in the classical sense: labels in contemporary generative grammar (the proposals summarised in (19)) do not correspond to distributional classes, nor do they allow for expansions and reductions of the classical Harrisian kind.

Evidently, reading immediate constituency from contemporary Minimalist, Nanosyntactic, Exoskeletal, or Cartographic trees is not tenable even from a pre-theoretic perspective. This is also true of non-transformational models, such as LFG, which have adopted the C > T/I > V clause structure for configurational languages. Arguably, furthermore, this has been true since the mid-1980s. Our claim is that the only way to interpret a sequence of functional phrases in a coherent way is to acknowledge the fact that they do not represent constituency at all in any classical way. We are not in presence of sequences of sequences of symbols ordered by the relation *follows-from*: a representation like any of those in (19) does *not* define the relation *is-a* for any pair Symbol-Sequence. Most importantly, *it does not aim to*. This is a crucial point, we think, that has flown under the radar of much contemporary theorising. Generative grammar, in a sharp departure from its origins, is no longer based on phrase structure grammar, nor is it equivalent to one. Distributional and selectional arguments of the kind defined in Section 1 play little if any role in contemporary generative syntactic argumentation: to the extent that selection is considered at all, it pertains to the relation between functional and lexical categories, and is assumed to correlate with projection such that, if A (*f*- or *c*-) selects B, then A projects. Selection is the main criterion (together with linear order of terminals, especially in cartographic analyses) used to sort the arrays in (19), which in the strongest interpretation are fixed-size arrays (with all projections always present even if not spelled out in a given language). Thus, for instance, the fact that determiners select nominals is taken by Larson (2020: §1.2) to mean that determiners project a DP. Selection and projection, however, are logically independent, and their correlation is a property of a specific theory of phrase structure. A weaker

18 Grimshaw (2000: 124–125) observes the problems that a uniformly Functional Category > Lexical Category brings about for selection, mentioning specifically the DP hypothesis: if D is the head of a DP dominating NP, a verb cannot locally select for properties of N, as D intervenes. The notion of *s-projection* in Abney (1987) and Grimshaw's own *extended projection* attempt to recapture locality by making all relevant properties of N project to DP, making N effectively count as the head of DP.

interpretation would allow these arrays to be dynamic arrays, where a subset of elements (here, projections) is used in a specific language or even a specific derivation: here, we can call the size of the underlying (universal) array the array's 'capacity' or 'physical size', determined by Universal Grammar.

As the theory became more complex, and functional categories flourished, generative syntax became grounded on strictly ordered sequences of functional expressions on top of lexical projections (in some versions of the theory, every projection has its own functional layer on top; e.g., Cinque 2023: 6). A stricter theory of syntactic configuration, where binarity is a foundational principle and structure building operations are concomitant to ‘labelling’ (Chomsky 2013, 2020) delivers structural descriptions organised as (sorted) *arrays*: one-dimensional ordered sequences of indexed addresses, where the source of ordering is given by Universal Grammar. In the case of the hierarchies in (19), each of these indexed addresses is a syntactic label drawn from a universal inventory and strictly ordered according to universal principles. Because of their departure from IC analysis, these structures can implement strictly monotonically growing spines of functional material: in addition to strict binarity (for theories that use tree diagrams, we speak of ‘binary branching’, but contemporary Minimalism contends that the generative operation Merge delivers unordered sets, not trees; see e.g. Chomsky 2013; Chomsky et al. 2023), contemporary generative grammar either strongly prefers or only delivers local structures of the form {head, non-head} at every derivational step (Chomsky 2013 et seq.; Cinque 2023; Kayne 2022, among many others). This commitment to monotonic growth generates structures that, as shown in Uriagereka (2012), are expressible in finite-state fashion without loss of information: this would not be possible under IC assumptions. Interestingly, Postal (1964) shows that IC systems of the kind formulated by Wells, Harris, and Hockett are all equivalent to context-free phrase structure grammars (and, like Chomsky, proceeds to emphasise the inadequacies of PSGs, advocating for more powerful transformational grammars).

4 Conclusions

We examined the IC assumptions that underpinned much work on phrase structure in the early days of generative grammar, as the influence of Harrisian structural syntax was still strong (despite goals and discovery methods differing, e.g. Chomsky 1964: 211–212, 215–216). The distinction between lexical and phrasal constituents was central to the theory, and phrase markers were seen as representations of more traditional conceptions of hierarchical relations between expressions. The goal of linguistic theory was mostly ‘projective’ (borrowing the term from Haider 1996): given a set of expressions, the grammar provides a unique structural description to each (reading of an) expression. These structural descriptions took the form of phrase markers, which are themselves illustrations of derivations: ordered sequences of strings related by the asymmetric relation *follows-from*, and which define the asymmetric relation *is-a* for substrings and members of V_N. This relation was seen then (and, we add, still is) as fundamental to correctly represent natural language structure in terms of IC analyses. Subsequent changes in the theory, in particular from the 1980s onwards, introduced other concerns in syntactic analysis, which progressively shifted generative grammar (transformational and non-transformational) from a theory based on IC analysis to a new kind of formalism that aims at defining a universal hierarchical sequence of categories in structural descriptions. In doing so, we argue, phrase markers became diagrams for *arrays*. Identifying and examining generative grammar's departure from IC analysis is the first step in a wider research agenda: that of examining more closely some technical and empirical consequences of this shift, and the new questions that become formulable under array-based syntax.

Acknowledgments: This paper owes its existence to an email exchange with Haj Ross (in 2021) about what it means to be a Harrisian grammarian. To Haj it is, then, humbly dedicated. We would like to thank three anonymous *Folia Linguistica* reviewers as well as Adam J. R. Tallman and Susan F. Schmerling for their comments and observations. Our deepest gratitude does also to Olga Fischer, Sune Gregersen, and Viviana Masia for their editorial work. The usual disclaimers apply.

References

Abels, Klaus & Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1). 25–74.

Abney, Steven. 1987. *The English noun phrase in its sentential aspect*. Cambridge, Massachusetts: MIT PhD thesis.

Adger, David. 2003. *Core syntax: A minimalist approach*. Oxford: OUP.

Barrie, Michael. 2022. *Generative syntax*. Berlin: Language Science Press.

Baunaz, Lena & Eric Lander. 2018. Nanosyntax: The basics. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring nanosyntax*, 4–56. Oxford: OUP.

Belletti, Adriana. 1990. *Generalized verb movement*. Torino: Rosenberg & Sellier.

Bloomfield, Leonard. 1926. A set of postulates for the science of language. *Language* 2(3). 153–164.

Borer, Hagit. 2005. *The normal course of events*. Cambridge: CUP.

Borsley, Robert. 2005. Against CoordP. *Lingua* 115. 461–482.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical functional syntax*, 2nd edn. Oxford: Wiley Blackwell.

Bruening, Benjamin. 2009. Selectional asymmetries between CP and DP suggest that the DP hypothesis is wrong. *University of Pennsylvania Working Papers in Linguistics* 15(1). 5.

Bruening, Bejnamin. 2020. The head of the nominal is N, not D: N-to-D movement, hybrid agreement, and conventionalized expressions. *Glossa* 5(1). https://doi.org/10.5334/gjgl.1031.

Caha, Pavel. 2020. Nanosyntax: Some key features. Ms. Available at: https://ling.auf.net/lingbuzz/004437.

Cardinaletti, Anna. 2004. Toward a cartography of subject positions. In Luigi Rizzi (ed.), *The structure of CP and IP*, 115–165. Oxford: OUP.
Chomsky, Noam. 1951. *Morphophonemics of modern Hebrew*. Pennsylvania: University of Pennsylvania MA Thesis [cited by the 1979 version, New York: Plenum].
Chomsky, Noam. 1955. *The logical structure of linguistic theory*. Mimeographed, MIT. [cited by the 1975 version, New York: Plenum].
Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2. 113–124.
Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and Control* 2. 137–167.
Chomsky, Noam. 1964. A transformational approach to syntax. In Jerry Fodor & Jerrold Katz (eds.), *The structure of language*, 211–245. New Jersey: Prentice Hall.
Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1970. Remarks on nominalization. In Roderick Jacobs & Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham: Ginn & Co.
Chomsky, Noam. 1975. Questions of form and interpretation. *Linguistic Analysis* 1(1). 75–109.
Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
Chomsky, Noam & Howard Lasnik. 1993. The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternfeld & Theo Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research*, 506–569. Berlin: de Gruyter. [Cited by the reprint in Chomsky (1995), 13–127.]
Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step – Essays in minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2006. *Language and mind*. Cambridge: CUP.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.
Chomsky, Noam. 2020. UCLA lectures. Available at: https://ling.auf.net/lingbuzz/005485.
Chomsky, Noam, T. Daniel Seely, Robert Berwick, Sandiway Fong, M. A. C. Huybregts, Hisatsugu Kitahara, Andrew McInnerney & Yushi Sugimoto. 2023. *Merge and the strong minimalist thesis*. Cambridge: CUP.
Cinque, Guglielmo. 1994. On the evidence for partial N-movement in the Romance DP. In Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi & Raffaella Zanuttini (eds.), *Paths towards universal grammar*, 85–110. Washington, D.C.: Georgetown University Press.
Cinque, Guglielmo. 2004. Restructuring' and functional structure. In Guglielmo Cinque (ed.), *Restructuring and functional heads*, 132–192. Oxford: OUP.
Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36(3). 315–332.
Cinque, Guglielmo. 2023. *On linearization*. Cambridge, Mass.: MIT Press.
Collins, Chris. 2021. A smuggling approach to the dative alternation. In Adriana Belletti & Chris Collins (eds.), *Smuggling in syntax*. Oxford: OUP.
Dalrymple, Mary, John Lowe & Louise Mycock. 2019. *The Oxford reference guide to lexical functional grammar*. Oxford: OUP.
Davies, William & Stanley Dubinsky. 2001. Remarks on grammatical functions in transformational syntax. In William Davies & Stanley Dubinsky (eds.), *Objects and other subjects*, 1–19. Dordrecht: Kluwer.

Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33(3). 409–442.
Embick, David. 2015. *The morpheme*. Berlin: de Gruyter.
Emonds, Joseph. 1970. *Root and structure-preserving transformations*. Cambridge, Massachusetts: MIT PhD thesis.
Grimshaw, Jane. 2000. Locality and extended projection. In Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.), *Lexical specification and insertion*, 115–133. Amsterdam: John Benjamins.
Haider, Hubert. 1996. Economy in syntax is projective economy. In Chris Wilder, Hans-Martin Gärtner & Manfred Bierwisch (eds.), *The role of economy principles in linguistic theory*, 205–226. Berlin: de Gruyter.
Harley, Heidi. 1995. *Subjects, events, and licensing*. Cambridge, Massachusetts: MIT PhD thesis.
Harris, Zellig. 1946. From morpheme to utterance. *Language* 22. 161–183.
Harris, Zellig. 1951. *Methods in structural linguistics*. Chicago: University of Chicago Press.
Harris, Zellig. 1957. Co-occurrence and transformation in linguistic structure. *Language* 33. 283–340.
Harris, Zellig. 1962. *String analysis of sentence structure*. The Hague: Mouton.
Harris, Zellig. 1964. Transformations in linguistic structure. *Proceedings of the American Philosophical Society* 108(5). 418–422. Cited by the reprint in Harris (1970), 472–481.
Harris, Zellig. 1968. *Mathematical structures of language*. New York: Wiley & Sons.
Harris, Zellig. 1970. *Papers in structural and transformational linguistics*. Dordrecht: Reidel.
Hockett, Charles. 1954. Two models of grammatical description. *Word* 10. 210–234.
Hopcroft, John & Jeffrey Ullman. 1969. *Formal languages and their relation to automata*. London: Addison-Wesley.
Hornstein, Norbert, Jairo Nunes & Kleanthes Grohmann. 2005. *Understanding minimalism*. Cambridge: CUP.
Hunter, Tim. 2021. The Chomsky hierarchy. In Nicholas Allott, Terje Lohndal & Georges Rey (eds.), *A companion to Chomsky*, 74–95. Oxford: Blackwell.
Jackendoff, Ray. 1977. *X-bar syntax*. Cambridge, Mass.: MIT Press.
Jespersen, Otto. 1985. [1937]. *Analytic syntax*. Chicago: University of Chicago Press.
Joshi, Aravind, Rao Kosaraju & H. Yamada. 1972. String adjunct grammars II. *Information and Control* 21. 235–260.
Kallmeyer, Laura & Aravind Joshi. 2003. Factoring predicate argument and scope semantics: Underspecified semantics with LTAG. *Research on Language and Computation* 1. 3–58.
Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.
Kayne, Richard. 2022. Antisymmetry and externalization. *Studies in Chinese Linguistics* 43(1). 1–20.
Koopman, Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85. 211–258.
Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 109–137. Dordrecht: Kluwer.
Krivochen, Diego Gabriel. 2023. *Syntax on the edge*. Leiden: Brill.
Laka, Itziar. 1990. *Negation in syntax*. Cambridge, Massachusetts: MIT PhD thesis.
Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19. 335–391.
Larson, Richard. 2014. *On shell structure*. London: Routledge.
Larson, Richard. 2020. The DP hypothesis and (a)symmetries between DP and CP. *Linguistic Analysis* 42(3–4). 1–41.
Löbel, Elisabeth. 1991. KP/DP syntax. *Theoretical Linguistics* 20(1). 37–70.
Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: CUP.
Marantz, Alec. 1997. No escape from syntax. *UPenn Working Papers in Linguistics* 4(2). 201–225.

Mateescu, Alexandru & Arto Salomaa. 1997a. Formal languages: An introduction and a synopsis. In Grzegorz Rozenberg & Arto Salomaa (eds.), *Handbook of formal languages*, 1–40. Berlin: Springer.
Mateescu, Alexandru & Arto Salomaa. 1997b. Aspects of classical language theory. In Grzegorz Rozenberg & Arto Salomaa (eds.), *Handbook of formal languages*, 175–252. Berlin: Springer.
McCawley, James. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4(3). 243–269.
McCawley, James. 1998. *The syntactic phenomena of English*, 2nd edn. Chicago: University of Chicago Press.
Müller, Stefan. 2018. Evaluating theories: Counting nodes and the question of constituency. *Language Under Discussion* 5(1). 52–67.
Müller, Stefan. 2020. *Grammatical theory*. Berlin: Language Science Press.
Oehrle, Richard. 2000. Context-sensitive node admissibility. *Grammars* 3. 275–293.
Pesetsky, David. 1995. *Zero syntax*. Cambridge, Mass.: MIT Press.
Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Head-driven phrase structure grammar. Chicago: University of Chicago Press.
Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.
Post, Emil. 1943. Formal reductions of the general combinatorial decision problem. *American Journal of Mathematics* 65(2). 197–215.
Postal, Paul. 1964. *Constituent structure: A study of contemporary models of syntactic description*. Bloomington: Indiana University Press.
Postal, Paul. 2010. *Edge-based clausal syntax*. Cambridge, Mass.: MIT Press.
Pullum, Geoffrey. 2019. What grammars are, or ought to be. In Stefan Müller & Petya Osenova (eds.), *Proceedings of the 26th international conference on head-driven phrase structure grammar*, 58–78. Stanford: CSLI.
Pullum, Geoffrey & Philippe Miller. 2022. NPs versus DPs: Why Chomsky was right. Ms. Available at: https://lingbuzz.net/lingbuzz/006845.
Radford, Andrew. 2009. *Analyzing English sentences*. Cambridge: CUP.
Rauh, Gisa. 2010. *Syntactic categories*. Oxford: OUP.
Reinhart, Tanya. 1976. *The syntactic domain of anaphora*. Cambridge, Massachusetts: MIT PhD thesis.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer.
Rizzi, Luigi & Guglielmo Cinque. 2016. Functional categories and syntactic theory. *Annual Review of Linguistics* 2. 139–163.
Rosenbaum, Peter. 1965. *The grammar of English predicate complement constructions*. Cambridge, Massachusetts: MIT PhD thesis.
Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, Massachusetts: MIT PhD thesis.
Ross, John Robert. 2012. A preliminary –but fattened- list of transformations. Ms. Available at: http://www-personal.umich.edu/~jlawler/haj/Preliminarybufattenedlistoftransformations.pdf.
Schmerling, Susan. 1983. Two theories of syntactic categories. *Linguistics and Philosophy* 6. 393–421.
Schmerling, Susan. 2018. *Sound and grammar*. Leiden: Brill.
Scott, Gary-John. 2002. Stacked adjectival modification and the structure of nominal phrases. In Guglielmo Cinque (ed.), *Functional structure in DP and IP*, 91–120. Oxford: OUP.
Speas, Margaret & Naoki Fukui. 1986. Specifiers and projection. In Naoki Fukui, Tova Rapoport & Elizabeth Sagey (eds.), *MIT working papers in linguistics: Papers in theoretical linguistics*, vol. 8, 128–172. Cambridge, Massachusetts: MIT Press.
Stowell, Tim. 1981. *Origins of phrase structure*. PhD dissertation, MIT.

Svenonius, Peter. 2008. The position of adjectives and other phrasal modifiers in the decomposition of DP. In Louise McNally & Chris Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics, and discourse*, 16–42. Oxford: OUP.

Ticio, Emma. 2010. *Locality domains in the Spanish determiner phrase*. Dordrecht: Springer.

Travis, Lisa. 1984. *Parameters and effects of word order variation*. Cambridge, Massachusetts: MIT PhD thesis.

Uriagereka, Juan. 2012. *Spell-out and the minimalist program*. Oxford: OUP.

Wells, Rulon. 1947. Immediate constituents. *Language* 23. 81–117.

Westcoat, Michael. 2005. English nonsyllabic auxiliary contractions: An analysis in LFG with lexical sharing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG05 conference*. Stanford: CSLI. Available at: https://web.stanford.edu/group/cslipublications/cslipublications/LFG/10/pdfs/lfg05wescoat.pdf.

Zagona, Karen. 1982. *Government and proper government of verbal projections*. Seattle: University of Washington PhD thesis.

Zwicky, Arnold. 1985. Heads. *Journal of Linguistics* 21. 1–29.