Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges*

Noam Chomsky
University of Arizona & M.I.T.
chomsky@mit.edu

Ángel J. Gallego
Universitat Autònoma de Barcelona
angel.gallego@uab.cat

Dennis Ott
University of Ottawa
dott@uottawa.ca

Abstract

This paper provides an overview of what we take to be the key current issues in the field of Generative Grammar, the study of the human Faculty of Language. We discuss some of the insights this approach to language has produced, including substantial achievements in the understanding of basic properties of language and its interactions with interfacing systems. This progress in turn gives rise to new research questions, many of which could not even be coherently formulated until recently. We highlight some of the most pressing outstanding challenges, in the hope of inspiring future research.

Keywords: Generative Grammar; faculty of language; basic properties; operations; interfaces; syntax

Resum. *La gramàtica generativa i la facultat del llenguatge: descobriments, preguntes i desafiaments*

Aquest treball proporciona una visió general dels aspectes clau actuals en el camp de la gramàtica generativa: l'estudi de la facultat del llenguatge humà. Es tractaran algunes de les visions a què aquest enfocament del llenguatge ha donat lloc, incloent-hi èxits importants en la comprensió de

les propietats bàsiques del llenguatge i les seves interaccions amb els sistemes d'interfície. Aquest progrés dona lloc a noves preguntes de recerca, moltes de les quals fins i tot no es podien formular de manera coherent fins fa poc. Destaquem alguns dels reptes més destacats amb l'esperança d'inspirar futures investigacions.

Paraules clau: gramàtica generativa; facultat de llenguatge; propietats bàsiques; operacions; interfícies; sintaxi

Table of Contents

1. Introduction

Generative Grammar (GG) is the study of linguistic capacity as a component of human cognition. Its point of departure is Descartes' observation that "there are no men so dull-witted or stupid […] that they are incapable of arranging various words together and forming an utterance from them in order to make their thoughts understood; whereas there is no other animal, however perfect and well endowed it may be, that can do the same" (*Discours de la méthode*, 1662). Studies in comparative cognition over the last decades vindicate Descartes' insight: only humans appear to possess a mental grammar—an "I-language," or internal-individual language system—that permits the composition of infinitely many meaningful expressions from a finite stock of discrete units (Hauser et al. 2002; Anderson 2004; Chomsky 2012a, 2017).

The term *Universal Grammar* (UG) is a label for this striking difference in cognitive capacity between "us and them." As such, UG is the research topic of GG: what is it, and how did it evolve in our species? While we may never find a satisfying answer to the latter question, any theory of UG must meet a criterion of *evolvability*: the mechanisms and primitives ascribed to UG (as opposed to deriving from independent factors) must be sufficiently sparse to plausibly have emerged as a result of what appears to have been a unique, recent, and relatively sudden event on the evolutionary timescale (Bolhuis et al. 2014; Berwick & Chomsky 2016).

GG's objectives open up many avenues for interdisciplinary research into the nature of UG. Fifty years ago, Eric Lenneberg published his now-classic work that founded the study of the biology of language, sometimes called "biolinguistics" (Lenneberg 1967). In conjunction with the then-nascent generative-internalist perspective on language (Chomsky 1956[1975], 1957, 1965), this major contribution inspired a wealth of research, and much has been learned about language as a result. The techniques of psychological experimentation have become far more

sophisticated in recent years, and work in neurolinguistics is beginning to connect in interesting ways with the concerns of GG (Berwick et al. 2013; Nelson et al. 2017; Friederici et al. 2017).

Important results have emerged from the study of language acquisition, which is concerned with the interaction of UG and learning mechanisms in the development of an I-language (Yang 2002, 2016; Yang et al. 2017). Work by Rosalind Thornton and others shows that children spontaneously produce expressions conforming to UG-compliant options realized in languages other than the local "target" language, without any relevant evidence; but they do not systematically produce innovative sentences that violate UG principles. This continuity between children's seemingly imperfect knowledge and the range of variation in adult grammars suggests that children are following a developmental pathway carved out by UG, exploring the range of possible languages and ultimately converging on a steady state (for review and references, see Crain & Thornton 1998, 2012; Crain et al. 2016; for a theory of the steady state as a probability distribution over I-languages, see Yang 2016). Converging conclusions are strongly suggested by the spontaneous creation of sign languages by deaf children without linguistic input (Feldman et al. 1978; Kegl. et al. 1999; Sandler & Lillo-Martin 2006).

On the whole, we believe that GG has made significant progress in identifying some of the computational mechanisms distinguishing man from animal in the way recognized by Descartes. In this paper, we offer our view of the current state of the field, highlighting some of its central achievements and the many remaining challenges, in the hope of inspiring future research. Section 2 discusses the fundamental, "non-negotiable" properties of human language that any theory of UG has to account for. Section 3 focuses on core computational operations and their properties. Section 4 turns to the interfaces of I-language and systems entering into language use, and how conditions imposed by these systems constrain syntactic computation. Section 5 reviews a number of challenges emerging from recent work, which call for resolution in order to meet minimalist desiderata. Section 6 concludes.

2. Basic Properties of I-language

A traditional characterization of language, going back to Aristotle, defines it as "sound with meaning." Building on this definition, we can conceive of an I-language as a system that links meaning and sound/sign in a systematic fashion, equipping the speaker with knowledge of these correlations. What kind of system is an I-language? We consider two empirical properties non-negotiable, in the sense that any theory that shares GG's goal of providing an explanatory model of human linguistic capacity must provide formal means of capturing them: *discrete infinity* and *displacement*.[1] Atomic units—lexical items, whose basic nature remains

---

1.   The latter notion is non-negotiable in its abstract sense: there can be multiple determinants of interpretation for some syntactic object. The mechanisms implementing this basic fact vary across theoretical frameworks, of course.

a subject of debate[2]—are assembled into syntactic objects, and such objects can occupy more than one position within a larger structure. The first property is the technical statement of the traditional observation that "there is no longest sentence," the informal notion "sentence" now abandoned in favor of hierarchically structured objects. The second property is illustrated by a plethora of facts across the world's languages. To pick one random illustration, consider the familiar active/passive alternation:

(1) a. Sensei-ga   John-o   sikar-ta.   (Japanese)
         teacher-NOM   John-ACC   scold-PST
         'The teacher scolded John.'

     b. John-ga   sensei-ni   sikar-are-ta.
        John-NOM   teacher-by   scold-PASS-PST
        'John was scolded by the teacher.'

The noun phrase *John* bears the same thematic relation to the verb *sikar* in both (1a) and (1b), but appears sentence-initially in the latter. On the assumption that thematic relations are established in a strictly local fashion—a guiding idea of GG since its inception—, this entails that the nominal is displaced from its original position in (1b).

　　To account for these elementary properties, any theory of GG must assume the existence of a computational system that constructs hierarchically structured expressions with displacement. The optimal course to follow, we think, is to assume a basic compositional operation MERGE, which applies to two objects X and Y, yielding a new one, $K = \{X,Y\}$. If X, Y are distinct (taken directly from the lexicon or independently assembled), K is constructed by External MERGE (EM); if Y is a term of X, by Internal MERGE (IM). If K is formed by IM, Y will occur twice in K, otherwise once; but the object generated is $\{X,Y\}$ in either case. IM thus turns Y into a *discontinuous object* (or *chain*), which can be understood as a sequence of occurrences of Y in K.[3] (2) illustrates for (1b) above (abstracting away from irrelevant details), where MERGE combines K and the internal NP *John-ga:*

(2) a. {sensei-ni,{sikarareta,John-ga}} $= K \; \rightarrow \;$ MERGE(K,John-ga)

     b. {John-ga,{sensei-ni,{sikarareta,John-ga}}}$= K'$

2.   For a sample, see Hale & Keyser (1993, 1999); Borer (2005); Marantz (2001, 2013); Mateu (2005); Ramchand (2008); Starke (2014).
3.   We assume that each syntactic object is a (possibly singleton) set of occurrences, where occurrences are individuated by their context (structural sister). This is the definition assumed in Chomsky (2000a: 115), going back to Quine (1940: 297). See also Nunes (2004: 50ff.) and Collins & Stabler (2016: sect. 4) for critical discussion and alternative conceptions.

MERGE, applying recursively so that any generated object is accessible to further operations,[4] thus suffices to account for the basic properties of discrete infinity and displacement. Furthermore, it is the computationally simplest operation (as opposed to, say, concatenation, which adds order) that implements the basic properties of an I-language, and as such a conceptually necessary, irreducible component of UG. MERGE(X,Y), yielding K = {X,Y}, imposes hierarchical structure (X, Y are terms of K, but not vice versa) but no order ({X,Y} = {Y,X}). Languages differ in how they ultimately linearize objects constructed by MERGE, an important research topic for the study of the interaction between core syntax and the sensorimotor systems involved in perception and articulation. In (1a) above, the VP is linearized with OV order (*John-o sikarta)*, whereas a corresponding English VP would surface with VO order (*scolded John*). Interpretation is not affected by this difference, suggesting that the relevant parameter should be a matter of externalization of internally generated expressions alone (see Travis 1984 for original ideas along these lines).

A corollary of restricting composition to MERGE is the *structure-dependence* of syntactic operations: if order is only established in the morphophonological component, no syntactic operation can make reference to it. This excludes a large class of logically possible languages as not humanly acquirable, namely languages whose rules and operations are defined in linear terms (e.g., "reverse the order of words in the sentence to yield a question"). There is evidence that hypothetical languages of this sort are indeed outside of the range of variation permitted by UG. Neurolinguistic studies conducted by Andrea Moro and colleagues suggest that invented "languages" whose rules operate over linear order are treated by speakers as a puzzle rather than linguistic data, as indicated by diffuse activity in many parts of the brain as opposed to the pattern of activity observed in ordinary language use (Musso et al. 2003). Similar results were found in the study of a linguistically gifted but cognitively impaired subject (see section 4 below).

There are many illustrations of structure-dependence from syntax-semantics and morpho-phonology (Rizzi 2013a; Everaert et al. 2015). AUX-raising was used in the earliest days of GG as a straightforward illustration of the Poverty of the Stimulus: the fact that the input (linguistic data) vastly underdetermines the I-language eventually attained. The argument then and now is that the language-learning child never entertains the hypothesis that yes/no questions are formed by moving the *linearly first* auxiliary in the clause—a hypothesis that would receive ample support from cases such as (3) and requires complex examples of the kind in (4) to be refuted. (The symbol '_' marks the gap left behind by the displaced auxiliary.)

(3)   Is the tall man from Italy _ happy?

(4)   Is the tall man [who is from Italy] _ happy?

---

4.   Recursion is a "deep" property of the generative procedure; to what extent constructions exhibiting category recursion are used in some particular language (e.g., English but not German permits recursive possessors) is an orthogonal issue. For related discussion, see Arsenijević & Hinzen (2012); Chomsky (2014).

The computation chooses the *structurally* first (highest) auxiliary for inversion, not the one that happens to be embedded in the subject (at arbitrary depth), despite the fact that identification of the linearly first auxiliary is computationally straightforward. No other hypothesis is ever considered by the child, and consequently cases such as (5) are not attested in children's production (Crain & Nakayama 1987; Crain et al. 2017):

(5) *Is the tall man [who _ from Italy] is happy?

The formally innocuous linearity-based "first auxiliary" hypothesis would furthermore mislead children acquiring verb-final German into postulating questions such as (7), deriving from the verb-final structure underlying (6).

(6) dass der dicke Mann [der   aus   Italien gekommen war] glücklich war
     that the fat   man   who from Italy   come   was happy   was
     '…that the fat man who had come from Italy was happy.'

(7) *War der dicke Mann [der   aus   Italien gekommen _] glücklich war?
      was the fat   man   who from Italy   come   happy   was

Instead, structure-dependence dictates that the *structurally* closest auxiliary raise, exactly as in English and, crucially, irrespective of linear order:

(8) War der dicke Mann [der   aus   Italien gekommen war] glücklich _?
     was the fat   man   who from Italy   come   was happy
     'Was the fat man who had come from Italy happy?'

Children acquiring German do not simply adopt an alternative "last auxiliary" hypothesis, which would falsely produce the result in (9), where the relative clause has undergone optional rightward extraposition. Instead, learners instinctively know that the correct form is (10)—the only form possible if AUX-raising operates over hierarchical structure.

(9) *War der dicke Mann glücklich war [der   aus   Italien gekommen _]?
      was the fat   man   happy   was who from Italy   come

(10) War der dicke Mann glücklich _ [der   aus   Italien gekommen war]?
     was the fat   man   happy   who from Italy   come   was
     'Was the fat man happy who had come from Italy?'

As before (and always, it seems), structure trumps linear order. The conclusion is as obvious to the language-learning child as it is to the theorist if linearity-based rules are simply not part of the hypothesis space, i.e. not permitted by UG. Children acquiring German have the same understanding of structure-dependence as children acquiring any other grammatical system, since it follows from the hierarchical organization of linguistic objects constructed by MERGE.

The phenomenon of AUX-raising illustrated above, alongside other classical illustrations of structure-dependence, has been the focus of attention of so-called "usage-based" approaches, which assume that basic facts of language are not rooted in UG but rather the emergent result of statistical analysis over vast amounts of data. Approaches of this kind assume that language acquisition is essentially a matter of memorization and minimal generalizations over a large database. We will not evaluate the specific claims made by these proposals here, as this task has been undertaken elsewhere (Berwick et al. 2011; Crain et al. 2017). The approaches fail invariably both at adequately capturing the phenomena they focus on and, more fundamentally, at addressing the only theoretically relevant question: why do languages universally adopt structure-dependent operations while avoiding, in all relevant cases, far simpler computational operations based on linear order? An approach that restricts generation to MERGE provides a principled solution to this long-standing puzzle; in fact, it provides the *optimal solution*, a straightforward consequence of the simplest computational operation.

In line with a long tradition in linguistics, we take the I-language to derive sound/ sign-meaning pairs: objects constructed by MERGE are mapped onto a semantic representation SEM, accessed by conceptual-interpretive systems, and a phonetic representation PHON, accessed by sensorimotor systems, the latter providing instructions to the vocal or gestural articulators. Each derivation thus yields a pair <SEM,PHON>, whose properties enter into complex thought and intentional planning (e.g., discourse organization) and perception/articulation (internal in self-talk, external in oral or gestural production). We return to these interfaces below.

Displacement as illustrated in (1b) above often has effects on both SEM and PHON: displaced objects are interpreted as chains of occurrences, and derived positions are typically privileged in production. Consider a standard example of *wh*-movement (from Sportiche 2013):

(11) Je me demande de quel   livre sur   elle-même$_i$ [cette loi]$_i$ a   entraîné
     I   wonder of which book about she-self   this law has triggered
     la   publication (α).
     the publication
     'I wonder which book about itself this law triggered the publication of.'

<div align="right">(French)</div>

The *wh*-phrase *de quel livre* 'of which book' is displaced by IM from its original position (α) as the complement of the noun *publication* to the left edge of the embedded clause, where it surfaces in the externalized form. At SEM, the resulting chain of occurrences is interpreted as an operator-variable dependency: *(I wonder) which book x about y is such that this law y has triggered the publication of x*. SEM provides access to the original copy of the *wh*-phrase that externally merged in the position marked (α) above, as evidenced by the fact that this is where the reflexive pronoun *elle-même* is interpreted: in the scope of its antecedent *cette loi*. Once again, a state of affairs that would otherwise be highly puzzling can be given a principled rationale in terms of MERGE and its effects at the interfaces.

The structural distance spanned by dependencies of this sort is not clause-bounded but of arbitrary depth. Some well-known evidence suggests that movement leaves intermediate copies, so that "long" dependencies are in effect composed of "shorter" sub-dependencies (see Boeckx 2007 for a review). All copies are available at SEM, rendering reconstruction operations of earlier theories obsolete. By contrast, mapping to PHON forces a choice about the realization of the discontinuous object created by IM. The typical choice is the highest position, with all lower copies remaining silent. If, when, and how this preference can be overriden by parametric and other factors remains an important research question (cf. Nunes 2004; Trinh 2011).

Whether other types of rearrangements commonly found in the world's languages, such as semantically vacuous scrambling, extraposition, clitic movement etc., likewise reflect narrow-syntactic computations or are part of the mapping to PHON (prior to the introduction of linear order, hence with displacement-like properties) is an open question. It is commonly assumed that effects on meaning pertaining to topic/comment and focus/background articulation necessarily indicate core-syntactic displacements, but the relevant notion of "meaning" encompasses pragmatic as well as externalization-related (e.g., prosodic) properties of expressions. "Meaning" properties in this broad sense plausibly emerge from holistic interpretation of <SEM,PHON> pairs, rather than narrow-compositional interpretation of SEM itself. We briefly return to related matters in section 5.

Does the basic operation MERGE meet the criterion of evolvability? Any answer to this question is necessarily preliminary, given our ignorance about the evolution of UG. Bolhuis et al. (2014) and Berwick & Chomsky (2016) suggest that MERGE plausibly arose as a cognitive innovation in an individual, which ultimately spread to a group. Whether or not this speculation is on the right track, given that MERGE is the minimal computational operation required to generate a discrete infinity of syntactic objects, its emergence is a necessary prerequisite for our species-specific linguistic mind. The evolutionary origins of the other central component of I-language—the lexicon and its atoms with all their semantic intricacies (Chomsky 2000b)—remain deeply mysterious.

3. Operations and Constraints

We assume that MERGE(X,Y) forms {X,Y}, and nothing else. We will occasionally refer to this operation as *simplest MERGE*, in order to distinguish it from proposals in the literature adopting a more complex operation (cf. Epstein et al. 2014; Fukui & Narita 2014; Collins 2017).

A computational system comprising a lexicon and MERGE applying freely will automatically satisfy some fundamental desiderata, such as recursive generation of infinitely many structures with internal constituency and discontinuous (displaced) objects. MERGE operates over syntactic objects placed in a *workspace*: the MERGE-mates X and Y are either taken from the lexicon or were assembled previously within the same workspace (for some relevant formal definitions, see Collins & Stabler 2016). There is no motivation for additional representations, such

as numerations or lexical arrays, as employed in earlier approaches that assumed trans-derivational comparisons (Chomsky 1993, 1995; cf. Collins 1997: sect. 4.6 on this point).

We assume that MERGE is strictly binary: given that this is what is minimally necessary to create hierarchical structure, we assume that it is the *only* operation defined by UG (although adjunction structures may necessitate a separate operation, a point to which we return in section 5). Generation by simplest MERGE thus entails a restrictive class of recursively defined, binary-branching and discrete-hierarchical structures. Anachronistically speaking, early work on "non-configurational" languages by Ken Hale (1983) suggested that there are languages without the binarity restriction, but subsequent work showed this postulation of additional, non-binary combination operations to be unjustified; see, e.g., Webelhuth (1992) on German, Legate (2002) on Warlpiri, and Kayne (1984, 1994) for additional arguments. While challenges remain, we take binarity and the absence of "flat" structures to be a theoretically desirable and empirically feasible property of MERGE-based generation.

Restriction to simplest MERGE entails an *Inclusiveness Condition* (IC) that precludes the introduction of extraneous objects—for instance, traces and the bar-levels of X-bar Theory and other labels, but not copies and the detection of headedness via search (more on this below). Unlike the production rules of phrase-structure grammars, simplest MERGE thus incorporates no notion of "projection" (Chomsky 2013, 2015). IC also bars introduction of features that are not inherent to lexical items, such as the discourse-related features (topic, focus, etc.) assumed in the cartographic tradition and other approaches (e.g. Rizzi 1997; López 2009). We suggest below that MERGE is generally not triggered but applies freely. Importantly, IC need not be stipulated as part of UG: it is a corollary of simplest MERGE.

Suppose having constructed $K = \{X,Y\}$, we proceed to merge K and some object W. W is either internal to K or external to it. If W is external, then it is taken from the lexicon or has been assembled independently; this is EM. If W is internal to K, then it is a term of K; this is IM (displacement). If $W = Y$, MERGE(K,Y) yields $K' = \{Y,\{X,\mathbb{Y}\}\}$, with two copies (occurrences) of Y in K'. Note that there is still only one, discontinuous object Y in K', not two distinct objects; for instance, a semantically ambiguous phrase such as *Mary's book* will not be interpreted differently in the multiple positions it occupies after IM (as in, e.g., *Mary's book arrived/ was published ~~Mary's book~~ last month*).

A widely-held but, we believe, unjustified assumption is that MERGE is a "Last Resort" operation, licensed by featural requirements of the MERGE-mates (cf. Chomsky 2000a and most current literature, e.g. Pesetsky & Torrego's 2006 *Vehicle Requirement on Merge*). Note that a trigger condition cannot be restricted to either EM or IM: the operation MERGE(X,Y) is the same in both cases, the only difference being that one of X, Y is a term of the other in one case, while X and Y are distinct in the other. Simplest MERGE is not triggered; featurally-constrained structure-building requires a distinct, more complicated operation (defined as *Triggered Merge* in Collins & Stabler 2016; see Collins 2017 for additional

discussion). The features invoked in the technical literature to license applications of MERGE are typically *ad hoc* and without independent justification, "EPP-features" and equivalent devices being only the most obvious case.[5] The same holds for selectional and discourse-related features; the latter in addition violate IC, as noted above (cf. Fanselow 2006). Featural diacritics typically amount to no more than a statement that "displacement happens"; they are thus dispensable without empirical loss and with theoretical gain, in that Triggered Merge or equivalent complications become unnecessary (cf. Chomsky 2001: 32, 2008: 151; Richards 2016; Ott 2017b).[6]

MERGE thus applies freely, generating expressions that receive whatever interpretation they are assigned by interfacing systems.[7] Surface stimuli deriving from the objects constructed by I-language can have any degree of perceived "acceptability" or "deviance," from perfect naturalness to complete unintelligibility. Since Chomsky 1955[1975] it has been recognized that no independently given notion of "well-formedness" exists for natural language in the way it is stipulated for artificial symbolic systems (Chomsky & Lasnik 1993: 508). Consequently, concerns about "overgeneration" in core syntax are unfounded; the only empirical criterion is that the grammar associate each syntactic object generated to a <SEM,PHON> pair in a way that corresponds to the knowledge of the native speaker.[8] In fact, "overgeneration" must be permitted on purely empirical grounds, since "deviant" expressions are systematically used in all kinds of ways. To pick a random illustration, the expression *John will ever agree* involving NPI *ever* must be generated to be usable in contexts such as *I doubt that [John will ever agree]*. Constructions such as Right-node Raising may have similar properties (see Larson 2018).

Do we need operations other than MERGE for the construction of syntactic objects? Agreement phenomena indicate that there is an operation AGREE that relates *features* of syntactic objects (Chomsky 2000a, 2001). The assumption of much current work is that AGREE is asymmetric, relating initially unvalued φ-features on a *Probe* to matching, inherent φ-features of a *Goal* within the

---

5. The "edge features" of Chomsky (2008) are equally dispensable while not technically equivalent, and were originally introduced to distinguish elements that enter into computation from those that do not, such as interjections and response particles (which Holmberg 2016 argues to be elliptical in many cases).

6. A trigger-free approach to MERGE also eliminates the motivation for counter-cyclic MERGE in subject/object raising, an extremely complex operation (Epstein et al. 2012); see Chomsky (in press).

7. We should be careful to distinguish "interpretive systems" from "performance systems." The interpretive sensorimotor and conceptual-intentional systems are systems of cognitive competence, involved in the determination of entailment and rhyme relations among expressions, for instance. Actual performance introduces all sorts of other complicating factors, such as memory constraints, irrationality, etc.

8. By contrast, the conception of syntactic computation as "crash-proof" (Frampton & Gutmann 2002, among others) is based the dubious assumption that an I-language defines a set of well-formed, intuitively acceptable/natural expressions. But there is no basis for this assumption, and the informal notion of "acceptability" involves a host of factors that under no rational conception are part of I-language.

Probe's search space (structural sister). These dependencies find their expression in morphological inflection in highly variable, language-specific ways. AGREE is structure-dependent: in (12) and (13) below, the verbal morphology indicates agreement with the *in situ* object regardless of whether the linear order is VO or OV (examples from Tallerman 2005).

(12) ni-k-te:moa šo:čitl. (Nahuatl)   (13) Uqa jo   ceh-ade-ia.   (Amele)
    1SG-3SG-seek flower   he   houses build-3PL-3SG.PST
    'I seek a flower.'   'He built houses.'

   AGREE furthermore obeys structurally-conditioned minimality: regardless of the eventual surface order of constituents in (14) and (15), upon entering the derivation the inflectional Probe above the verb phrase locates the hierarchically closest Goal (underlined below) in each case—the singular subject in (14) vs. the plural one in (15), the latter subsequently displaced to the left.

(14) Die Kinder   hat / *haben [$_{vP}$ <u>die Lehrerin</u> die Kinder erschreckt].
    the children has   have   the teacher   startled
    'The teacher startled the children.'   (German)

(15) Die Kinder   haben / *hat [$_{vP}$ <u>die Kinder</u> die Lehrerin erschreckt].
    the children have   has   the teacher startled
    'The children startled the teacher.'

Embedding the plural subject NP of (15) within a larger singular NP expectedly gives rise to singular agreement, despite identical adjacency relations at the surface.

(16) [Die Geschichte über   [die Kinder]] hat / *haben [vP <u>NP_{sg}</u> die Lehrerin
    the story   about the children has   have   the teacher
    erschreckt].
    startled
    'The story about the children startled the teacher.'

   Empirically, AGREE or some equivalent operation is clearly required; we set aside here many intricacies of agreement phenomena uncovered in much detailed work on the topic (e.g. Bobaljik 2008; Harbour et al. 2008; Legate 2008). It is commonly assumed that IM is parasitic on AGREE, but this, like the assumption that applications of MERGE are licensed by formal features, requires a more complicated, separate movement operation. It is also empirically unfounded, since the effects of AGREE can be observed in the absence of IM and vice versa. Consider (18), where the matrix verb *parecen* 'seem' agrees with the *in situ* NP *varios sobornos a políticos* 'many bribes to politicians' (as well as with the participle *descubiertos* 'discovered').

(17) Parecen   haber   sido descubiertos   varios sobornos a políticos.
      seem.3PL have.INF been discovered.3PL many bribes   to politicians
      'Many bribes to politicians seem to have been discovered.'   (Spanish)

The NP can raise into the matrix clause but it need not, unlike in languages such as English. Cases of this short show that IM and AGREE are independent operations.[9] IM without AGREE is illustrated by cases such as (14) above.

Objects constructed in core syntax must be mapped onto representations that can be accessed by C-I and SM systems: SEM and PHON, respectively. Consequently, there must be an operation TRANSFER that hands constructed objects over to the mapping components. The mapping to PHON is complex, involving the computation of stress and a prosodic contour, "flattening" of the hierarchical structure, etc. (see Collins 2017 for a partial theory of this mapping, Idsardi & Raimy 2013 for general discussion, and Arregi & Nevins 2012 for a detailed case study in 'post-syntax'). The mapping to SEM is more direct, given that hierarchical structure is the input to semantic interpretation; just how complex it is depends on the obscure question of where the boundary between the generative procedure and C-I systems is to be drawn.

A further open question is what the effects of TRANSFER are on the syntactic derivation. Ideally, TRANSFER should impose some degree of *cyclicity* on the system, such that for a given syntactic object K assembled in the course of the derivation, further computation cannot modify K. This is achieved if TRANSFER renders the objects to which it applies impenetrable to later operations, thereby providing an upper bound to the internal complexity of syntactic objects operated on at any given stage of the derivation. In Chomsky 2000a and subsequent works it is suggested that the derivational *phases* subject to TRANSFER correspond to the thematic domain (the verb phrase, vP) and the propositional domain (the clause, CP). A common assumption in the literature is that TRANSFER to PHON (or *Spell-Out*) eliminates structure, such as the interior of a phase, from the derivation. This cannot be literally correct, however: transferred phases are not spelled out in their original position but can be realized elsewhere, such as when a larger object containing the phase is displaced (Obata 2010). To illustrate, in (18) the NP α contains the clausal phase β:

(18) [$_\alpha$ the verdict [$_\beta$ that Tom Jones is guilty]]

Suppose that subsequent to TRANSFER of β, α raises to a higher position, as in (19):

(19) [$_\alpha$ the verdict [$_\beta$ that Tom Jones is guilty]] seems to have been reached (α) by the jury

---

9. Further arguments are needed to establish the absence of covert raising in such cases (with English-style IM but pronunciation of the original copy); see Wurmbrand (2006) on German and Icelandic. But such vacuous covert displacements are highly dubious on grounds of learnability alone.

The clausal phase β is pronounced in its derived position internal to displaced α; it is not pronounced in its original position (or omitted from the final string). This means that there is no *Spell-Out*, and no structure is eliminated: there is only TRANSFER, which renders β inaccessible to subsequent manipulation.[10]

At the C-I interface, global principles of interpretation such as Condition C of the Binding Theory and the unboundedness of operator-variable dependencies (including "reconstruction" effects, as in (11) above) suggest the same conclusion: transferred phases remain accessible, but they cannot be modified at later cycles. This is a version of the *Phase Impenetrability Condition* (PIC) that permits Probe-Goal relations across phase boundaries, as long as these only manipulate the Probe. Examples include the well-known quirky-subject configurations in which C-T agrees (at least optionally) with an internal argument *in situ* and cases of long-distance agreement across finite-clause boundaries (D'Alessandro et al. 2008; Richards 2012).[11]

While permitting Probe-Goal relations and interpretive dependencies, PIC blocks IM of X "out of" a phase P on the plausible assumption that the resulting discontinuity of X alters P's internal structure.[12] Suppose X is raised from within P by IM. If syntactic objects are defined as sets of occurrences, it follows that P subsequently no longer contains X, since it does not contain the set of X's occurrences. Consequently, inter-phasal IM is barred by the PIC, as it affects the internal constitution of previously-transferred P. PIC thus requires raising of X to the edge of P before or at TRANSFER, as well as the assumption that the edge remains accessible at the next phase. In this way, the PIC gives rise to successive-cyclic movement and its reflexes in externalization.

If smaller units such as NPs, PPs, etc. are also phases (as argued in Uriagereka 1999, Abels 2003, Den Dikken 2007, Marantz 2007, Bošković 2014, and various other works), PIC enforces cyclic movement of any internal element that will undergo modification at a later stage of the derivation. While technically coherent, this inflation of phasal categories creates significant additional complexity and threatens to render the notion of phase-based derivation vacuous. The fact that the effects associated with successive-cyclic movement seem to be absent from these categories (Gallego 2012; Van Urk 2016) supports the hypothesis that vP and CP are the only phases.

The verbal and clausal phases in essence capture the "duality of interpretation" stated in terms of the D-structure/S-structure distinction of earlier theories. EM

---

10. We thus avoid the *assembly problem* of Collins & Stabler (2016), first discussed in Uriagereka (1999).

11. See Epstein et al. (2016a) for a theory of "phase cancellation" that may permit a stronger formulation of the PIC, with no access to what has already been transferred. For alternative ways to cancel, extend, or parametrize phases, see Gallego (2010a), den Dikken (2007), Alexiadou et al. (2014), and Chomsky (2015).

12. The *No-Tampering Condition* (NTC) sometimes assumed in the literature is a general desideratum of computational efficiency, but the case of IM shows that it cannot hold in its strictest form: if X is a term of Y contained in W, MERGE(X,W) *affects* both X (now a discontinuous object) and W (now no longer containing X), but doesn't *change* X or Y, e.g. by replacing either with a distinct object. This suggests that the NTC is reducible to the PIC (Gallego 2020).

within the vP phase gives rise to configurations expressing generalized argument structure, whereas IM at the CP cycle yields chains that enter into the determination of scope/discourse properties (Chomsky 2004, 2007; Gallego 2013a, in progress). While this is a reasonable approximation of the effects of EM and IM at the C-I interface, apparent exceptions (such as semantically vacuous displacements) pose interesting research questions.

To be sure, the basic operations MERGE, AGREE, and TRANSFER require much further formal explication; we will address some relevant issues in the following two sections.[13] Despite many remaining questions, we think that it is important to appreciate the fact that an austere system as outlined so far can accommodate a significant range of facts about natural language that are equally fundamental and surprising from a naïve point of view, such as hierarchical structure and structure-dependence, the cross-linguistically variable externalization of head-complement structures, the ubiquity of displacement and "reconstruction," and the duality of interpretation.

4. Interfaces

At the completion of each derivational cycle, the object W constructed in narrow syntax is subject to TRANSFER to the interfaces, mapping W onto SEM and PHON, accessed by C-I and SM systems, respectively. Let us refer to the mapping from narrow syntax to PHON as *externalization* (EXT). How and when does EXT take place? There are several possibilities. It could be that EXT takes place "all at once," applying to the final output of the narrow-syntactic derivation. Or it could be that the units rendered inaccessible by PIC are spelled out partially, while not being eliminated from the syntactic representation (permitting phasal objects to be moved as part of larger objects, as discussed above).

The interpretive and perceptual/articulatory systems accessing PHON and SEM impose constraints on the expressions freely constructed by MERGE that map onto these representations. For instance, the C-I system imposes a general requirement of *Full Interpretation*: all terms of a syntactic object must be interpreted, none can be ignored.[14] As a result, (20) cannot be interpreted at C-I as either "Who did John see?" or "John saw Mary," ignoring the theta-less object *Mary* or the vacuous operator *who*, respectively.

(20) {who,{John,{T,{see,Mary}}}}

----

13. We will not discuss here the operation of FEATURE INHERITANCE (F-I), introduced in Chomsky 2008 in order to account for the deletion of φ-features of phase heads. Ouali (2008) explores three possible manifestations of this operation, whereas Gallego (2014) argues that F-I can be eliminated under the Copy Theory of Movement. For reasons given in Richards (2007), F-I, like AGREE, must apply at the phase level, avoiding countercyclicity (Chomsky 2007: 19 fn. 26).
14. Sportiche (2015) argues that Full Interpretation permits "neglect" of elements that are meaningless or multiply represented. On this view, agreement features valued in the course of the derivation remain without consequence at C-I; no additional mechanism that removes these features is required.

So-called "crash-proof" models seek to bar generation of structures such as (21), given the intuitive "ill-formedness" of the derivative string (Frampton & Gutmann 2002). We think this is a mistake, for both conceptual and empirical reasons (see note 8 and related discussion above). On methodological grounds, constraints imposed on MERGE are typically redundant with more general interface conditions, such as Full Interpretation in the case of (20) (Chomsky 1986). The same is true for theta-theoretic violations, e.g. when the derivation fails to supply a strongly transitive verb with an object: the incompleteness is independently detected at the C-I interface, and there is no need to block generation of the "deviant" object, e.g. by complicating MERGE.[15] Furthermore, "deviant" expressions typically do have some interpretation, however inexpedient it may be in real-life usage.

More specific constraints are imposed by C-I on particular elements within SEM, such as those governed by the principles of Binding Theory. Thus, different types of pronouns receive interpretations that relate them to c-commanding antecedents in specific ways, accounting for the fact that *Himself likes John* does not mean "John likes himself," the impossibility of a coreferent interpretation of "John" and "him" in *John likes him*, etc. While many aspects of Binding Theory remain to be addressed for a system obeying IC, principled explanations of core cases in terms of C-I principles appear to be within reach (Chomsky 2008; Reuland 2011).[16]

What about the other interface, which relates the core computational system to articulatory and perceptual systems involved in EXT? As noted above, EXT is necessarily much more complex than the mapping to SEM, in that hierarchical objects must be translated into an altogether distinct, sequential format: while linear order plausibly plays no role in the syntactic and semantic processes yielding expressions and their interpretations, it is plainly required for vocal or gestural articulation. This is not the only complication: EXT violates just about every natural computational principle and carries out extensive modifications (e.g. by introducing boundary tones, prosodic contours and stress placement, etc., all in violation of IC), in ways that are furthermore highly variable across languages. What is more, the mapping must be sufficiently flexible to accommodate the contingencies of all possible modalities. For instance, speech requires strict temporal ordering, while gestural articulation permits a degree of simultaneity between manual and non-manual signs as well as within manual signs (Sandler & Lillo-Martin 2006; Vermeerbergen et al. 2007). The morphophonological properties superimposed as part of EXT also seem to be the locus of much, perhaps all cross-linguistic variation (in accordance with Chomsky's 2001 *Uniformity Principle)*.[17]

15. An important remaining question is how to handle apparent idiosyncrasies in selection. Some of these may well turn out upon closer scrutiny to be less idiosyncratic than standardly assumed, as argued recently by Melchin (2018) for *eat/devour*-type contrasts. Idiosyncratically-selected functional prepositions plausibly fall under a general theory of morphological case realized as part of externalization.

16. Chomsky (2007, 2008) suggests that reflexive binding might reduce to AGREE of one Probe with multiple Goals (cf. Hiraiwa 2005; López 2007). For more on this idea, see Hasegawa (2005); Gallego (2010b).

17. For related discussion and developments in the study of parametric variation, see Biberauer et al. (2014); Eguren et al. (2016); Kayne (2013); Picallo (2014).

Psycholinguistic and neurolinguistic inquiries have the potential to shed light on the status of EXT. One example is Smith & Tsimpli's (1995) work on a subject they call Chris, whose cognitive capacities are extremely limited but who has extraordinary linguistic capacities that allow him to pick up languages very quickly (at least superficially, without significant understanding). Smith and Tsimpli investigated Chris's reactions to invented languages of two types, one that conformed to UG principles and another that used principles that are not available to UG, such as linearity-based operations. It turned out that Chris was completely unable to deal with the language based on simple computational procedures using linear order, but would master easily an invented language that conformed to UG principles in employing structure-dependent rules. Subsequent studies by Smith and Tsimpli (corroborated by Musso et al.'s 2003 findings mentioned above) suggest that normals can likewise deal relatively easily with languages conforming to UG principles, but can handle the non-UG-conforming systems relying on linear order only if they were expressly presented as a puzzle rather than a language. While preliminary, these findings strike us as suggestive.

These observations support the speculation that those properties of language that pertain exclusively to perception and articulation are ancillary, perhaps altogether external to I-language, whereas the core computational system may be close to uniform (Berwick & Chomsky 2016; but see Irurtzun this volume).[18] EXT relates very different systems, a computational system constructing hierarchical expressions on the one hand and sequential production/perception systems on the other. While the computational system appears to have evolved recently and suddenly, the SM systems had at that point been in place for hundreds of thousands of years (see, e.g., Fitch 2010: chapter 8).[19] Given that the linkage between these two systems is an inherently "messy" affair, EXT is a plausible source of linguistic variation—perhaps the only one.

Where does all of this leave us with regard to the question of evolvability? MERGE and the inventory of lexical atoms it operates over must be part of UG and as such represent evolutionary innovations specific to the human linguistic mind. What about AGREE and TRANSFER? We believe that while no firm conclusions can be drawn at this point, it is plausible that these operations are rooted in principles of efficient computation. Chomsky (2013, 2015) suggests that AGREE instantiates *minimal search* within the syntactic object, in which case its core properties (structure-dependence, minimality) would reduce to general properties of computation. With regard to TRANSFER and the interface mappings, we noted above that the mapping to PHON is necessarily complex, while the mapping to SEM may be near-trivial. A plausible speculation is that EXT and its variable properties reflect not UG specifications but rather the absence thereof, if the linkage

18. We say "close" because even a computationally minimal core syntax might permit a degree of variation when multiple derivational options are consistent with efficiency of computation. See Richards (2008) and Obata et al. (2015) for proposals along these lines.

19. See also Huybregts (2017) for relevant recent discussion (expanding on observations in Uriagereka 2012: 254) of the evolutionary relevance of aerially isolated click phonemes.

established between the computational system proper and externalization systems was a problem that had to be solved subsequently to the evolution of I-language.

5. Open Questions and Future Directions

In this section, we turn to a number of theoretical issues and outstanding questions that have emerged in recent work. While we will outline what seem to us to be plausible steps towards resolving these questions, our primary intention here is to highlight their relevance to future research in GG.

We begin by returning to the operation MERGE, which, despite its apparent simplicity, raises many questions. A narrow conception of MERGE permits only two logical options: binary EM and IM. Various further options have been proposed in the literature, such as Parallel Merge/Sideward Movement, a species of "multidominance" structures (Nunes 2004; Citko 2005), and countercyclic Late Merge (Lebeaux 1988; Fox 2002), which replaces a displaced object with a larger one. Are these options corollaries of the availability of simplest MERGE, as has sometimes been claimed, or do they require additional mechanisms, raising new evolvability problems? We believe that there are reasons for skepticism towards these extensions beyond a narrow conception of MERGE, which warrant further scrutiny in future research.

All syntactic objects in the lexicon and in the workspace WS are *accessible* to MERGE; there is no need for a SELECT operation (as in, e.g., Chomsky 1995). WS represents the stage of the derivation at any given point. The basic property of recursive generation requires that any object already generated be accessible to further operations. WS can contain multiple objects at a given stage, so as to permit formation of {XP,YP} structures (subject-predicate constructions) by EM. A derivation may, but need not, terminate whenever WS contains a single object; if it terminates in any other situation, no coherent interpretation can be assigned.

Beyond these fundamentals, many questions arise. For instance, does MERGE(X,Y) *add* {X,Y} to WS = [X,Y] (where X, Y are LIs or complex elements), yielding WS′ = [X,Y,{X,Y}]? Or does it rather *replace* X and Y in WS with {X,Y}, yielding WS′ = [{X,Y}] (as assumed in Chomsky 1995: 243)? The latter view is more restrictive, and arguably more in line with basic desiderata for optimal generation: the generative procedure constructs a *single* object to be mapped onto PHON and SEM, not a multiplicity of objects; and considerations of computational efficiency suggest that WS should be kept minimal throughout a derivation.[20] The same conclusion is suggested by the fact that a workspace WS′ = [X,Y,{X,Y}] derived by MERGE(X,Y) would not ensure that subsequent operations can apply in a determinate fashion: any rule applying to X or Y would ambiguously refer to the individual objects X, Y or to the terms of K = {X,Y}.

---

20. A strong hypothesis about the generative procedure would be that operations never extend WS (i.e. increase the cardinality of elements contained in it). Except for the case where two elements taken from the lexicon are combined, EM and IM keep WS constant or reduce it. For related considerations (but very different conclusions), see De Vries (2009).

Indeterminacy of rules in this sense is formally unproblematic and in fact a familiar property of phrase-structure grammars; but a sensible question to ask is whether it should be permitted in an optimal I-language at all, given that it raises various technical complications (for instance with regard to the distinction between copies and repetitions, to which we return below). If the answer is negative, we are led to a view of simplest MERGE as mapping WS = [X,Y] onto WS′ = [{X,Y}], reducing its complexity and avoiding indeterminate rule application. For further elaboration on this conception of MERGE as a function mapping workspaces onto workspaces, going back to Bobaljik (1995), see Collins & Stabler (2016) and Chomsky (this volume); for an alternative conception of derivations that does away with work-spaces, see Collins (2017).

This restrictive view of MERGE, which seeks to curtail the complexity of WS, bars operations such as Parallel Merge (which establishes a ternary relation between the shared element X, its MERGE-mate Y, and the object Z containing Y) and Late Merge (which requires substitution of X by some more complex object; see Epstein et al. 2012).[21] This leaves EM and IM as the only possible instantiations of simplest MERGE. We believe that future work should address these and other questions raised by the above considerations, in order to establish a restrictive "null theory" of the generative procedure that adheres to plausible—yet at present necessarily tentative—desiderata of computational efficiency.

Regardless of which implementation of recursive generation we adopt, a further central question is how a MERGE-based system can distinguish copies (created by IM) from repetitions of identical elements (created by EM), so that we correctly distinguish the two instances of the noun phrase *the man* in *The man saw the man* from those in the unaccusative construction *The man arrived the man*. Suppose MERGE(K,W), where W is a term of K, creates Z. Z now contains two (or more) *copies* of W. But upon accessing Z, how do the external interpretive systems know whether multiple instances of W are copies of a single object or independent objects (*repetitions* of W)?[22]

Different answers to this question have been pursued, e.g. in terms of multi-dominance structures (Gärtner 2002) or an operation COPY that duplicates W prior to IM (Chomsky 1993; Nunes 2004). But complex graph-theoretic objects are not defined by simplest MERGE, and no COPY operation is necessary given that copies are simply a by-product of IM (on standard set-theoretic assumptions). Another possibility is that the system keeps track of how often the relevant object was assembled (or accessed in the lexicon) and communicates this information to the interfaces as part of TRANSFER (see Kobele 2006 and Hunter 2011 for related proposals). Along these lines, Chomsky (2007, 2012b) proposes that the distinction is established by the phasal nature of syntactic computation. At TRANSFER, phase-level memory suffices to determine whether a given pair of identical terms

---

21. See Sportiche (2015) for an alternative treatment of the facts motivating Late Merge analyses in terms of "neglect" at the interface.
22. Earlier theories sidestepped the problem by assuming a rewriting of lower copies as distinct symbols (traces), linking these to their antecedent via coindexing, in radical violation of IC.

Y, Y′ was formed by IM.[23] If it was, then Y and Y′ are copies; if it was not (i.e., it was formed by EM), Y and Y′ are independent repetitions. This captures the basic intuition that if some syntactic object is introduced into the derivation "from the outside," it is a distinct object; if it is added "from within," it is a copy. Phases would then play the crucial role of limiting memory to the current cyclic domain (the principal desideratum of phase theory), preventing unbounded search and thus rendering the detection of repetitions vs. copies computationally feasible. For critical discussion of this approach, which remains to be formalized, as well as related proposals, see Collins & Groat (2018).

A further important question is whether objects constructed by MERGE are necessarily endocentric and identified by a determinate *label*, as in earlier phrase-structural models incorporating X-bar Theory. The assumption of universal endo-centricity carried over to the Bare Phrase Structure model of Chomsky (1995), where MERGE(X,Y) is taken to yield a labeled object $\{L,\{X,Y\}\}$, $L \in \{X,Y\}$. But this is a departure from simplest MERGE, rooted in the intuitive appeal and pedagogical convenience of tree notation. In its simplest form, MERGE has no "built-in" projection mechanism, hence does not yield labeled objects (Chomsky 2013, 2015; Collins 2017). Unlike displacement and linear order, projection is not an empirically detectable property of linguistic expressions but a theory-internal concept. Encoding a label as part of the object constructed by MERGE raises various non-trivial questions (Seely 2006)—for instance, why can the label not undergo head movement on its own, or be pronounced? These problems vanish if labels *qua* syntactic objects do not exist, but the question of endocentricity remains in a different form: is it relevant to the syntactic derivation and/or to the interfacing systems?

Chomsky (2013) argues that the answer to this question is positive, and that an algorithm LABEL is required to supplement MERGE. For some syntactic object K, LABEL(K) locates within K the first element where search "bottoms out:" the structurally most prominent lexical item. LABEL is thus not an entirely new operation, but, like AGREE, an instantiation of minimal search. For K = {H,XP}, where H is an LI and XP a complex object, H will be chosen as the label. The first step in a derivation necessarily relates two atomic objects, yielding K = {H,R}. What is the label of K in this case? If R is a feature-less root, as assumed by many contemporary approaches, it is plausibly ignored by LABEL, and H will be correctly chosen as the label of K. On this conception, LABEL locates a *feature* of H, which renders the traditional notion of "head" irrelevant for labeling purposes. This approach to labeling raises intricate questions about the nature of lexical items (and the distribution of their properties across components, as assumed by models such as Distributed Morphology), which we set aside here.

X-bar-theoretic universal endocentricity has conceptually and empirically questionable consequences. To begin with, it is trivially falsified by every case of IM,

---

23. Identity must take features into account, so that, for instance, in a double-object construction with two identical objects (*The king sold a slave a slave*), an object NP raised to the phase edge can be correctly associated with its lower copy. The distinction is trivial if the NPs are distinguished by structural vs. inherent case-marking.

which yields an unlabelable {XP,YP} configuration (putting aside head movement). Another case in point is the DP hypothesis, a corollary of X-bar Theory. Bruening (2009) shows that while selection by a higher verb clearly targets C (the head of the clause), there is no selection for D (only for properties of N, such as number); and unlike C, D is not universal. The challenge, then, is to accommodate D-type elements while retaining the nominal character of the overall phrase. One possibility, suggested in Chomsky (2007) and developed by Oishi (2015), is that nominals are headed by a nominalizer n, analogous to v as the head of the verb phrase, with D, where present, occupying some lower position. Another is that determiners are in fact internally complex elements, as suggested by their morphology in many languages; see, e.g., Leu (2015).

If K = {X,Y} and neither X nor Y is a lexical item (e.g., when X is a "specifier" in earlier terminology), no head is detected by LABEL. Building on Moro (2000), Chomsky (2013) argues that this situation can motivate displacement of X: if X merges (internally) to some object W containing K, K will no longer contain X (X being the set of its occurrences), and consequently Y will act as the label of K. Chomsky suggests that W and X must share a feature if the resulting configuration is to be "stable," an idea that Chomsky (2015) extends to EPP and ECP effects (see also Rizzi 2015). Such feature sharing is involved in subject/object raising, for instance, where the raising XP enters into an AGREE relation with the head it raises to (T/v*, respectively; see Gallego 2017 for an alternative, and Epstein et al. 2016b for further discussion).

Again building on Moro's work, Ott (2012), Chomsky (2013, 2015), and Blümel (2017) argue that the need to break the symmetry of {XP,YP} configurations (motivated by LABEL) can drive displacement of XP, yielding phenomena such as successive-cyclic movement, raising to object, and others. Such proposals assume that MERGE applies freely; but derivations in which relevant applications fail to apply will not yield the required outcome. Plausibly, efficiency of computation precludes "superfluous" applications of MERGE that have no effect on the eventual output (such as string-vacuous IM with no effect on interpretation, which would entail massive structural ambiguity of any given sentence). For proposals along these lines and relevant evidence, see e.g. Fox (2000), Chomsky (2001, 2008a), Reinhart (2006), Struckmeier 2016.

Note that unlike classical X-bar Theory, a LABEL-based system allows for the possibility that a constructed object K remains unlabeled (exocentric), e.g. when K is a root clause or created by operations that are not head-oriented in any plausible sense, such as syntactic scrambling. Further illumination of these issues will require a theory that answers the question of where detectable endocentricity is required: in the syntactic derivation (e.g., for purposes of interpreting local selectional relations), at the interfaces (e.g., for the computation of prosody), both, or not at all (Collins 2017)? These questions remain open for now and are in urgent need of clarification.

A further important research question is whether structure-building mechanisms beyond simplest MERGE are necessary, such as Chomsky's (2004) PAIR-MERGE for adjuncts and De Vries's (2012) PAR-MERGE for parenthetical expressions.

Adjuncts and parentheticals have distinct properties, among them strong opacity for extraction. Thus, while (21) is ambiguous between a complementation and an adjunction structure, (22) is unambiguous, since only the former permits IM of the *wh*-phrase. And while an NP such as *a book about NP* readily permits *wh*-extraction of NP (23), an analogous extraction from a corresponding parenthetical appositive NP yields no coherent interpretation (24).

(21) John decided on the boat.

(22) What did John decide on _?

(23) What did John read a book about _?

(24) *What did John read something, a book about _?

Chomsky (2004) proposes that adjunction is the result of an operation PAIR-MERGE, which yields asymmetric (ordered) pairs rather than symmetric (unordered) sets, permitting the identification of an adjunct in a phrase-modifier configuration. PAIR-MERGE may also be required for unstructured coordination (as in *John is tall, happy, hungry, bored with TV, etc.*), a construction that was recognized as problematic in the earliest work in GG, due to the apparent absence of internal hierarchical organization: even unrestricted rewriting systems cannot generate these expressions, nor can transformations (see Lasnik & Uriagereka 2012 for a critical review of some proposals in Chomsky & Miller 1963).[24]

PAIR-MERGE is a formally distinct operation from simplest MERGE, hence raises problems of evolvability. Ideally, it could be shown to be dispensable. We do not take up the challenge here; for some suggestive work on adjunction that does not invoke special operations (but at the cost of introducing other stipulations), see Hunter (2015). As for parenthesis, it seems to us that the only principled approach consistent with evolvability considerations relegates the phenomenon entirely to discourse pragmatics, obviating the need to enrich UG with special operations. On this view, parenthetical expressions (which are frequently elliptical) are generated independently and interpolated or juxtaposed only in production (see Ott & Onea 2015; Ott 2016a,b).

Traditionally, adjunction is also assumed to be involved in head movement (HM),[25] but such an approach has several unwelcome consequences (Chomsky 2015: 12ff.; also Carstens et al. 2016). HM violates principles of minimal computation and cannot be implemented by simplest MERGE, given its countercyclic character. It also typically lacks semantic effects, at least for the core cases of

---

24. A possible analysis of unstructured coordination that avoids these problems could take each AP in the above example to be an elliptical 'afterthought' expression in the sense of Ott & De Vries (2016), Ott (2016). This would capture the central properties of the construction: infinite iterability and individual predication of each AP of the subject. For reasons of space, we cannot explore this idea further here.

25. See Epstein et al. (2016a) on PAIR-MERGE as a mechanism for affixation.

verb raising. This vacuity and the fact that the configurations standardly described in terms of HM are highly variable across languages suggest that at least some instances of HM might fall within the mapping to PHON (as suggested in Chomsky 2001 and supported by specific arguments in Zwart 2017 and elsewhere), although there are interesting arguments to the contrary (e.g., Roberts 2010).[26,27] Other cases might reduce to core-syntactic IM, in line with proposals in Toyoshima (2000) and Matushansky (2006). We believe that a fresh take on the relevant phenomena is needed, based on the recognition that traditional implementations of HM are in fact problems restated in technical terms rather than solutions.

We noted above that simplest MERGE applies freely, and that features which are not introduced into the derivation by LIs, such as those pertaining to informational functions of XPs, violate IC. "Cartographic" analyses, where such features take center stage as the driving force behind displacements to the peripheries, are essentially construction-based approaches, with the notion "construction" recast in terms of features and phrase-structure rules generating cascades of projections. But informational notions such as "topic" or "focus," like grammatical functions or thematic roles, are properties of configurations and their syntactic/discursive context, not of individual syntactic objects (Chomsky 1965; Hale & Keyser 1993); consequently, they should neither be represented in the lexicon, nor in the narrow syntactic derivation (cf. Uriagereka 2003; Fortuny 2008; López 2009; Gallego 2013a, in progress).

The Cartographic Program pursued by Cinque, Rizzi and many others has revealed remarkable facts and generalizations, such as Cinque's (1999) hierarchy of adverbial positions and Rizzi's (1997) structure for the left periphery. But the postulated structures raise serious problems, as acknowledged by Cinque & Rizzi (2010: 63). As we observed above, any linguistic theory must minimally meet the conditions of acquirability and evolvability. UG must permit acquisition of I-language, and it must have evolved in the human lineage—and if current best

---

26. For a different, syntactic approach to HM, see Chomsky (2015). Core-syntactic HM is presupposed by many approaches to diverse phenomena, such as Donati's (2006) analysis of free relatives, where the *wh*-element is analyzed as a D head that determines the label of the embedded clause after IM. See Ott (2011) for an alternative that is consistent with a non-syntactic conception of HM, but relies on specific assumptions concerning the interaction of TRANSFER and LABEL.

27. An interesting challenge to the idea that HM could be relegated to EXT is provided by Spanish VOS constructions, which suggest that verb movement can resolve minimality conflicts (see discussion around (14)-(16) above). Consider (i) below, where the internal argument *cada coche* 'each car' has moved to a position at the vP edge, from where it c-commands the vP-internal external argument *su propietario* 'its owner,' enabling a bound-variable interpretation of the subject-internal pronoun.

(i) Recogió   [$_{vP}$ cada coche [$_{vP}$ su propietario (v) [recogió cada coche]]]   (Spanish)
    picked-up   each car   its owner
    'Each car was picked up by its owner.' (lit.: 'Its owner picked up each car.')

What is surprising is that this configuration does not preclude AGREE between C-T and the external argument (as it should under a conception of minimality without the notion of equidistance: Chomsky 1993, 2000). The facts are discussed in Gallego (2010, 2013b), where it is argued that nominative Case assignment to the *in situ* subject in such cases is parasitic on verb movement. If HM were merely a phonological operation, its apparent role in licensing Probe-Goal dependencies would be unexpected.

guesses are correct, it must have evolved recently. The cascades of projections postulated for various areas of clause structure cannot possibly be learned: there is no conceivable evidence that a child could rely on to infer these hierarchical sequences from experience. But attributing complex functional hierarchies to UG raises an evolutionary puzzle: it seems virtually unimaginable that the complex cartographic templates could have evolved as irreducible properties of UG. The conclusion is that cartographic sequences of positions are problems, not solutions. As aptly discussed by Rizzi (2013b), the challenge is to derive the descriptive generalizations from more elementary principles that are motivated independently.

There is some promising work in this direction, such as Ernst's (2002) non-templatic analysis of adjunct ordering that derives Cinque's universal template from interpretive properties of adverbial expressions, rendering a "hard-coded" functional sequence obsolete. Developing alternatives to templatic approaches to the clausal peripheries will require, we believe, a re-evaluation of the extent to which the superficial complexity of "sentences" in fact reflects amalgamation of independent expressions in discourse, rather than syntactic composition. In contrast to Cinque's (1983) early work on "topic constructions," the cartographic tradition assumes that all sorts of peripheral elements, including left- and right-dislocated constituents, are structurally integrated into the clause structure. As a result, the puzzling properties of dislocated elements that distinguish them from displaced constituents (such as *wh*-phrases) are merely restated, not explained, including their universal extra-peripheral ordering. An alternative, developed in Ott (2014, 2016b, 2017a), denies the reality of structurally complex peripheries by analyzing dislocated elements, unlike fronted or extraposed XPs, as structurally independent elliptical expressions that are interpretively related to their host clauses by principles of discourse organization and cross-sentential anaphora. On this alternative approach, cartography's peripheral functional sequence remains only as an artifact of description.

We adumbrated above the idea that the core computation yields hierarchically-structured, language-invariant expressions (entering into "thought" processes of various kinds at the interface with C-I systems) whereas the mapping that feeds externalization-related SM systems is necessarily more involved and indirect. This asymmetry between the two interfaces leads Chomsky (2014) to adopt the following hypothesis:

(25) I-language is optimized relative to the C-I interface alone, with EXT ancillary.

"Optimized" here refers to the kinds of considerations introduced above: relying only on simplest MERGE and no more complex operations. As we pointed out, this strong thesis is consistent with the general fact that operations of I-language operate over structures, not strings (with concomitant beneficial implications for language acquisition), and that structured objects provide the input to compositional interpretation. At the same time, challenges for (25) emerge from recent work suggesting a rather direct involvement of morphophonological factors in the syntactic computation. Richards (2016) develops an elaborate theoretical framework in which the articulation systems impose universal constraints that, in conjunction

with independent language-specific differences, can account for central aspects of cross-linguistic variation (see also Mathieu 2016 for a related proposal). In this model, metrical requirements of affixes and other conditions imposed by PHON can effect the application of MERGE and other operations.[28] Given the impressive results achieved by Richards' system, his work poses an interesting challenge to the hypothesis that EXT is an ancillary process. The same is true for recent work arguing for the relevance of linear order to various syntactic and semantic processes (Kayne 2011, 2018; Barker 2012; Bruening 2014; Willer Gold 2018), contrary to our suggestions above. If and how these challenges can be reconciled with (25) is an important topic for future research.[29]

As noted above, a related open question pertaining to the overall organization of the system is whether the narrow-syntactic computation includes an operation AGREE in addition to MERGE, or whether featural interactions are restricted to EXT. The former view is based on the assumption that AGREE mediates assignment of structural Case and serves to eliminate semantically redundant φ-features from the syntactic object, as required by a particularly strong version of the Full Interpretation principle (Chomsky 2000a *et seq*., building on observations of Vergnaud 1977[2006] and George & Kornfilt 1981). Another possibility is that case is a purely morphological phenomenon (Marantz 1991; McFadden 2004), and that uninterpretable features are simply neglected at the C-I interface (in the spirit of Sportiche 2015). The latter scenario is consistent with relegating AGREE to EXT, where it would then serve the sole purpose of determining the morphological form of initially underspecified inflectional elements (cf. Bobaljik 2008, and Preminger 2014 for an opposing view; also Landau 2016 for an argument from Control). Also in view of the cross-linguistically highly variable expression of inflection, AGREE seems to fit rather naturally with other operations pertaining to EXT. We believe that there are interesting arguments in either direction and leave the matter here as an important topic for future research.

These and many other issues concerning the overall architecture of the computational system(s) underlying human linguistic capacity remain to be adequately addressed and explored. The mere fact that they can be coherently stated testifies to the progress GG has made over the years, providing ample fertile ground for further stimulating research.

---

28. Richards explicitly discusses instances of derivational opacity, where phonological factors trigger movements whose effects are later undone by subsequent operations. This entails that the morphophonology in his model cannot simply act as an output filter, but must be directly involved in the narrow-syntactic derivation.

29. Kayne (2018) presents a series of arguments for the inclusion of linear order in core-syntactic operations, proposing an operation *ip-merge* that yields an ordered pair expressing a relation of immediate precedence. Kayne argues furthermore that the operation is constrained such as to only construct LCA-compliant syntactic objects (in the sense of Kayne 1994). This logic strikes us as inconsistent: where LCA or some similar principle determines order, it is wholly redundant to impose order independently in narrow syntax. Kayne's empirical arguments also strike us as unconvincing, as they appear to pertain primarily to pragmatics/discourse organization and production/processing, hence EXT. For reasons of space, however, we have to leave a proper discussion of Kayne's arguments to another occasion.

6. Conclusion

Even within the expressly narrow focus of GG on linguistic competence, virtually every aspect of (I-)language remains a problem. Nevertheless, significant progress has been made since the 1950s, and in recent years the establishment of a minimal formal toolkit meeting basic desiderata of explanatory and evolutionary adequacy has become a feasible goal. As always, it remains to be seen to what extent such a toolkit can be reconciled with the empirical challenges and puzzles that inevitably arise wherever we look. As documented above, an approach based on the operation MERGE raises new problems on its own, both empirical and conceptual. In fact, in many cases it remains to be determined where to even look for solutions, e.g. when we ask whether heavy-NP shift falls within the MERGE-based system of core computation or is part of externalization. In our view, this conclusion makes the challenges ahead no less exciting, but should rather fuel our appreciation of the fascinating research questions that present themselves once we approach human language as an object of the natural world.

References

Abels, K. 2003. *Successive Cyclicity, Anti-locality, and Adposition Stranding*. PhD dissertation, University of Connecticut.

Alexiadou, A., E. Anagnostopoulou & S. Wurmbrand. 2014. Movement vs. Long-distance Agree in Raising: Disappearing Phases and Feature Valuation. In H.-L. Huang, E. Poole & A. Rysling (eds.). *Proceedings of NELS 43*, 1-12. Amherst, MA: GLSA.

Anderson, S.R. 2004. *Doctor Dolittle's Delusion*. New Haven, CT: Yale University Press.

Arregi, K. & A. Nevins. 2012. *Morphotactics*. Dordrecht: Springer.

Arsenijević, B. & W. Hinzen. 2012. On the Absence of X-within-X Recursion in Human Grammar. *Linguistic Inquiry* 43: 423-440.

Barker, C. 2012. Quantificational Binding does not Require C-command. *Linguistic Inquiry* 43: 614-633.

Berwick, R.C. & N. Chomsky. 2011. The Biolinguistic Program: The Current State of its Development. In A.M. Di Sciullo & C. Boeckx (eds.). *Biolinguistic Investigations*, 19-41. Oxford: Oxford University Press.

Berwick, R.C. & N. Chomsky. 2016. *Why Only Us*. Cambridge, MA: MIT Press.

Berwick, R.C., P. Pietroski, B. Yankama & N. Chomsky. 2011. Poverty of the Stimulus Revisited. *Cognitive Science* 35: 1207-1242.

Berwick, R., A. Friederici, N. Chomsky & J. Bolhuis. 2013. Evolution, Brain, and the Nature of Language. *Trends in Cognitive Sciences* 17: 89-98.

Biberauer, T., A. Holmberg, I. Roberts & M. Sheehan. 2014. Complexity in Comparative Syntax: The View from Modern Parametric Theory. In F. Newmeyer & L. Preston (eds.). *Measuring Grammatical Complexity*, 103-127. Oxford: Oxford University Press.

Bobaljik, J. D. 1995. In Terms of Merge: Copy and Head Movement. In R. Pensalfini & H. Ura (eds.). *Papers on minimalist syntax* (= *MIT Working Papers in Linguistics* 27). Cambridge, MA: MITWPL.

Bobaljik, J. D. 2008. Where's Phi? Agreement as a Postsyntactic Operation. In D. Adger, D. Harbour & S. Béjar (eds.). *Phi Theory: Phi-Features Across Interfaces and Modules*, 295-328. Oxford: Oxford University Press.

Bolhuis, J., I. Tattersall, N. Chomsky & R.C. Berwick. 2014. How Could Language Have Evolved? *PLoS Biology* 12: e1001934.

Blümel, A. 2017. *Symmetry, Shared Labels and Movement in Syntax*. Berlin: De Gruyter.

Boeckx, C. 2007. *Understanding Minimalist Syntax: Lessons from Locality in Long-Distance Dependencies*. Oxford: Blackwell.

Borer, H. 2005. *Structuring Sense* (2 volumes). Oxford: Oxford University Press.

Bošković, Z. 2014. Now I'm a Phase, Now I'm Not a Phase: On the Variability of Phases with Extraction and Ellipsis. *Linguistic Inquiry* 45: 27-89.

Bruening, B. 2009. Selectional Asymmetries between CP and DP Suggest that the DP Hypothesis is Wrong. *U. Penn Working Papers in Linguistics* 15.1: 26-35.

Bruening, B. 2014. Precede-and-Command Revisited. *Language* 90: 342-388.

Carstens, V., N. Hornstein & T.D. Seely. 2016. Head-Head Relations in 'Problems of Projection'. *The Linguistic Review* 33: 67-86.

Chomsky, N. 1956[1975]. The Logical Structure of Linguistic Theory. Mimeograph, Harvard University and MIT. Partial revised version published by Plenum Press, 1975.

Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1986. *Knowledge of Language*. New York: Praeger.

Chomsky, N. 1993. A Minimalist Program for Linguistic Theory. In K. Hale & S.J. Keyser (eds.). *The View from Building 20*, 1-52. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2000a. Minimalist Inquiries: The Framework. In R. Martin, D. Michaels, J. Uriagereka & S.J. Keyser (eds.). *Step by Step*, 89-155. Cambridge, MA: MIT Press.

Chomsky, N. 2000b. *New Horizons in the Study of Language and Mind.* Cambridge: Cambridge University Press.

Chomsky, N. 2001. Derivation by Phase. In M. Kenstowicz (ed.). *Ken Hale: A Life in Language*, 1-52. Cambridge, MA: MIT Press.

Chomsky, N. 2004. Beyond Explanatory Adequacy. In A. Belletti (ed.). *Structures and Beyond*, 104-131. Oxford: Oxford University Press.

Chomsky, N. 2007. Approaching UG from Below. In U. Sauerland & H.-M. Gärtner (eds.). *Interfaces + Recursion = Language?*, 1-30. Berlin: De Gruyter.

Chomsky, N. 2008. On Phases. In R. Freidin, C. P. Otero & M.L. Zubizarreta (eds.). *Foundational Issues in Linguistic Theory*, 134-166. Cambridge, MA: MIT Press.

Chomsky, N. 2012a. Some Simple Evo-devo Theses: How True Might They Be for Language? In R.K. Larson, V. Déprez & H. Yamakido (eds.). *The Evolution of Human Language: Biolinguistic Perspectives*, 45-62. Cambridge: Cambridge University Press.

Chomsky, N. 2012b. Foreword. In Á.J. Gallego (ed.). *Phases*, 1-7. Berlin: De Gruyter.

Chomsky, N. 2013. Problems of Projection. *Lingua* 130: 33-49.

Chomsky, N. 2014. Minimal Recursion: Exploring the Prospects. In T. Roeper & M. Speas (eds.). *Recursion: Complexity in Cognition*, 1-15. Berlin: Springer.

Chomsky, N. 2015. Problems of Projection: Extensions. In E. Di Domenico, C. Hamann & S. Matteini (eds.). *Structures, Strategies and Beyond*, 1-16. Amsterdam: John Benjamins.

Chomsky, N. 2017. The Language Capacity: Architecture and Evolution. *Psychonomic Bulletin and Review* 24: 200-203.

Chomsky, N. 2019. Some Puzzling Foundational Issues: The Reading Program. *Catalan Journal of Linguistics Special Issue*: 263-285.
<https://doi.org/10.5565/rev/catjl.287>

Chomsky, N. In press. Puzzles About Phases. In L. Franco & P. Lorusso (eds.). *Linguistic Variation: Structure and Interpretation.* Berlin: De Gruyter.

Chomsky, N. & G. Miller. 1963. Introduction to the Formal Analysis of Natural Languages. In R.D. Luce, R.R. Bush & E. Galanter (eds.). *Handbook of Mathematical Psychology*, vol. II, 269-321. New York: John Wiley.

Chomsky, N. & H. Lasnik. 1993. The Theory of Principles and Parameters. In J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann (eds.). *Syntax: An International Handbook of Contemporary Research*, 506-569. Berlin: De Gruyter.

Cinque, G. 1983. 'Topic' Constructions in Some European Languages and 'Connectedness'. In K. Ehlich & H. van Riemsdijk (eds.). *Connectedness in Sentence, Discourse and Text*. Tilburg: Katholieke Hogeschool.

Cinque, G. 1999. *Adverbs and Functional Heads*. Oxford: Oxford University Press.

Cinque, G. & L. Rizzi. 2010. The Cartography of Syntactic Structures. In B. Heine and H. Narrog (eds.). *The Oxford Handbook of Linguistic Analysis*, 51-65. Oxford: Oxford University Press.

Citko, B. 2005. On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge. *Linguistic Inquiry* 36: 475-497.

Collins, C. 1997. *Local Economy*. Cambridge, MA: MIT Press.

Collins, C. 2017. Merge(X,Y) = {X,Y}. In L.S. Bauke & A. Blümel (eds.). *Labels and Roots*, 47-68. Berlin: De Gruyter.

Collins, C. & E. Groat. 2018. Copies and Repetitions. Ms., NYU. <http://ling.auf.net/lingbuzz/003809>

Collins, C. & E. Stabler. 2016. A Formalization of Minimalist Syntax. *Syntax* 19: 43-78.

Crain, S. & M. Nakayama. 1987. Structure-dependence in Grammar Formation. *Language* 63: 522-543.

Crain, S. & R. Thornton. 1998. *Investigations in Universal Grammar*. Cambridge, MA: MIT Press.

Crain, S. & S. Thornton. 2012. Syntax Acquisition. *WIREs Cognitive Science* 3: 185-203.

Crain, S., L. Koring, and R. Thornton. 2017. Language Acquisition from a Biolinguistic Perspective. *Neuroscience and Biobehavioral Reviews* 81(B): 120-149.

D'Alessandro, R., S. Fischer & G.H. Hrafnbjargarson. 2008. *Agreement Restrictions*. Berlin: DeGruyter.

Den Dikken, M. 2007. Phase Extension: Contours of a Theory of the Role of Head Movement in Phrasal Extraction. *Theoretical Linguistics* 33: 1-41.

Donati, C. 2006. On *Wh*-head Movement. In L. Cheng & N. Corver (eds.). *Wh-Movement: Moving On*, 21-46. Cambridge, MA: MIT Press.

Eguren, L., O. Fernández-Soriano & A. Mendikoetxea (eds.). 2016. *Rethinking Parameters*. Oxford: Oxford University Press.

Epstein, S.D., H. Kitahara & T.D. Seely. 2012. Structure Building That Can't Be. In M. Uribe-Etxebarria & V. Valmala (eds.). *Ways of Structure Building*, 253-270. Oxford: Oxford University Press.

Epstein, S.D., H. Kitahara & T.D. Seely. 2014. Labeling by Minimal Search: Implications for Successive-cyclic A-Movement and the Conception of the Postulate 'Phase'. *Linguistic Inquiry* 45: 463-481.

Epstein, S.D., H. Kitahara & T.D. Seely. 2016a. Phase Cancellation by External Pair-Merge of Heads. *The Linguistic Review* 33: 87-102.

Epstein, S.D., H. Kitahara & T.S. Seely. 2016b. What Do We Wonder is Not Syntactic? In S. Epstein, H. Kitahara & T.D. Seely (eds.). *Explorations in Maximizing Syntactic Minimization*, 222-239. New York: Routledge.

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

Everaert, M., M. Huybregts, N. Chomsky, R. Berwick & J. Bolhuis. 2015. Structures, Not Strings: Linguistics as Part of the Cognitive Sciences. *Trends in Cognitive Sciences* 19: 729-743.

Fanselow, G. 2006. On Pure Syntax (Uncontaminated by Information Structure). In P. Brandt & E. Fuß (eds.). *Form, Structure, and Grammar*, 137-157. Berlin: Akademie Verlag.

Feldman, H., S. Goldin-Meadow & L. Gleitman. 1978. Beyond Herodotus: The Creation of Language by Linguistically Deprived Deaf Children. In A. Lock (ed.). *Action, Gesture, and Symbol*. London: Academic Press.

Fitch, W.T. 2010. *The Evolution of Language*. Cambridge: Cambridge University Press.

Fortuny, J. 2008. *The Emergence of Order in Syntax*. Amsterdam: John Benjamins.

Fox, D. 2000. *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.

Frampton, J. & S. Gutmann. 2002. Crash-Proof Syntax. In S.D. Epstein & T.D. Seely (eds.). *Explanation and Derivation in the Minimalist Program*, 90-105. Oxford: Blackwell.

Friederici, A.D., N. Chomsky, R.C. Berwick, A. Moro & J.J. Bolhuis. 2017. Language, Mind and Brain. *Nature Human Behaviour* 1: 713-722.

Fox, D. 2002. Antecedent-contained Deletion and the Copy Theory of Movement. *Linguistic Inquiry* 33: 63-96.

Fukui, N. & H. Narita. 2014. Merge, Labeling, and Projection. In A. Carnie, D. Siddiqi & Y. Sato (eds.). *The Routledge Handbook of Syntax*, 3-23. New York: Routledge.

Gallego, Á.J. 2010a. *Phase Theory*. Amsterdam: John Benjamins.

Gallego, Á.J. 2010b. Binding Through Agree. *Linguistic Analysis* 34: 163-192.

Gallego, Á.J. 2012. *Phases*. Berlin: De Gruyter.

Gallego, Á.J. 2013a. A Configurational Approach to the Left Periphery. Paper presented at *23rd Colloquium on Generative Grammar*, Universidad Complutense de Madrid, May 9-11.

Gallego, Á.J. 2013b. Object Shift in Romance. *Natural Language and Linguistic Theory* 31: 409-451.

Gallego, Á.J. 2014. Deriving Feature Inheritance from the Copy Theory of Movement. *The Linguistic Review* 31: 41-71.

Gallego, Á.J. 2017. The EPP in Labeling Theory: Evidence from Romance. *Syntax* 20: 384-399.

Gallego, Á.J. 2020. Strong and Weak 'Strict Cyclicity' in Phase Theory. In A. Bárány et al. (eds.). *Syntactic architecture and its consequences*. Language Science Press.

Gallego, Á.J. In progress. A Chomsky Hierarchy-based approach to Information Structure. Ms., Universitat Autònoma de Barcelona.

Gärtner, H.M. 2002. *Generalized Transformations and Beyond*. Berlin: Akademie Verlag.

George, L. & J. Kornfilt. 1981. Finiteness and Boundedness in Turkish. In F. Heny (ed.). *Binding and Filtering*, 105-127. Cambridge, MA: MIT Press.

Hale, K. 1983. Warlpiri and the Grammar of Non-configurational Languages. *Natural Language and Linguistic Theory* 1: 5-47.

Hale, K. & S.J. Keyser. 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In K. Hale & S.J. Keyser (eds.). *The View from Building 20*, 53-109. Cambridge, MA: MIT Press.

Hale, K. & S.J. Keyser. 1999. A Response to Fodor and Lepore, 'Impossible Words'. *Linguistic Inquiry* 30: 453-466.

Harbour, D., D. Adger & S. Béjar (eds.). 2008. *Phi Theory*. Oxford: Oxford University Press.

Hasegawa, H. 2005. Reflexive Binding as Agreement and its Interaction With the Phase System. In N. Imanashi (ed.). *The World of Linguistic Research*, 53-69. Tokyo: Kaitakusha.

Hauser M.D., N. Chomsky N. & W.T. Fitch. 2002. The Faculty of Language: What is It, Who Has It, and How Did It Evolve? *Science* 298: 1569-1579.

Hiraiwa, K. 2005. Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture. PhD dissertation, MIT.

Holmberg, A. 2016. *The Syntax of Yes and No*. Oxford: Oxford University Press.

Hunter, T. 2011. Insertion Minimalist Grammars: Eliminating Redundancies Between Merge and Move. In M. Kanazawa, A. Kornai, M. Kracht & H. Seki (eds.). *The Mathematics of Language*, 90-107. Berlin: Springer.

Hunter, T. 2015. Deconstructing Merge and Move to Make Room for Adjunction. *Syntax* 18: 266-319.

Huybregts, R. 2017. Phonemic Clicks and the Mapping Asymmetry: How Language Emerged and Speech Developed. *Neuroscience and Biobehavioral Reviews* 81(B): 279-294.

Idsardi, W. & E. Raimy. 2013. Three Types of Linearization and the Temporal Aspects of Speech. In I. Roberts & M.T. Biberauer (eds.). *Challenges to Linearization*, 31-56. Berlin: De Gruyter.

Kayne, R.S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R.S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 2011. Why Are There No Directionality Parameters? In *Proceedings of WCCFL 28*, 1-23. Somerville, MA: Cascadilla Proceedings Project.

Kayne, R. 2013. Comparative Syntax. *Lingua* 130: 132-151.

Kayne, R. 2018. The Place of Linear Order in the Language Faculty. Ms., NYU. <http://as.nyu.edu/faculty/richard-s-kayne.html>

Kegl, J., A. Senghas & M. Coppola. 1999. Creation Through Contact: Sign Language Emergence and Sign Language Change in Nicaragua. In M. De Graff (ed.). *Language Creation and Language Change*. Cambridge, MA: MIT Press.

Kobele, G. 2006. *Generating Copies*. PhD dissertation, University of California, Los Angeles.

Landau, I. 2016. Agreement at PF: an Argument from Partial Control. *Syntax* 19: 79-109.

Larson, B. 2018. Right-node Raising and Ungrammaticality. *Studia Linguistica* 72: 214-260.

Lasnik, H. & J. Uriagereka. 2012. Structure. In R. Kempson, T. Fernando & N. Asher (eds.). *Handbook of Philosophy of Science vol. 14: Philosophy of Linguistics*, 33-61. Oxford: Elsevier.

Lebeaux, D. 1988. *Language Acquisition and the Form of the Grammar*. PhD dissertation, University of Massachusetts, Amherst.

Legate, J. 2002. *Warlpiri: Theoretical Implications*. PhD dissertation, MIT.

Legate, J. 2008. Morphological and Abstract Case. *Linguistic Inquiry* 39: 55-101.

Lenneberg, E. 1967. *Biological Foundations of Language*. New York: John Wiley.

Leu, T. 2015. *The Architecture of Determiners*. Oxford: Oxford University Press.

López, L. 2007. *Locality and the Architecture of Syntactic Dependencies*. New York: Palgrave.

López, L. 2009. *A Derivational Syntax for Information Structure*. Oxford: Oxford University Press.

Marantz, A. 1991. Case and Licensing. In G.F. Westphal, B. Ao & H.-R. Chae (eds.). *Proceedings of ESCOL '91*, 234-253. Ithaca, NY: Cornell Linguistics Club.

Marantz, A. 2001. Words. Ms., MIT.

Marantz, A. 2007. Phases and Words. In S.-H. Choe (ed.). *Phases in the Theory of Grammar*, 191-222. Seoul: Dong In.

Marantz, A. 2013. Locality Domains for Contextual Allomorphy Across the Interfaces. In O. Matushansky & A. Marantz (eds.). *Distributed Morphology Today*, 95-116. Cambridge, MA: MIT Press.

Mateu, J. 2005. Impossible primitives. In M. Werning, E. Machery & G. Schurz. (eds.). *The Compositionality of Meaning and Content: Foundational Issues*, 213-225. Frankfurt: Ontos.

Mathieu, E. 2016. The *Wh*-parameter and Radical Externalization. In L. Eguren, O. Fernández-Soriano and A. Mendikoetxea (eds.). *Rethinking Parameters*, 252-290. Oxford: Oxford University Press.

McFadden, T. 2004. *The Position of Morphological Case in the Derivation*. PhD dissertation, University of Pensylvania.

Melchin, P. 2018. *The Semantic Basis for Selectional Restrictions*. PhD dissertation, University of Ottawa.

Moro, A. 2000. *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.

Musso, M., A. Moro, V. Glauche, M. Rijntjes, J. Reichenbach, C. Büchel & C. Weiller. 2003. Broca's Area and the Language Instinct. *Nature Neuroscience* 6: 774-781.

Nelson, M.J., I. El Karoui, K. Giber, X. Yang, L. Cohen, H. Koopman, S.S. Cash, L. Naccache, J.T. Hale, C. Pallier & S. Dehaene. 2017. Neurophysiological Dynamics of Phrase-structure Building During Sentence Processing. *PNAS* 114: E3669-E3678.

Nunes, J. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.

Obata, M. 2010. *Root, Successive-Cyclic and Feature-Splitting Internal Merge: Implications for Feature-Inheritance and Transfer*. PhD dissertation, University of Michigan, Ann Arbor.

Obata, M., S. Epstein & M. Baptista. 2015. Can Crosslinguistically Variant Grammars be Formally Identical? Third-factor Underspecification and the Possible Elimination of Parameters of UG. *Lingua* 156: 1-16.

Oishi, M. 2015. The Hunt For a Label. In H. Egashira, H. Kitahara, K. Nakazawa, A. Nishimae, T. Nomura, M. Oishi & I. Suzuki (eds.). *In Untiring Pursuit of Better Alternatives*, 322-334. Tokyo: Kaitakusha.

Ott, D. 2011. A Note on Free Relative Clauses in the Theory of Phases. *Linguistic Inquiry* 42: 183-192.

Ott, D. 2012. *Local Instability*. Berlin: De Gruyter.

Ott, D. 2014. An Ellipsis Approach to Contrastive Left-dislocation. *Linguistic Inquiry* 45: 269-303.

Ott, D. 2016a. Fragment Anchors Do Not Support the Syntactic Integration of Appositive Relative Clauses: Reply to Griffiths & de Vries 2013. *Linguistic Inquiry* 47: 580-590.

Ott, D. 2016b. Ellipsis in Appositives. *Glossa* 1: article 34.

Ott, D. 2017a. The Syntax and Pragmatics of Dislocation: A Non-templatic Approach. In A. Monti (ed.). *Proceedings of the 2017 annual conference of the Canadian Linguistic Association*. <http://cla-acl.ca/actes-2017-proceedings/>

Ott, D. 2017b. Strong Generative Capacity and the Empirical Base of Linguistic Theory. *Frontiers in Psychology* 8: 1617.

Ott, D. & E. Onea. 2015. On the Form and Meaning of Appositives. In *Proceedings of NELS 45*, 203-212.

Ouali, H. 2008. On C-to-T φ-feature Transfer. In D'Alessandro, R., S. Fischer & G.H. Hrafnbjargarson (eds.). *Agreement Restrictions*, 159-180. Berlin: Mouton de Gruyter.

Pesetsky, D. & E. Torrego. 2006. Probes, Goals, and the Nature of Syntactic Categories. In Y. Otsu (ed.). *Proceedings of the Seventh Tokyo Conference on Psycholinguistics*, 25-60. Tokyo: Hituzi Syobo Publishing Company.

Picallo, M.C. (ed.). 2014. *Linguistic Variation in the Minimalist Framework*. Cambridge: Cambridge University Press.

Preminger, O. 2014. *Agreement and Its Failures*. Cambridge, MA: MIT Press.

Quine, W.V. 1940. *Mathematical Logic*. New York: W.W. Norton & Co.

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

Reinhart, T. 2006. *Interface Strategies: Optimal and Costly Computations*. Cambridge, MA: MIT Press.

Reuland, E. 2011. *Anaphora and Language Design*. Cambridge, MA: MIT Press.

Richards, M. 2007. On Feature Inheritance: An Argument from the Phase Impenetrability Condition. *Linguistic Inquiry* 38: 563-572.

Richards, M. 2008. Two Kinds of Variation in a Minimalist System. In F. Heck, G. Müller & J. Trommer (eds.). *Varieties of Competition*, 133-162. Linguistische Arbeitsberichte 87, Universität Leipzig.

Richards, M. 2012. On Feature Inheritance, Defective Phases, and the Movement–Morphology Connection. In Á.J. Gallego (ed.). *Phases*, 195-232. Berlin: De Gruyter.

Richards, N. 2016. *Contiguity Theory*. Cambridge, MA: MIT Press.

Rizzi, L. 1997. The Fine Structure of the Left Periphery. In L. Haegeman (ed.). *Elements of Grammar*, 281-337. Dordrecht: Kluwer.

Rizzi, L. 2013a. Introduction: Core Computational Principles in Natural-language Syntax. *Lingua* 130: 1-13.

Rizzi, L. 2013b. Notes on Cartography and Further Explanation. *Probus* 25: 197-226.

Rizzi, L. 2015. Cartography, Criteria, and Labeling. In U. Shlonsky (ed.). *Beyond Functional Sequence*, 314-338. Oxford: Oxford University Press.

Roberts, I. 2010. *Agreement and Head Movement*. Cambridge, MA: MIT Press.

Sandler, W. & D. Lillo-Martin. 2006. *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Seely, T. D. 2006. Merge, Derivational C-command, and Subcategorization in a Label-Free Syntax. In C. Boeckx (ed.). *Minimalist Essays*, 182-217. Amsterdam: Benjamins.

Smith, N. & I.-M. Tsimpli. 1995. *The Mind of a Savant: Language Learning and Modularity*. Oxford: Blackwell.

Sportiche, D. 2013. Binding Theory. Structure Sensitivity of Referential Dependencies. *Lingua* 130: 187-208.

Sportiche, D. 2015. Neglect. Ms., UCLA.

Starke, M. 2014. Cleaning Up the Lexicon. *Linguistic Analysis* 39: 245-256.

Struckmeier, V. 2016. Against Information Structure Heads: A Relational Analysis of German Scrambling. *Glossa* 2: Article 1, 1-29.

Tallerman, M. 2005. *Understanding Syntax*. New York: Routledge.

Toyoshima, T. 2000. *Head-to-Spec Movement and Dynamic Economy*. PhD dissertation, Cornell University.

Travis, L. 1984. *Parameters and Effects of Word Order Variation*. PhD dissertation, MIT.

Trinh, T. 2011. *Edges and Linearization*. PhD dissertation, MIT.

Uriagereka, J. 1999. Multiple Spell-Out. In S. Epstein & N. Hornstein (eds.). *Working Minimalism*, 251-282. Cambridge, MA: MIT Press.

Uriagereka, J. 2003. Evidential Contexts. In J. Guéron & L. Tasmowski (eds.). *Tense and Point of View*, 367-394. Paris: Université Paris X.

Uriagereka, J. 2012. *Spell-out and the Minimalist Program*. Oxford: Oxford University Press.

Van Urk, C. 2016. On the distribution of reflexes of successive cyclicity. Paper presented at *Brussels Conference on Generative Linguistics* 9, Brussels, December 13-14.

Vergnaud, J. R. 1977 [2006]. Letter to Noam Chomsky and Howard Lasnik. Reprinted in R. Freidin & H. Lasnik (eds.). *Syntax: Critical Concepts in Linguistics*, vol. 5: 21-34. New York: Routledge.

Vermeerbergen, M., L. Leeson & O. Crasborn. 2007. Simultaneity in Signed Languages: A String of Sequentially Organised Issues. In M. Vermeerbergen, L. Leeson & O. Crasborn (eds.). *Simultaneity in Signed Languages*, 1-26. Amsterdam: John Benjamins.

De Vries, M. 2009. On Multidominance and Linearization. *Biolinguistics* 3: 344-403.

De Vries, M. 2012. Unconventional Mergers. In M. Uribe-Etxebarria & V. Valmala (eds.). *Ways of Structure Building*, 143-166. Oxford: Oxford University Press.

Webelhuth, G. 1992. *Principles and Parameters of Syntactic Saturation*. Oxford: Oxford University Press.

Willer Gold, J. et al. 2018. When Linearity Prevails over Hierarchy in Syntax. *PNAS* 115: 495-500.

Wurmbrand, S. 2006. Licensing Case. *Journal of Germanic Linguistics* 18: 175-236.

Yang, C. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

Yang, C. 2016. *The Price of Linguistic Productivity*. Cambridge, MA: MIT Press.

Yang, C., S. Crain, R. C. Berwick, N. Chomsky, J. Bolhuis. 2017. The Growth of Language: Universal Grammar, Experience, and Principles of Computation. *Neuroscience & Biobehavioral Reviews* 81(B): 103-119.

Zwart, J.W. 2017. An Argument Against the Syntactic Nature of Verb Movement. In L.R. Bailey & M. Sheehan (eds.). *Order and Structure in Syntax I: Word Order and Syntactic Structure*, 29-47. Berlin: Language Science Press.