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

Noam Chomsky
Massachusetts Institute of Technology

Ángel J. Gallego
Universitat Autònoma de Barcelona

Dennis Ott
University of Ottawa

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 Decartes' 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 simply 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 us? While we may never find a satisfying answer to the latter question, any theory of UG seeking to address the former must meet a criterion of *evolvability*: any mechanisms and primitives ascribed to UG rather than derived from independent factors must plausibly have emerged in what appears to have been a unique 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 to appear).

[*] To appear in Á.J. Gallego & D. Ott (2018). *Generative Syntax: Questions, Crossroads, and Challenges*. Special issue of *Catalan Journal of Linguistics*

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. in press). 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; see Yang 2016 on the attained state as a probability distribution over I-languages). Converging conclusions follow from 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 some of 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 under minimalist desiderata. Section 6 summarizes.

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 nature remains the subject of much 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] 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 dramatically across theoretical frameworks, of course.

[2] See Hale & Keyser 1993, 1999, Borer 2005, Marantz 2001, 2013, Mateu 2005, Ramchand 2008, and Starke 2014 for a sample.

(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 (displaced from his base position) in the latter. On the assumption that thematic relations are established in a uniform and 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 expressions with displacement. The optimal course to follow, we think, is to assume the 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 collection of occurrences of Y within the structured object assembled. (2) illustrates for (1b) above (abstracting away from irrelevant details), where the NP *John-ga* is internally merged to K:

(2) a. {sensei-ni,{sikarareta,John-ga}} = K → MERGE(K,John-ga)
b. {John-ga,{sensei-ni,{sikarareta,John-ga}}} = K′

MERGE, applying recursively so that any already generated object is accessible to further operations,[3] thus suffices in principle to model the basic properties of discrete infinity and displacement. Furthermore, it is the computationally simplest operation 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 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).

[3] Recursion is thus a "deep" property of the generative procedure; to what extent constructions displaying category recursion are used in some particular language (e.g., English but not German permits recursive possessors, as in *Maria's neighbor's friend's house*) is an entirely different issue. See Arsenijević & Hinzen 2012, Chomsky 2014, and references therein.

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 spectrum of variation defined by UG. Neurolinguistic studies conducted by Andrea Moro and associates 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 had been 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 (see Rizzi 2013 and Everaert et al. 2015 for surveys). 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 under-determines 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?

The computation choses 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 straightfoward. 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. in press):

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

The linearity-based but formally straightforward "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 *highest* auxiliary raise, exactly like 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 inversion 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 arrangement. 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 nature of linguistic objects imposed 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, a task undertaken elsewhere (see Berwick et al. 2011, Crain et al. in press and references therein). The approaches fail invariably at adequately capturing the phenomena they focus on, and more fundamentally at addressing the only theoretically relevant question: why do languages universally adopt operations on structure while avoiding, in all relevant cases, the far simpler computational operations 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, as it is 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 intentional planning (e.g., discourse organization) and articulation (internal in self-talk, external in oral or gestural production) at the *interfaces* of I-language and performance systems. 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 are pronounced in their derived position. Consider a standard example of *wh*-movement in French (from Sportiche 2013):

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

The *wh*-phrase *de quel livre* 'of which book' is displaced from its original position (α) as the complement of the noun *publication*, by internally merging at the left edge of the embedded clause, where it surfaces in the externalized form. At SEM, the resulting chain (a collection 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 unnecesary. By contrast, mapping to PHON involves a choice as to where to pronounce the single discontinuous object created by IM. The typical choice is the highest position, with all lower copies remaining unpronounced. To what extent and under which conditions this preference can be overriden by parametric and other factors remains an important research question (Trinh 2011, Richards 2016).

Whether other types of displacement commonly found in the world's languages—semantically vacuous scrambling, extraposition, clitic movement etc.—likewise reflect narrow-syntactic computations or rather part of the mapping to PHON is an open question. It is commonly presupposed that effects on meaning pertaining to topic/focus articulation necessarily indicate core-syntactic displacements, but the notion of "meaning" invoked in this argument encompasses pragmatic as well as externalization-related (e.g., prosodic) properties of expressions. "Meaning" properties in this broad sense may well emerge from holistic interpretation of <SEM,PHON> pairs, rather than narrow 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 current understanding of the evolution of UG. Bolhuis et al. (2014) and Berwick & Chomsky (2016) suggest that MERGE, given its computational primitiveness, is indeed a plausible candidate for sudden emergence 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 necessary to construct a discrete infinity of syntactic objects, its emergence in our species is a minimal prerequisite for the linguistic mind. The emergence of the other central component of I-language—the lexicon and its atoms with all their semantic complexity—remains a deep mystery.

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 (for discussion and review, see Epstein et al. 2014, Fukui & Narita 2014, Collins in press).

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 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 this is the *only* operation defined by UG (although perhaps adjunction structures 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 and 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.

Simplest MERGE automatically satisfies 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 incoporates no notion of "projection" (Chomsky 2013, 2015). IC also bars introduction of movement-inducing features that are not inherent to lexical items, such as the discourse-related features (topic, focus, etc.) assumed in the cartographic tradition (Rizzi 1997 and much work since) and elsewhere (e.g. López 2009). We suggest below that simplest MERGE is generally not triggered but applies freely, and return to the specific problems raised by cartography in section 5 below. IC need not be stipulated: it is an axiom of Simplest MERGE.

Having constructed K = {X,Y}, we may want 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 ("movement"). 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 the unaccusative construction *Mary's book arrived Mary's book*).

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 2000 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 in press 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" being only the most obvious case.[4] 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 & Šimík to appear).[5]

MERGE thus applies freely, generating expressions that receive whatever interpretation they are assigned by external systems. We should be careful to distinguish "interpretive systems" from "performance systems." The interpretive SM (sensorimotor) and C-I (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. 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. 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), as has been an assumption since Chomsky 1955[1975]. Consequently, concerns about "overgeneration" in core syntax are unfounded; the only empirical criterion is that the grammar associate each SO generated to a <SEM,PHON> pair in a way that accurately reflects the knowledge of the native speaker.[6]

Do we need operations other than MERGE for the construction of syntactic objects? Agreement phenomena indicate that there is an algorithm that relates *features* of syntactic objects, called AGREE (Chomsky 2000, 2001). The assumption of much current work is that AGREE is

[4] The "edge features" of Chomsky 2008 are equally dispensable while not technically equivalent, and originally introduced to distinguish elements that enter into computation from those that do not, such as interjections.

[5] 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 to appear.

[6] The claim that syntactic computation should be "crash-proof" (cf. Frampton & Gutmann 2002, among others) is based the dubious assumption that an I-language defines a set of well-formed, hence "acceptable" 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.

asymmetric, relating initially unvalued φ-features on a *Probe* to matching, inherent φ-features of a *Goal* within the 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)
1sg-3sg-seek flower
'I seek a flower.'

(13) Uqa jo ceh-ade-ia. (Amele)
he houses build-3pl-3sg.pst
'He built houses.'

Empirically, AGREE is clearly required. Whether it is part of the narrow-syntactic computation—as assumed by theories linking AGREE to assignment of structural Case (Chomsky 2000 *et seq.*)—or rather part of the mapping to PHON in order to determine the morphological shape of underspecified forms remains an open question. We set aside here many intricacies of agreement phenomena uncovered in much detailed work on the topic (see, e.g., Bobaljik 2008, Harbour et al. 2008, Legate 2008, and references therein).

In earlier work it was 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 the Spanish sentence in (14), where the matrix verb *parecen* 'seem' agrees with the *in situ* NP *varios sobornos a políticos* 'many bribes to politicians', as well as the participle *descubiertos* 'discovered', irrelevantly for Case assignment.

(14) Parecen haber sido descubiertos varios sobornos a políticos. (Spanish)
seem.3pl have.inf been discovered.3pl many bribes to politicians
'Many bribes to politicians seem to have been discovered'

The NP can raise into the matrix clause but it need not, unlike in other languages such as English. These and many other cases in various languages show that IM and AGREE are independent operations.[7]

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 "flattening" of hierarchical structure and computation of stress, prosody etc. (see Collins in press for a partial theory of this mapping, and Idsardi & Raimy 2013

[7] 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 movements are dubious on grounds of learnability alone.

for general discussion). 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.

What the effects of TRANSFER on the syntactic representation are is a matter of debate. Ideally, it 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 should not 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 2000 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 (see Obata 2010 for discussion). To illustrate, in (15) the NP α contains the clausal phase β:

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

Suppose we were to eliminate β and then raise α to a higher position, obtaining (16):

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

The clausal phase β is pronounced in its derived position internal to displaced α; it is not pronounced in its original position (or eliminated from the final string). What this means is that literal *Spell-Out* does not exist and no structure is eliminated: there is only TRANSFER, which renders β inaccessible to subsequent manipulation.[8]

At the C-I interface, global principles of interpretation such as Condition C of the Binding Theory or the unbounded character of operator-variable dependencies (including "reconstruction" effects, as in (11) above) suggest the same conclusion: transferred phases remain accessible, but they cannot be modified by syntactic operations (MERGE, AGREE) at later cycles. This is a version of the *Phase Impenetrability Condition* (PIC). Note that the PIC may permit Probe-Goal relations across phase boundaries, as long as these only affect properties of the Probe. Examples are 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 (see D'Alessandro et al. 2008 and Richards 2012 for relevant discussion).[9]

---

[8] We thus avoid what Collins & Stabler (2016) dub the *assembly problem*, first discussed in Uriagereka 1999.

[9] 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.

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 constitutes an alteration of P's internal structure.[10] 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 all 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, and an assumption that the edge remains accessible at the next phase. In this way, PIC gives rise to familiar indications of cyclic movement (Boeckx 2007).

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

The verbal and clausal phases, transitive/unergative vP and clausal CP, in essence capture the "duality of interpretation" stated in terms of the D-structure/S-structure distinction of earlier theories. EM within the vP phase gives rise to configurations that represent generalized argument structure (see Hale & Keyser 1993 and related work), whereas IM at the CP cycle creates chains that enter into the determination of scope/discourse properties of expressions at C-I (Chomsky 2004, 2007, Gallego 2013, 2016). While this is a reasonable approximation, apparent exceptions (such as semantically vacuous A-movement and scrambling) pose interesting research questions.

The basic operations MERGE, AGREE, and TRANSFER raise many further questions, which require sharpening; we will address some in the following two sections.[11] Despite many open issues, it is important to appreciate that a minimal system as outlined so far can accommodate a significant range of facts about natural language that are equally basic 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.

[10] 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 may suggest that the NTC is reducible to the PIC (see Gallego 2017).

[11] 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).

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 not eliminated from the syntactic representation (permitting phasal objects to be moved as part of larger objects, as discussed above).

The interpretive and perceptual/articulatory performance systems accessing PHON and SEM impose indirect 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.[12] As a result, (17) 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.

(17) {who,{John{saw,Mary}}}

So-called "crash-proof" models seek to bar generation of structures such as (17), given the intuitive "ill-formedness" of the derivative string (Frampton & Gutmann 2002). We think this is a mistake, for both conceptual and empirical reasons (cf. note 2). On methodological grounds, constraints imposed on MERGE are typically redundant with more general interface conditions, such as Full Interpretation in the case of (17) (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: this failure 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 so as to be sensitive to selectional features or the like.[13] Furthermore, "deviant" expressions typically do have some interpretation, however inexpedient.

More specific constraints are imposed by C-I on particular elements within W, such as the types of nominals 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

---

[12] 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 mechanism that removes these features is required.

[13] 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 (2017) for *eat*/*devour*-type contrasts. Idiosyncratically selected functional prepositions plausibly fall under a general theory of morphological case realized as part of EXT.

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).[14]

What about the other interface, which relates the core computational system to articulatory and perceptual systems involved in externalization? 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. 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 highly variable across languages. 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. The mapping must be sufficiently general to accommodate the contingencies of both 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 variation between languages (in accord with Chomsky's 2001 *Uniformity Principle*).[15] A strong thesis is that EXT is altogether ancillary to I-language, drawing on distinct properties and operations (Berwick & Chomsky 2011, 2016).

Psycholinguistic and neurolinguistic inquiries have the potential to shed considerable light on the issue. 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 arguably are not available to UG, such as linear order. It turned out that Chris was completely unable to deal with the language using 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 suggested that normals could likewise relatively easily deal with languages conforming to UG principles, but could handle the non-UG conforming systems using linear order only if they were expressly presented to them as a puzzle rather than a language. While preliminary, these results strike us as suggestive (see also Musso et al. 2003).

These observations support the speculation that those properties of language that pertain exclusively to perception and articulation are ancillary, perhaps external to I-language, whereas the core computational system is close to uniform (Berwick & Chomsky 2016).[16] EXT relates very different systems, a computational system constructing hierarchical expressions on the one

---

[14] 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.

[15] For related discussion and developments in the study of parametric variation, see Biberauer et al. 2014, Picallo 2014, Eguren et al. 2016, Kayne 2013, and references therein.

[16] 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 of this sort.

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. 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, as noted above. See Huybregts 2017 for relevant recent discussion, including speculations concerning the evolutionary relevance of aerially isolated click phonemes.

Where does all of this leave us with regard to the question of evolvability? As noted above, 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, there are proposals that link these operations to considerations of efficient computation. Chomsky (2013, 2015) suggests that AGREE instantiates *minimal search* within the syntactic object, in which case its formal properties (structure-dependence, minimality) need not be part of UG but would rather reflect natural properties of computation. For TRANSFER and the interface mappings, the question is how complex they are: necessarily quite complex for PHON, but perhaps near-trivial for SEM. A plausible speculation is that EXT and its cross-linguistically variable properties reflect not UG specifications but rather a lack thereof, if the linkage established between the computational system proper and externalization systems was a problem to be solved subsequent 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 posulated in some earlier and recent work; cf. Chomsky 1995, Collins & Stabler 2016). 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}] (cf. Chomsky 1995:243)? The latter view is more restrictive, and arguably more in line with basic desiderata about 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.[17] 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 rules can apply in a determinate fashion: any rule referencing X or Y would ambiguously refer to the individual terms and their occurrences within K = {X,Y}. 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 on which see 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 ensuring that subsequent rules can unambiguously refer to X and Y.

This restrictive view of MERGE as a strictly binary rule seeks to curtail the complexity of WS in effect rules out 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).[18] This leaves EM and IM as the only instantiations of Simplest MERGE.[19] We believe that future work should address the various questions raised by these considerations, in order to establish a restrictive "null theory" of the generative procedure that adheres to plausible 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)?

Different answers to this question have been pursued, e.g. in terms of multidominance 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

---

[17] 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.

[18] See Sportiche 2015 for an alternative treatment of the facts motivating Late Merge analyses in terms of "neglect" at the interface.

[19] We leave open here what the consequences of the stricter view would be for the theory of head movement; see Carstens et al. 2016, Epstein et al. 2016a, Chomsky 2015.

operation is necessary given that IM automatically yields copies (on standard set-theoretic assumptions). The question, thus, is really how repetitions are recognized as such. One 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 Y, Y′ was formed by IM.[20] 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 to 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.[21]

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 endocentricity 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}}, with L ∈ {X,Y}. But this is a departure from Simplest MERGE, rooted in part due in the intuitive appeal and pedagogical convenience of tree diagrams. In its simplest form, MERGE has no "built-in" projection mechanism, hence does not yield labeled objects (Chomsky 2013, 2015, Collins to appear). 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 (cf. 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 a *labeling 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

---

[20] 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.

[21] Earlier theories avoided the problem by assuming a rewriting of lower copies as distinct symbols (traces), linking these to their antecedent via coindexing, but in radical violation of IC.

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, which yields an unlabelable {XP,YP} configuration (putting aside head movement). Another case in point is the DP hypothesis, a problematic corollary of X-bar Theory. For instance, 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, e.g. 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 by Oishi (2015) (see also Chomsky 2007) 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 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. 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 and v*, respectively; see Gallego in press for an alternative, and Epstein et al. 2016b for further relevant discussion).

Again building on Moro's work, Ott (2012), Chomsky (2013, 2015), and Blümel (2017) argue that the need to "break" {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 conform to interface conditions. 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.

It is quite possible that a constructed object K can, in at least some cases, remain unlabeled (exocentric), e.g. when K is a root clause or created by operations that are not in any plausible way head-driven, e.g. syntactic scrambling. Further illumination of these issues will require a theory of where endocentricity is required: in the syntactic derivation (say, for purposes of interpreting local selectional relations), at the interfaces (say, for the computation of prosody), both, or not at all (Collins to appear)? These questions remain open for now and are in 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 (18) is ambiguous between a complementation and an adjunction structure, (19) 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 (20), an analogous extraction from a corresponding parenthetical appositive NP yields no coherent interpretation (21).

(18) John decided on the boat. (ambiguous)

(19) What did John decide on _? (unambiguous)

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

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

Chomsky (2004) proposes that adjunction is the result of a separate operation PAIR-MERGE, which yields asymmetric objects (ordered pairs) rather than symmetric ones (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 problem that goes back to early work of the 1960s. Even unrestricted rewriting systems cannot generate these structures, nor can transformations (see Lasnik & Uriagereka 2012 for a critical review of some proposals in Chomsky & Miller 1963).[22]

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. That is, 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),[23] but such an approach has several unwelcome consequences (Chomsky 2015:12ff.). It has long been observed that HM violates principles of minimal computation and cannot be implemented by Simplest MERGE, given its countercyclic application. It also typically lacks semantic effects, at least for the core cases of verb movement. This vacuity and the fact that the configurations standardly described in terms of HM are highly variable across languages suggest that HM might fall within the mapping to PHON (as suggested in Chomsky 2001 and supported by specific arguments in Zwart to appear and elsewhere), although there are interesting arguments to the contrary (see Roberts 2010 and references therein).[24]

---

[22] A possibility that remains to be explored is that such constructions involve ellipsis.

[23] See Epstein et al. 2016a on PAIR-MERGE as a mechanism for affixation.

[24] 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.

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 ("subject," "object," etc.; cf. Chomsky 1965) or thetamatic roles ("agent," "theme," etc.; cf. Hale & Keyser 1993) are properties of configurations and their syntactic or discursive context, not of individual syntactic objects; consequently, they should neither be represented in the lexicon, nor in the narrow syntactic derivation (cf. Fortuny 2008; Gallego 2013, 2016; López 2009; Uriagereka 2003).

The Cartographic Program pursued by Cinque, Rizzi and many others has revealed some remarkable facts and generalizations, such as Cinque's (1999) hierarchy of adverbial positions and Rizzi's (1997) template for the left periphery. But there are serious problems, which should lead us to recognize that the theoretical conclusions of the program cannot be correct. As we remarked above, any linguistic theory is going to have to meet two conditions: the conditions of acquirability and evolvability. UG must permit acquisition of I-language, and UG must have evolved in the human lineage—and if current best guesses are correct, it must have evolved recently. Work in the cartographic tradition fails to meet these criteria. 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 learn these templates from experience. Consequently, the templates must be attributed directly to UG. But this solution of the acquisition question raises an evolutionary question: it seems virtually unimaginable that the complex cartographic templates could have evolved as part of UG. The conclusion is that cartographic generalizations must derive from more elementary principles that are available independently, obviating the need to enrich UG with positional templates.

There is some promising work in this direction, such as Ernst's (2002) non-templatic analysis of adjunct ordering that derives the effects of the template from interpretive properties of adverbial expressions. Developing alternatives to templatic approaches to the clausal peripheries will require, we believe, a re-evaluation of whether all peripheral surface complexity in fact reflects syntactic composition, rather than amalgamation of independent expressions in discourse. In contrast to early work on "topic constructions" (e.g., Cinque 1983), the cartographic tradition assumes that virtually all sorts of peripheral phrases, including left- and right-dislocated elements, are part of core 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, 2017, 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. In this way, the approach renders peripheral templates – and their adverse implications for evolution and acquisition – obsolete.

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:

(22) I-language is optimized relative to the C-I interface alone, with externalization a secondary phenomenon.

The adjective "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 (22) emerge from recent work suggesting a rather direct involvement of morphophonological factors in the syntactic computation. Richards (2016) develops an elaborate theoretical framework according to 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 on *wh*-movement). In this model, metrical requirements of affixes and other conditions imposed by PHON can effect the application of MERGE and other operations.[25] Given the impressive results achieved by Richards' system, his work poses an interesting challenge to the hypothesis that EXT is an ancillary process. For further challenges, see Kayne 2011 and Bruening 2014, among others. Whether or not the results of these works can be reconciled with (22) is an important topic for future research.

A related open question pertaining to the overall organization of the system is whether AGREE operates as part of the narrow-syntactic computation, or is restricted to the mapping to PHON. The former, traditional view is primarily based on the assumption that AGREE, induced by semantically redundant φ-features, mediates core-syntactic assignment of structural Case (Chomsky 2000 *et seq.*, building on observations of Vergnaud 1977/2006 and George & Kornfilt 1981). Another possibility is that case is a purely morphological phenomenon (i.e., there is no Case; see Marantz 1991, McFadden 2004), and that uninterpretable features are simply neglected at the C-I interface, without requiring elimination (in the spirit of Sportiche 2015). The latter scenario is consistent with relegating AGREE to EXT, where it would then serve the purpose of determining the eventual form of initially underspecified inflectional elements (cf. Bobaljik 2008, and Preminger 2014 for an opposing view). Also in view of the cross-linguistically highly variable expression of agreement, 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.

---

[25] 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. For further discussion, see Ott forthcoming.

6. Conclusions

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 emergence of GG in 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 we have tried to document in this survey, the formal toolkit centered around 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 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, and in many other cases.

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

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

Anderson, S.R. 2004. *Doctor Dolittle's Delusion. Animals and the Uniqueness of Human Language*. New Haven, CT: Yale University Press.

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

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

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

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

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

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

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

Bolhuis, J., I. Tattersall, N. Chomsky, and 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 and 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. Its Nature, Origin, and Use*. New York: Praeger.
Chomsky, N. 1993. "A Minimalist Program for Linguistic Theory". In K. Hale and S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 1-52. Cambridge, MA: MIT Press.
Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, N. 2000. "Minimalist Inquiries: The Framework". In R. Martin, D. Michaels, J. Uriagereka and S.J. Keyser (eds.), *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, MA: MIT 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. The Cartography of Syntactic Structures vol. 3*, 104-131. Oxford: Oxford University Press.
Chomsky, N. 2007. "Approaching UG from Below". In U. Sauerland and H.-M. Gärtner (eds.), *Interfaces + Recursion = Language?*, 1-30. Berlin: De Gruyter.
Chomsky, N. 2008. "On Phases". In R. Freidin, C. P. Otero, and M.L. Zubizarreta (eds.), *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, 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, and 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. Developing the Framework*, 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 and M. Speas (eds.), *Recursion: Complexity in Cognition*, 1-15. Berlin: Springer.
Chomsky, N. 2015. "Problems of Projection: Extensions". In E. Di Domenico, C. Hamann, and S. Matteini (eds.), *Structures, Strategies and Beyond*, 1-16. Amsterdam: John Benjamins.
Chomsky, N. to appear. "Puzzles About Phases". In L. Franco and G. Bellucci (eds.), *Linguistic Variation: Structure and Interpretation. A Festschrift in Honour of M. Rita Manzini.* Berlin: De Gruyter.

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

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

Chomsky, N. and 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 and H. van Riemsdijk (eds.), *Connectedness in Sentence, Discourse and Text*. Tilburg: Katholieke Hogeschool.

Cinque, G. 1999. *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. 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. in press. “Merge(X,Y) = {X,Y}”. In L.S. Bauke and A. Blümel (eds.), *Labels and Roots*. Berlin: De Gruyter.

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

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

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

Crain, S. and S. Thornton. 2012. “Syntax acquisition”. *WIREs Cognitive Science* 3 : 185-203

Crain, S., L. Koring, and R. Thornton. In press. “Language Acquisition from a Biolinguistic Perspective”. *Neuroscience and Biobehavioral Reviews*.

D’Alessandro, R., S. Fischer, and 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 and N. Corver (eds.), *Wh-Movement: Moving On*, 21-46. Cambridge, MA: MIT Press.

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

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

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

Epstein, S.D., H. Kitahara, and T.S. Seely. 2016b. “What Do You Wonder’ is Not Syntactic”. In S. Epstein, H. Kitahara and 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 and 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 and E. Fuß (eds.), *Form, Structure, and Grammar: A Festschrift to Günther Grewendorf on Occasion of His 60th Birthday*, 137-157. Berlin: Akademie Verlag.

Feldman, H., S. Goldin-Meadow and L. Gleitman. 1978. Beyond Herodotus: The creation of language by linguistically deprived deaf children. In A. Lock (ed.), *Action, gesture, and symbol: The emergence of Language*. London: Academic 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. and S. Gutmann. 2002. “Crash-Proof Syntax”. In S.D. Epstein and T.D. Seely (eds.), *Explanation and Derivation in the Minimalist Program*, 90-105. Oxford: Blackwell.

Friederici, A.D. to appear. *Language in Our Brain*. Cambridge, MA: MIT Press.

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

Fukui, N. and H. Narita. 2014. “Merge, Labeling, and Projection”. In A. Carnie, D. Siddiqi, and 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. Developing the Framework*. Berlin: De Gruyter.

Gallego, Á.J. 2013. “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. 2014. “Deriving Feature Inheritance from the Copy Theory of Movement”. *The Linguistic Review* 31: 41-71.

Gallego, Á.J. in press. “The EPP in Labeling Theory: Evidence from Romance”. *Syntax*.

Gallego, Á.J. 2016. “The Basic Elements of the Left Periphery”. Ms., UAB.

Gallego, Á.J. 2017. “Strong and weak “strict cyclicity” in Phase Theory”. To appear in a *Festschrift*.

Gärtner, H.M. 2002. *Generalized Transformations and Beyond. Reflections on Minimalist Syntax*. Berlin: Akademie Verlag.

George, L. and 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. and S.J. Keyser. 1993. “On Argument Structure and the Lexical Expression of Syntactic Relations”. In K. Hale and S.J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 53-109. Cambridge, MA: MIT Press.

Hale, K. and S.J. Keyser. 1999. “A Response to Fodor and Lepore, ‘Impossible words’”. *Linguistic Inquiry* 30: 453-466.

Harbour, D., D. Adger and S. Béjar (eds.). 2008. *Phi Theory. Phi-Features Across Modules and Interfaces*. 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. A Festschrift for Kinsuke Hasegawa on Occasion of his Seventieth Birthday*, 53-69. Tokyo: Kaitakusha.

Hauser M.D., N. Chomsky N. and 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.
Hunter, T. 2011. "Insertion Minimalist Grammars: Eliminating Redundancies Between Merge and Move". In M. Kanazawa, A. Kornai, M. Kracht, and H. Seki (eds.), *The Mathematics of Language. Lecture Notes in Computer Science*, vol. 6878: 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*.
Idsardi, W. and E. Raimy. 2013. "Three Types of Linearization and the Temporal Aspects of Speech". In I. Roberts and 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.
Kegl, J., A. Senghas, and 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: Creolization, Diachrony, and Development*. Cambridge, MA: MIT Press.
Kobele, G. 2006. Generating Copies: An Investigation Into Structural Identity in Language and Grammar. PhD Dissertation, University of California, Los Angeles.
Lasnik, H. and J. Uriagereka. 2012. "Structure". In R. Kempson, T. Fernando and 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, and 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 and A. Marantz (eds.), *Distributed Morphology Today: Morphemes for Morris Halle*, 95-116. Cambridge, MA: MIT Press.
Mateu, J. 2005. "Impossible primitives". In M. Werning, E. Machery, and 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. A Study on the Syntax-Morphology Interface. PhD Dissertation, University of Pensylvania.

Melchin, P. 2017. “Argument Omission and Verb Meaning”. Ms., University of Ottawa.

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

Musso, M., A Moro, V. Glauche, M. Rijntjes, J. Reichenbach, C. Büchel and 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, and 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, and 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 and I. Suzuki (eds.), *In Untiring Pursuit of Better Alternatives: A Festschrift for Shigeo Tonoike*, 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: Split Topicalization and Quantifier Float in German*. 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. 2017. “The Syntax and Pragmatics of Dislocation: A Non-templatic Approach”. Paper presented at the *Annual Meeting of the Canadian Linguistics Association*, Ryerson University Toronto, May 27-29.

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

Ott, D. forthcoming. “Review of N. Richards, *Contiguity Theory*”. *Language*.

Ott, D. and R. Šimík. to appear. “What Drives Syntactic Computation? Alternatives to Formal Features”. *Glossa*.

Ouali, H. 2008. “On C-to-T ϕ-feature Transfer”. In R. D’Alessandro et al. (eds.), *Agreement Restrictions*, 159-180. Berlin: Mouton de Gruyter.

Pesetsky, D. and 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.

Ramchand, G. 2008. *Verb Meaning and the Lexicon. A First Phase Syntax*. 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 and 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. Developing the Framework*, 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. Handbook in Generative Syntax*, 281-337. Dordrecht: Kluwer.

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

Roberts, I. 2010. *Agreement and Head Movement. Clitics, Incorporation, and Defective Goals*. Cambridge, MA: MIT Press.

Sandler, W. and 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. and 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.

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

Trinh, T. 2011. Edges and Linearization. An Investigation Into the Pronunciation of Chains. PhD Dissertation, MIT.

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

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

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". In R. Freidin and H. Lasnik (eds.), *Syntax: Critical Concepts in Linguistics* vol. 5: 21-34. New York: Routledge.

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

De Vries, M. 2012. "Unconventional Mergers". In M. Uribe-Etxebarria and 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.

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: How Children Learn and Break Rules of Language*. Cambridge, MA: MIT Press.

Yang, C. 2017, S. Crain, R.C. Berwick, N. Chomsky, J. Bolhuis. In press. "The Growth of Language: Universal Grammar, Experience, and Principles of Computation". *Neuroscience and Biobehavioral Reviews*.

Zwart, J.W. to appear. "An Argument Against the Syntactic Nature of Verb Movement". In L.R. Bailey and M. Sheehan (eds.), *Order and Structure in Syntax I: Word Order and Syntactic Structure*, 29-48. Berlin: Language Science Press.