

tl;dr Developmental Minimalist Syntax

Michael Diercks, with Madeline Bossi
michael.diercks@pomona.edu
Pomona College and UC, Berkeley
Draft date: May 16, 2021

This brief note is a highly truncated summary of a longer in-development manuscript, which is [available on lingbuzz](#). We propose Derivational Minimalist Syntax (DMS), an interpretive principle claiming that the sequence of structure-building operations in the derivation of a sentence directly correlates with the sequence in which those structures are acquired by children (with Merge-based syntax being the primary mechanism for grammaticalizing language input). This correlation arises because mechanisms/operations of Universal Grammar are language acquisition devices; specifically, they are mechanisms for children to grammaticalize patterns they encounter in their input during acquisition. Thus, adult grammatical knowledge contains earlier stages of grammatical knowledge, and retains that knowledge, which is evidenced in the derivational sequence of a Minimalist syntactic analysis. On this approach, counter-cyclic operations in adult grammars are the result of grammatical patterns being acquired “earlier” or “later” than a strict application of bottom-up structure building would require. This brief note is intended to make the core concepts of the larger manuscript more accessible.

1 Building Phrase Structure in the Minimalist Program

For over two decades generative syntacticians operating in the Chomskyan tradition have been following the general framework laid out by [Chomsky \(1995, 2000a, 2001\)](#) that is known as the Minimalist Program, an iteration of the generative syntax tradition. By many measures, this has been an extraordinary success, both in terms of the range of empirical phenomenon that our theoretical mechanisms now capture, as well as the depth of detail that those same mechanisms allow us to describe and explain with precision within a given grammatical construction in a given language.

On the Minimalist approach, syntactic structures are built from bottom-up, with a verb (or more abstract root) merged with a complement to form a basic predication relation, and all additional syntactic structure is built on top of this in a cyclic operation where new material is merged with the root of the existing structure. In (1) we informally use the ‘plus’ symbol to annotate a Merge operation.

$$(1) \quad V + DP \rightarrow [_{VP} \ V \ DP] + v \rightarrow [_{vP} \ v \ [_{VP} \ V \ DP]] \rightarrow (\text{etc})$$

As (1) shows, a verb merges with a noun phrase (DP), resulting in a VP, and the resulting VP is then a candidate for merging with another phrase.

This creates a strict cycle of structure building: sequential applications of Merge building hierarchical syntactic structures step-by-step. Specifically, that standard implementation of this follows what is usually referred to as the “Extension Condition,” a requirement that applications of Merge extend the root of the structure. “A natural requirement for efficient computation is a ‘no-tampering condition’ (NTC): Merge of X and Y leaves the two SOs [syntactic objects] unchanged. If so, then Merge of X and Y can be taken to yield the set X, Y, the simplest possibility worth considering. Merge cannot break up X or Y, or add new features to them. Therefore Merge is invariably ‘to the edge’ ” (Chomsky, 2008, 138).

Merge does not alone account for all of syntactic phenomena (there are other posited components of UG, and other empirical aspects of morphosyntax), but it is where we direct our attention for the moment. Focusing on that core structure-building operation, this requires a strict cycle that is monotonic, with each new cycle extending the root of the structure without any changes to the existing structure at that point of the derivation.

As an operating framework of analysis, the Minimalist Program has been quite successful, facilitating a large amount of analytical and theoretical work in a broad range of languages. That said, what might be viewed as a bit of a black mark is the persistence of counter-cyclic analyses of empirical phenomena in the literature. By “counter-cyclic” here we refer to step in the derivation that does not proceed according to the strict cyclic Merge described above. On the way we are using the term here, this includes an operation that happens *later* than it ought to according to the strict cycle: a classic instance of this is Late Merger, where some head/phrase is merged into a structure in violation of the Extension condition, inserting itself into existing structure. We also include (in our definition of countercyclicity) operations that happen earlier than they ought to according to the strict cycle of narrow syntax, and/or without clear motivation at the point in which they appear to occur. These are often termed look-ahead problems, where something happens derivationally before the supposed trigger of the operation has been merged into the structure.

It is notable that despite the lack of large-scale theoretical incorporation of counter-cyclicity, we are quite familiar with *both* look-ahead problems and instances of Late Merger, in addition to other more nuanced instances of counter-cyclic analysis, such as upward-probing Agree (Baker 2008, Bjorkman and Zeijlstra 2019), delayed probing (Zeller 2015), and deletion of structure (Müller 2017, Pesetsky 2019), among others. Chomsky has recently spoken out against such countercyclic analyses, which make up a large proportion of the “foundational issues which are unsettled and I think are rather troublesome, and that bear directly on a number of very important issues in current work and I think raise questions about the legitimacy of problems and challenges we have faced” (Chomsky, 2019, 264). Sportiche (2019) expressed similar skepticism in his article titled, “Somber prospects for Late Merger,” arguing that the empirical benefits of the Late Merger analyses don’t outweigh the potential overgeneration that such an operation introduces into the theory. Any syntactician who has submitted a counter-cyclic analysis for review has encountered the obligatory push-back, which is (of course!) deserved: the Minimalist Program has no widely-accepted theory of counter-cyclic operations, and based on foundational assumptions

deducing the properties of language from interface conditions (Chomsky 2000a, 2001) there is not good reason that Late Merger ought to exist. Apparent counter-cyclicity is, from a standard Minimalist perspective, always an analytical problem that is perhaps only tolerated because no possible alternative (cyclic) derivation is apparent (to both the researcher and reviewers/editors) that might explain the patterns. Chomsky himself expressed it recently in this way:

Over the years, whenever some descriptive device has been introduced, and whatever it is (PSG, transformations, X-bar theory, parameters, phases, whatever it might be), almost always it tends to be used pretty extravagantly, well beyond the basis, of any solid foundation for the rule ... the advantages were that it led to a lot of discoveries, there were lots of insights about language that came out of it. They're not solutions, they're problems, and it's good to have problems, and led to explorations of new domains that hadn't been looked at. All of that's positive, and that's commonly true for the promiscuous use of devices that are invented. The negative aspect is that it doesn't lead us to the goal of trying to understand UG and the language faculty, and it's also misleading in that it tends to present problems, which are interesting problems, as if they were solutions, and they are not solutions: they are ways of stating a problem that we have to look at. (Chomsky, 2019, 270-272)

His overall perspective here seems undeniable: if a theoretical mechanism is made available, grammar analysts will make use of it, whether or not it is deservedly in our theory. That is the nature of research, and it has positive effects, as it often opens us up to exploration in new theoretical areas. But if those mechanisms are in fact ultimately not well-founded, what do we do with the resulting empirical puzzles?

The core question this monograph investigates can be phrased like this: why does counter-cyclicity exist? The glib answer, of course, is that the analysts couldn't think of a better idea: it is one of these promiscuously-used analytical mechanisms that actually doesn't belong in the theory, so apparent counter-cyclic analyses are simply a persistent empirical problems to be solved. On this approach, counter-cyclic processes are necessarily only apparently counter-cyclic, because (by assumption) Merge is the only structure-building operation.

But as we've mentioned, certain kinds of counter-cyclic analyses have managed to persist in the literature despite the relatively widely-accepted theoretical assumptions that counter-cyclicity ought not exist. How do we resolve this collective cognitive dissonance? We assume that counter-cyclicity is real as an empirical phenomenon (broadly-speaking), and take that to argue that inclusion of counter-cyclic mechanisms in our conceptualization of Universal Grammar. That said, we aren't arguing here for a specific technical account for countercyclic mechanisms (presumably, versions of the already-proposed mechanisms would work for the constructions they have been developed for). Rather, we propose an extra-grammatical motivation for counter-cyclic operations, correlating the derivation of sentences in adult competence to sequences of language acquisition.

This brief note is a highly truncated summary of an in-development manuscript, which is [available on lingbuzz](#). Cross-references to sections/chapters refer to the sections in the full manuscript. Our hope is to communicate the main ideas of the full manuscript without requiring readers to engage that whole manuscript. That said, the proposals here are expansive and those who are interested in evaluating the proposals are recommended to read the full

manuscript.

2 Developmental Minimalist Syntax

Developmental Minimalist Syntax (DMS) is a relatively simple proposal about a possible interpretation of the findings of the Minimalist Program. In short, DMS claims that adult knowledge of the syntax of a speaker's native language(s) is well-modeled in a bottom-up fashion precisely because bottom-up structure building mimics the ontological (i.e. organism-internal) timeline of grammaticalization of syntactic structures during child language acquisition. That is to say, the general model of the acquisition of syntax is that 1) structurally lower elements are acquired before structurally higher element and 2) later stages of syntactic knowledge incorporate previous stages of knowledge (i.e. as a general rule, later-acquired syntactic structures are added to existing knowledge, rather than replacing that existing knowledge). The result is that earlier stages of knowledge are retained in our grammatical representations as we mature, and as a result adult grammatical structures also encode the grammaticalization pathways that we traverse in reaching our adult grammar.

DMS is at its core an interpretive principle; it is a proposal that the syntactic analysis of adult syntactic structures has direct implications for our understanding of the ontological development of language. The step-by-step derivation of a sentence using Minimalist principles (e.g. Merge, Agree, phases)—which DMS adopts without alteration—mirrors the acquisition timeline for a child acquiring those linguistic structures:

(2) Developmental Minimalist Syntax (an interpretive principle)

The Minimalist derivation of adult language structures recapitulates the ontological development of those same syntactic structures.

DMS essentially proposes a correlation between empirical domains—that structural hierarchy, syntactic displacement, and the mechanisms which generate it in adult language (e.g. bottom-up structure building and syntactic movement via Copy + Internal Merge) are directly correlated with stages in language development.

More specifically, if (4) is on the right track, the theoretical constructs and derivational processes proposed to account for observed syntactic phenomena in effect also model the mechanisms and pathways of child language acquisition. For example, (2) leads to the claim that the early stage of the derivation of a sentence in adult grammar—in which a subject, verb, and object all exist in their base positions in *vP*—corresponds to an ontological stage wherein the components of *vP* are the full extent of a child's (grammaticalized) syntactic knowledge of their language. Higher structures are then systematically *added* on top of existing structures as a child gains more knowledge about the target language. In other words, the widely-accepted *vP*-internal subject hypothesis—in which all clausal subjects have an underlying base position within *vP* regardless of their surface position—is a remnant of an earlier stage in child language acquisition when the child represented the subject in that *vP*-internal position. Such remnant structures can be thought of as *ontological fossils* (i.e. fossils from an earlier state of that organisms language knowledge), a term (and a concept) inspired by work on the evolution of human syntax in [Progo-](#)

[vac \(2015\)](#) (discussed in §2.4 of Chapter 2). In this way, syntactic movement is always an instance of adding a new grammatical representation of the position of an element, but without demolishing the initial grammatical representation: adult knowledge encodes not only the final word order, but also the previous stages of knowledge.

Stepping back from the theory for a moment helps to put this interpretive principle in perspective. The fact that Minimalist syntactic models build structures bottom-up—starting first with verbs and objects—is quite unintuitive from certain perspectives. There is no obvious way that this bottom-up structure building relates to the psychological reality of language processing; perception of most of the world’s languages clearly cannot proceed in that fashion—or else garden path sentences could not exist—nor can language production proceed bottom-up: a speaker can certainly begin a sentence without knowing how they are going to end it. Non-specialists and new linguistics students often find these theoretical constructs quite unintuitive given the lack of obvious correlation to more surface-evident language patterns like perception and production.

Generative syntacticians generally do not assume that a model of language must have a psychological reality related to language processing; per the now-famous distinction proposed in [Chomsky \(1965\)](#), syntacticians are modeling speakers’ knowledge about the syntax of their language (i.e. ‘competence’) rather than how they speak or perceive sentences (i.e. ‘performance’). Given that Minimalist models of syntax work quite well to model a wide range of constructions in a large number of the world’s languages spanning a disparate range of language families, a crucial question arises: why does bottom-up structure building model grammatical knowledge so well as a process, a sequence of apparently ‘mental’ operations, when it clearly does not relate to language processing or production? DMS represents an attempt to answer precisely that question by claiming that there is in fact some type of psychological reality to bottom-up structure building, as it directly encodes the pathway by which children arrive at adult-like syntactic generalizations about the target language. In other words, bottom-up structure building models adult language so well because adult linguistic knowledge is essentially a crystallized history of the successive generalizations that children make when learning language.

DMS effectively brings Minimalist theories of Universal Grammar back to their historical roots: UG as a Language Acquisition Device. As we lay out in more detail below, this readily allows for many acquisition processes to proceed based on statistical learning, and to proceed in an item-based fashion, but at the point when grammaticalization occurs (i.e. the child generalizes beyond specific examples to a grammatical rule, structure, or generalization), that these generalizations take a very specific form. Specifically, the Minimalist proposals about UG, on this approach, are proposals about how children grammaticalize a pattern. We argue below that the canonical operation for forming a new grammatical generalization about the structure of a sentence is Merge: essentially, new knowledge of their language’s structure is incorporating by merging a new functional head atop the child’s existing grammatical structure of a sentence.

The claim that children acquire language in some kind of bottom-up syntactic fashion has always been appealing (and, as we will summarize below, boasts relatively strong empirical support), and hence there is a host of generative work that has proposed this (summarized below and discussed in depth in Chapter 2 of the full monograph). But there have been persistent problems with such claims, specifically, empirical domains where obvious predictions of such accounts are

not upheld. What we want to show, however, is that there are compelling correlations between (on the one hand) the empirical domains that have proven to be problems for a bottom-up theory of acquisition and (on the other hand) the empirical domains in adult language grammar that are problematic for a strictly cyclic derivation of syntactic structure.

For example, some presumably CP-level structures are acquired by children before they “ought to be” according to a bottom-up acquisition theory. Likewise, some lower-level structures seem to be acquired later than they “ought to be.” Rather than calling these weaknesses of this account, though, we claim that such divergences from the core structure building process in fact correlate with long-standing anomalies within Minimalist syntactic analyses: namely, countercyclic operations. We will show that “early” acquisition (i.e. earlier than predicted per a strict bottom-up acquisition process) correlates with look-ahead and with otherwise unmotivated movement processes (like most head movement of verbs). In contrast, “late” acquisition (i.e. later than predicted per DMS) correlates with constructions in adult language that require Late Merger operations (see, for example, [Lebeaux 1988](#); [Takahashi and Hulsey 2009](#)).

So our proposal in ways attempts to reinvigorate older ideas, but also claims that advances in the Minimalist Program can give those ideas new life; syntax is generally acquired bottom-up, but not exclusively so. On the flip side, we suggest that facts about acquisition can offer some explanatory basis to the persistence of counter-cyclic analyses in models of adult grammar. Moreover, we claim that the minimalist model of analysis uncovers this specific kind of grammatical knowledge (and acquisition phenomenon), rather than countercyclicality being a problem to be eliminated from the Minimalist Program.

2.1 Core Principles of DMS

Recall our interpretive principle from above:

(3) **Developmental Minimalist Syntax (an interpretive principle)**

The Minimalist derivation of adult language structures recapitulates the ontological development of those same syntactic structures.

This kind of approach leads to a specific conceptualization of Minimalist proposals about Universal Grammar (UG) as largely being proposals about the nature of the Grammaticalization Mechanisms that are employed on language. Put another way—in a slightly weaker claim—accurate Minimalist proposals about UG are in fact detailed articulations of the specific structures that result from grammaticalization. A number of conclusions fall out from this, which we outlined in this work as 9 principles of DMS.

(4) DMS Principle #1:

The theory of Universal Grammar (composed of at least Merge and Agree) is effectively a description of the nature of grammaticalization in language acquisition.

(5) DMS Principle #2:

In acquisition, new syntactic structures typically incorporate existing structures (with some principled exceptions).

- (6) DMS Principle #3:
Sequence of structure building in the Minimalist derivation of a sentence correlates with the timeline of acquisition.
- (7) DMS Principle #4:
Syntactic movement is reanalysis.

Essentially, the principles above articulate the proposal and consequences of Minimalist syntactic operations being articulations of the precise way in which grammatical representations are formed. As we summarized in Chapter 2 and as we briefly address in this tl;dr summary, there is a wide range of work showing that the empirical predictions of these principles are broadly attested: hierarchically lower structures are acquired before higher structures.

Many of the particular details, while important, are aside from the main claim here, which is to claim that as a result of this correlation, the entire derivation of adult grammatical structures recapitulates their acquisition pathways; adult grammatical knowledge of syntax encodes the pathways by which those structures were acquired. The principles above may not necessarily lead to the conclusions below, but on the strictest interpretation of the Minimalist program (and, finding agreement with the proposals of both usage-based grammarians and typologists), we find these conclusions:

- (8) DMS Principle #5:
Syntactic categories are emergent ([Wiltschko, 2014](#))
- (9) DMS Principle #6:
Parameters are emergent ([Roberts 2019](#), among others)

These are not new claims, even within the Minimalist framework. But they are the most naturally compatible approaches to categories and parameters for a DMS-style approach to Language.

To our knowledge, the complete suite of claims above have not been proposed in the configuration done here, or with the degree of predictive force that (2) carries; that said, for the most part the preceding principles/claims have direct and/or indirect precedents in the literature. If the canonical result of DMS as articulated in (2) is that syntax is acquired in a bottom-up fashion, counter-cyclic processes are the exceptions that prove the rule. It is well-established that nothing like rigid stepwise acquisition occurs among actual children learning actual natural languages, as we might envision on a simplistic application of Minimalist derivations to acquisition. Rather, acquisition is fluid and gradual, with some clear exceptions to ‘bottom-up’ acquisition: these exceptions are central to our claims, as discussed in the sections that follow (and in Chapters 3 and 4 in the manuscript). That said, it is nonetheless clear that (as a general rule) hierarchically lower structures reach adult-like states of grammatical knowledge before hierarchically higher ones.

Table 5.1: Examples of early pivot schemas / small clauses in Western European languages (Rakhlin and Progovic, 2020, 3)

English:	Eve climb. Go truck. See Adam. Put baby. Hurt doggie. Kittie go home.			
German:	Weint die Katze. cry.3SG the cat	Macht das Baby. make.3SG the baby	Schlafen mein Bruder. sleep.INF my brother	Mache ich auch do.3SG I too
French:	Ouvre la porte. Open.3SG the door	Monter Grégoire climb.INF Gregoire	Mangé (l)e chien eat.3SG the dog	Est tombé voiture have-3SG fallen car
Russian:	Ubasku simat'. shirt.ACC take.off.INF	Midet' idjot bear.NOM go.3SG	Osik kusiit. donkey.NOM eat.3SG	

2.2 Support for DMS

Empirical Support for DMS, brief summary

Section 3 of Chapter 2 outlines the empirical support in the literature for bottom-up structures building in acquisition as in adult grammars. First, there is a clear stage in acquisition where children's grammars consist of little other than a predicate and an argument: Tomasello (2003) calls these *pivot schemas*, Progovic (2015) and Rakhlin and Progovic (2020) simply call them small clauses. We argue (as do Rakhlin and Progovic 2020), that this is the equivalent of the head/complement relation between verbs and their internal arguments. Table 5.1 shows a collection of these early constructions across English, German, French, and Russian. Rakhlin and Progovic (2020) note that there is an absolutive-like quality at this stage, where there is a clear preference for objects and intransitive subjects as the single argument of the predicate, though there is often ambiguity as to the role of the single argument (subject/object), as (for example) in English subject-like postverbal arguments and object-like preverbal arguments both appear.

Rakhlin and Progovic (2020, 6) note that “[a]round 2 years of age, children undergo a cognitive shift when they begin to refer explicitly to causes of events, instead of reporting events as simply happening.” This is perhaps most evident in the causative errors that appear around this stage, where children over-generalize the transitive schema to produce errors as in (10).

- (10) Daddy go me around (2;8); I come it closer so it won't fall (2;3); Mommy can you stay this open (2;6); Drink me (3;1); I am going to fall this on her (2;9).
(Rakhlin and Progovic, 2020, 6)

While errors like (10) clearly show errors in lexical knowledge (i.e. *go*, *come* are not transitive verbs), they in fact are very strong evidence for the acquisition of transitivity and agentivity in general. The key for us, here (as for Rakhlin and Progovic 2020), is that these kinds of errors in fact reflect acquisition of the grammatical notion of agentivity.

There is a stage in acquisition where children are consistent in producing verbs with their arguments in systematic ways, but do not inflect those verbs appropriately for tense/aspect/etc (see Grinstead 2016 for an overview). This stage has variably been referred to as the root infinitive stage, optional infinitive stage, or root default stage, the latter term being the one we will adopt (following Vainikka and Young-Scholten 2011). So although children occasionally make use of finite verbs during the root default stage, this stage is characterized by the frequent use of

nonfinite forms and/or default forms in finite contexts. As (11) shows, children acquiring a range of genetically distinct languages exhibit characteristics of the root default stage (adopted from Legate and Yang 2007 and Kallestinova 2007):⁷

- | | | |
|------|--|---------|
| (11) | a. Papa have it. | English |
| | b. Thee drinken.
tea drink-INF | Dutch |
| | c. Dormir petit bébé.
sleep-INF little baby | French |
| | d. Mein Kakao hinstelln.
my cocoa out-INF | German |
| | e. Lashevel al ha-shulxan.
sit-INF on the-table | Hebrew |
| | f. Mama spat'.
mommy sleep-INF | Russian |

Recalling the predictions above, CP-level structures (across languages) are consistently structurally higher than TP and vP, and are predicted to be acquired later, sequentially-speaking. There does appear to be a consistent, cross-linguistic tendency for CP structures to be acquired later than TP structures. In general, TP-level structures are acquired in adult-like ways before CP-level structures (collectively) reach that point. Table 5.2 shows a summary of these kinds of sequences across a range of languages.

This is a highly truncated summary of the empirical evidence for bottom-up acquisition of syntax (generally). §3 of Chapter 2 gives a fuller summary, drawing on the literature reviews in Rakhlin and Progovac (2020), Vainikka and Young-Scholten (2011), and Bossi 2017 in addition to other findings in the literature. The conclusion is that that, in a general sense, acquisition of syntax does proceed ‘bottom-up’ as predicted by DMS and as argued by various other researchers. We don’t mean to imply that there aren’t complications: in fact, exceptions to this overall trend are central to the DMS proposal as articulated in (2), as summarized below, and as discussed at length in Chapters 3 and 4 of the full monograph.

Theoretical Precedent for DMS

There is an impressively broad range of theoretical precedent for DMS, to the point that very little that is proposed here is in fact new: to our knowledge, it has not been put together in this fashion, and it has not been explicitly linked with Minimalist derivations of adult grammars. But even some of the more novel claims about counter-cyclicity here have rather direct precedents as well. It does seem, however, that many of these ideas have either been forgotten, or ignored. We think DMS as articulated here has promise to bring syntactic and acquisition theorizing closer together again, but we want to be clear not to overstate the novelty of the ideas here: most have been floating around the literature in one form or another.

⁷There is a quite large range of research on this topic; ? provides a relatively recent overview.

Table 5.2: L1 Acquisition of IP and CP in 12 languages ([Vainikka and Young-Scholten, 2011](#), 77)

Language	IP-elements acquired earlier [before or around age 2]	CP-elements acquired later [after age 2]
English	tense	relative clauses
	auxiliary verbs	sentential complementation
Polish	tense/aspect	relative clauses
		complex sentences
Scandinavian	negation	relative pronoun
French	clitic pronouns	subordinate clauses
	tense	relative clauses
	negation	
Hebrew	tense	relative clauses
	negation	causal and temporal linking
	agreement	of clauses
Turkish	verb inflections	conjunctions
Georgian	agreement inflections	two-clause constructions
Mandarin	modals	topicalisation
	aspect marking	discourse particles
Japanese	verbal inflection	relative clauses
Kaluli	tense	discourse particles
Sesotho	tense/aspect	relative clauses
		topicalisation
K'iche 'Maya	aspect	yes/no question particle
	negation	

Theories of acquisition of syntax (especially those similar to DMS) were discussed and heavily debated in the 1990s (e.g. Rizzi's 1994 Truncation Hypothesis, Radford's 1990 Small Clause Hypothesis, and the work from Harald Clahsen and colleagues (e.g. Clahsen et al. 1993/1994; Clahsen 1990/1991; Clahsen et al. 1994, among others). Lebeaux (1988, 2000) in fact sketches a theory quite similar to what we discuss here; while his initial proposal of late merger operations has certainly taken hold among syntacticians (albeit somewhat controversially), it's fairly clear that his ideas about connections with language acquisition have not. The literature does contain some cursory comments about some aspects of his proposals, they haven't led to a holistic conception of syntax-acquisition links in the way that we propose here. In part, this may be due to Lebeaux's (2000) insistence that some of his proposals were in fact radically different than Chomsky's Minimalist proposals. At least insofar as the Minimalist Program has evolved since that time, in our minds they are not nearly as incompatible as Lebeaux (2000) suggests (see discussion in §5 of Chapter 2).

All of the preceding work notes the clear 'bottom-up' trend of acquisition of syntax. To our knowledge, though, these core ideas have not been seriously considered in recent decades (note the absence of more recent theories of acquisition of syntax in two recent generative textbooks on language acquisition: Guasti 2016 and *ud Deen and Becker 2020*). A major exception is the work of Anne Vainikka and Martha Young-Scholten on their theory *Organic Grammar* (focused mainly on second language acquisition, but also explicitly discussing first language acquisition) (e.g. Vainikka and Young-Scholten 2007, 2011, 2013). We discuss their claims at length in Chapter 2. An additional exception that shares many similarities with *Organic Grammar* comes from the suggestions about language acquisition advanced in Progovac (2015) as part of a discussion of language evolution, and the follow-up work on acquisition itself from Rakhlin and Progovac (2017, 2020). Progovac (2015) and Rakhlin and Progovac (2020) are perhaps the two strongest sets of claims that Minimalist derivations in adult grammars correlate to extra-grammatical properties of cognition, and their work is referenced liberally throughout the full monograph. Our main contribution is to expand the level of detail of those initial claims, and especially engaging potentially problematic aspects of acquisition (where 'bottom-up' acquisition is not clearly the case, or is clearly *not* the case) and arguing that these properties of acquisition are *also* encoded in adult grammars. These are mainly found in our discussions of counter-cyclicity (Chapter 3) and phases (Chapter 4).

All of this is to say that there is a good theoretical foundation for these claims, and the resulting conclusions are of significance both for Minimalist syntacticians and also for acquisitionists.

2.3 Master Tree Inventory

This summary to this point has overlooked a central question: what are the specific grammatical representations being acquired in this bottom-up fashion? It cannot be each and every possible sentence in a language, as that would simply return us to a behaviorist model of language, and completely overlooks the productive aspect of language knowledge (capable of producing novel sentences). In Construction Grammar accounts, this role is taken on by abstract syntactic constructions (see Tomasello 2003 for detailed discussion); Vainikka and Young-Scholten (2011)

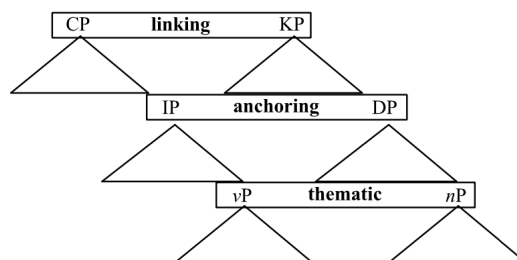


Figure 5.1: Universal Spine (Ritter and Wiltschko, 2014, (5)) Screenshot for draft, will be redrawn

propose the concept of a ‘master tree’ to serve this purpose in their Organic Grammar model. On this approach, what is being acquired by a child is the maximal clause structure that is possible in the language: knowledge of individual sentences in production and perception are derived from this master tree. As we discuss in §5 of Chapter 2, we adopt this approach, though with more detail relating ideas of master trees to standard Minimalist syntax.

It is well established that there are quite general structural hierarchies across all languages. This is evident by the apparently universal structural hierarchy of adverbials across languages, where adverbials of certain semantics appear at predictable structural heights (cf. Ernst 2014 and Cinque (1999)). Cinque (1999) models this as a highly detailed hierarchy of functional projections that is universal because it is a part of UG, with particular adverbs introduced in the specifier positions of matching particular functional heads.

(12) The universal hierarchy of clausal functional projections (Cinque, 1999, 106)

[*frankly* Mood_{speech act} [*fortunately* Mood_{evaluative} [*allegedly* Mood_{evidential} [*probably* Mod_{epistemic} [*once* T(Past) [*then* T(Future) [*perhaps* Mood_{irrealis} [*necessarily* Mod_{necessity} [*possibly* Mod_{necessity} [*possibly* Mod_{possibility} [*usually* Asp_{habitual} [*again* Asp_{repetitive(I)} [*often* Asp_{frequentative(I)} [*intentionally* Mod_{volitional} [*quickly* Asp_{celerative(I)} [*already* T(anterior) [*no longer* Asp_{terminative} [*still* Asp_{continuative} [*always* Asp_{perfect(?)} [*just* Asp_{retrospective} [*soon* Asp_{proximative} [*briefly* Asp_{durative} [*characteristically(?)* Asp_{generic/progressive} [*almost* Asp_{prospective} [*completely* Asp_{SgCompletive(I)} [*tutto* Asp_{PlCompletive} [*well* Voice [*fast/early* Asp_{celerative(II)} [*again* Asp_{repetitive(II)} [*often* Asp_{frequentative(II)} [*completely* Asp_{SgCompletive(II)}

Wiltschko (2014), Ritter and Wiltschko (2014), and Ramchand and Svenonius (2014) argue that accounts like Cinque’s are both theoretically undesirable (too much content is assumed to be innate) and empirically inaccurate, as it predicts more consistency than we actually find across languages. In particular, Wiltschko (2014) and Ritter and Wiltschko (2014) argue that some languages lack TP, but nonetheless have other kinds of inflectional morphology at similar clause levels that perform similar abstract syntactic functions. Therefore, particular grammatical categories within particular languages are emergent properties of language, and not universal. Ritter and Wiltschko (2014) instead propose that adverb hierarchies (in addition to other consistencies in clause structure across languages) derive from a more abstract, simpler ‘Universal Spine:’ an abstract set of domains that wherein units of language are associated with specific grammatical properties, or functions (using those descriptors in non-technical senses). Their structure is shown in Figure 5.1.

Wiltschko (2014) claims that the Universal Spine can explain longstanding correlations between nominal and clausal domains (Abney 1987, among many others). Because there is one Universal Spine (but grammatical categories are themselves emergent), lexical categories (namely, nouns and verbs) will therefore share similarities in their extended projections precisely because the same abstract universal structure-building mechanisms underlie both: the universal spine and the categorization mechanism by which particular grammatical categories are formed in particular languages. As discussed by Wiltschko (2014), the *thematic domain* is where the event is introduced (i.e. predicate + participants). The *anchoring domain* anchors the event to the utterance, and discourse roles are introduced in the *linking domain*, linking the existing structure to the discourse.

In this way, we suggest that instead of a single ‘master tree,’ a more appropriate model will include a *master tree inventory*, the collection of all master trees for all lexical categories in a language. Beyond this, we likewise expect that there are also master trees for specific constructions of general application within languages (e.g. passives), though we don’t discuss this in much depth in the manuscript. This is the direct parallel of the ‘constructicon’ (parallel to the lexicon) proposed by Construction Grammarians (Goldberg, 2006; Jurafsky, 1996). The account proposed under DMS differs starkly in the underlying nature of constructions, of course: we follow the Minimalist Program in proposing a particular structure of constructions, rather than assuming that any kind of string can form a construction. Presumably, online processing accesses and references this grammatical knowledge. But notably, the UG-mechanisms proposed under the Minimalist Program are not involved in online processing on this account, but rather are involved in the grammaticalization of the hierarchical knowledge that is referenced and used in language processing. Minimalist theorizing has always had to assume distinct ‘workspaces’ (Chomsky, 2001), as there are many points in the derivation of a full sentence that fully-built structures must be merged into the structure of a clause (for example, DP subjects merge into Spec, vP having already been built, as is the case for AdvPs, adjunct clauses, etc). We already know, therefore, that our model has to accommodate structure-building in distinct ‘workspaces’ which are then combined into a specific structure of a specific sentence. On the DMS approach, knowledge of possible structures for the extended projection of different lexical projections is stored in ‘master trees,’ which are conceived of as separate representations in the master tree inventory, and that knowledge is referenced in the construction of individual structures in individual sentences. Those separate representations relate, in some ways, to the separate workspaces wherein sub-structures are built in standard minimalist theory: for example, separate ‘workspaces’ are needed for DPs to be built before they are merged into a sentential structure. On our account, these separate ‘workspaces’ correspond to distinct master trees for sentential structure (vP/TP/CP) and nominal structures (DPs). In some ways, then, the master tree inventory naturally fits with established Minimalist assumptions.

3 Deriving Countercyclicity

While DMS as articulated in (2) predicts a generalized bottom-up pathway of acquisition of syntax, there are additional syntax-acquisition correspondences that are predicted by (2). Centrally, we are not simply claiming a structural-directional correlation (timeline of acquisition to hier-

archical structure). Rather, we are proposing that a Minimalist derivation of a sentence is itself what correlates to acquisition. Of course, in most instances, this is a straightforward ‘bottom-up’ derivation (at least, per standard Minimalist assumptions). But our claim in this work is that counter-cyclic analyses are in fact real parts of syntax (both empirically, and theoretically) and therefore a Minimalist derivation is not, in fact, as strictly “bottom-up” as one might think at first. The exceptional instances are what can be referred to (broadly) as counter-cyclic operations.

First, we consider Late Merger operations. Late Merger is shown schematically in (13), where a structure like (13a) has been built, and only after the structure has been built does merger of LP occur, but instead of extending the root (XP), instead LP merges lower in the existing structure, resulting in the structure in (13b).

- (13) a. $[_{XP} X [_{YP} Y [_{ZP} Z]]] + LP \rightarrow$
 b. $[_{XP} X [_{YP} Y [\boxed{LP} L [_{ZP} Z]]]]$

Such operations are generally thought to be illicit, violating the Extension Condition by tampering with the root structure rather than extending it. Nonetheless, analyses like these do in fact appear to be well-motivated for certain empirical phenomena, and analyses dependent on Late Merger can be found again and again in the literature (examples are discussed in §2.1 of Chapter 1 and in Chapter 3). The most prominent instances are cases of anti-reconstruction, but given the complexity of the data patterns, we simply refer the reader to the full manuscript. We likewise discuss Zulu object marking as another instance of counter-cyclicity, but likewise we do not assume broad familiarity with those patterns, and so we refer the interested reader to our full manuscript (see §3 of Chapter 3). So for the purposes of this tl;dr summary, we will leave the structural description of late merger operations at the schematic level.

Another example that we consider in the general family of counter-cyclic operations is what is often termed ‘look-ahead,’ where an operation that is presumably contingent upon structures that are merged later in a derivation begins to apply before those later structures are merged. A prominent example is that *wh*-movement raises through lower phase edges before the target position of movement has been merged into the structure (matrix CP).

- (14) $[_{CP[WH]} \text{What}_k \text{ do } [_{TP} \text{you do think } [_{CP} \boxed{\text{what}_k} [_{CP} \text{that Alex ate } \text{what}_k]]]?$

There is a broad range of empirical evidence arguing for intermediate stages of movement of a *wh*-phrase, here shown at the edge of the embedded CP phase. The look-ahead problem arises in that (per the requirements of cyclic spell-out, i.e. derivation by phase) the *wh*-phrase moves to the edge of the embedded CP at the point that embedded C is merged; at this point in the derivation, however, the matrix C is not yet part of the structure, and there’s no clear immediate motivation for *what* to move to the edge of the embedded CP (the matrix CP is marked as [+WH], the embedded CP is most plausibly [-WH]). This requires the computational system to “look ahead” in the derivation, foreseeing the eventual need to raise to matrix CP, and ensuring the exit from the lower CP to avoid being trapped in the spelled-out phase.

In Chapter 3 we consider three case studies of counter-cyclicity, arguing that they have natural explanations within DMS. DMS Principle #7 (15) articulates this claim specifically, which adds nothing new from (2), but makes explicit the stance about counter-cyclicity. And (15) in turn entails (16).

- (15) DMS Principle #7:
Counter-cyclic phenomena in adult language grammars correlate to counter-cyclic acquisition processes.
- (16) DMS Principle #8:
Counter-cyclic syntactic operations exist.

(15) summarizes our claim that there are instances where adult-like acquisition of a grammatical element is not available to a child. This can be for various reasons. In instances of so-called “early” acquisition of some element, this will correlate with look-ahead. Our case study of these kinds of instances is *wh*-movement. The counter-cyclic puzzle is that *wh*-phrases (in adult grammar) appear to move before they should, i.e. before the final landing position has been added to the structure. This correlates to a prediction, per DMS, that children ought to acquire *wh*-movement “before they should,” i.e. before they have acquired the correlating CP structures that are the target of *wh*-movement. We show in §2 of Chapter 3 that this is the case. In this instance, then, counter-cyclicity arises because the full adult structure has not developed yet, but their input (PLD, primary linguistic data) offers a large supply of overt evidence for *wh*-words being sentence-initial. Therefore, children easily conclude that *wh*-words are sentence-initial at stages earlier than the final grammaticalization of CP.

Likewise, we propose that there are instances where children acquire structures “late,” meaning, a pattern is acquired that is located a specific structural height in a tree, but only after additional structures beyond that point have been acquired. We claim that this does not eliminate the possibility of merging at that position, but instead that “late merger” operations of the sort discussed in §2 of Chapter 1 and §2.1 of Chapter 3 are in fact available grammaticalization operations, but they have predictable grammatical effects in adult grammars, as discussed in the Late Merger literature (e.g. Takahashi and Hulsey 2009).

But why would Late Merger be necessary in the first place? Why could children not simply acquire the relevant structures on the strict cycle of Merge alone? This is where we think the link between adult structures and child acquisition may in fact be quite informative in broader research on cognition. In all of the areas where we have seen plausible syntactic proposals, the late-merged element is a structure of relevant complexity that is either plausibly (or obviously) not grammaticalized in acquisition until after subsequent, more complex clause structures are acquired. We outline the empirical arguments for this in Chapter 3, but in short, we have seen proposals for late mergers of DP-content (DPs exist early in acquisition, plausible, but fully complex DPs are not acquired until late), adjuncts (modifiers/ornaments of existing clause structures), adjunct clauses/relative clauses (requiring acquisition of full clause structures and constructions bearing on discourse/pragmatic structure, e.g. topic/focus): all of these are plausibly XPs that are not grammaticizable as Merged into the structure at the point on acquisition where they occur, because they contain a complexity of structure or complexity of semantics that the child has not yet acquired sufficient background to acquire, or which is cognitively inaccessible to a child during earlier periods of development.

In short, counter-cyclicity emerges when a child is not ready to grammaticalize a pattern that they are encountering on the usual cycle of strict structure-building via Merge. If they can access the data patterns themselves (e.g. *wh*-questions) but they do not yet have the grammatical

knowledge to fully incorporate the data pattern they observe, a look-ahead problem is created: aptly named, because the child can in fact “look ahead” to what the final structure looks like. In the case of *wh*-movement, this simply means stages of knowledge that place *wh*-phrases at the left edge of their existing (incomplete) grammatical structures. This is because DMS is not the blind computational system that the Minimalist itself is: under DMS, children are in fact “looking ahead,” as the UG-mechanisms are being employed to grammaticalize the patterns they are observing.

Conversely, a child may be unready to grammaticalize a pattern not because they don’t have the requisite grammar knowledge built to incorporate the pattern, but because they instead don’t yet have the cognitive ability to analyze a pattern to be grammaticalized in the first place. We will suggest in §3 of Chapter 3 that this is what is at work in instances like Zulu conjoint/disjoint and object marking constructions.

4 Phases as Acquisition Workspaces

The final principle that we propose for DMS, continuing on the list from above, has to do with phases.

- (17) DMS Principle #9:
Phases are acquisition workspaces.

The idea here is that children are working on specific grammaticalization tasks in a particular sequence: we propose that the sequence is precisely the functional domains proposed by [Ritter and Wiltschko 2014](#), mentioned above in Figure 5.1. First to be grammaticalized by a child are the thematic properties of an event (i.e. predicates and arguments: *vP*), then anchoring the event in time/space (i.e. inflection, per [Ritter and Wiltschko 2014](#): *Asp/TP/Infl*), and then linking the resulting situation/proposition (depending on the particular definition of those terms) to the discourse context (i.e. the *CP* domain). The properties of ‘phases’ in adult grammars are ontological fossils, the grammatical result of acquisition proceeding in this fashion. A phase in this sense is a domain that contains many heads, and which may share various properties (due to the underlying acquisition task a child is undertaking during a given phase). Phases are not, in fact, linked with particular functional heads. This approach has potential not just to derive some puzzling properties of phases in adult grammars (variable sizes of phases, simultaneity of operations within phases) but it also has promise to provide some clarity to the enduring metaphor that phases promote “computational efficiency,” despite our model being one of competence, not performance. We have suggested that many of these metaphors are in fact fairly accurate, but the efficiency that our model is achieving is in fact acquisition-related computation, and not online processing in adults.

These ideas are discussed at length in Chapter 4. Given the amount of theoretical background covered on phases in service of provided clear explanations of our proposals, we simply refer the reader to our full manuscript for details.

5 Potential Extensions

Despite the length of our discussion in the full monograph, we can't legitimately say to have provided a full empirical argument for DMS: the predictions are simply too expansive. We hope, however, that between the case studies that we provide on counter-cyclicity, the broad empirical support from the acquisition literature, and the theoretical precedents, that the reader finds these claims are at least on solid footing. We do think it is worth pointing out possible additional areas of investigation that may align well with the DMS approach.

The syntax-phonology intersections implicated by DMS are fairly far-reaching and non-standard. As stated succinctly by Richards (2017, 23), "The current consensus about this relationship in Minimalist circles, as [we] understand it, is that a phonological derivation begins once the syntactic derivation of a spellout domain is completed. The details of the phonological derivation are not often a focus of interest for syntacticians ... it is generally assumed that the derivation begins with a syntactic tree and performs a series of operations to convert that tree into a representation that can be used by the phonological interface." DMS, however, suggests that something entirely different is possible in the syntax-phonology interface. Rather than phonology necessarily coming *late* in the logical sequence of operations, by interpreting the UG mechanisms specifically as a Grammar Acquisition Device, phonological forms of morphemes/words/sentences are accessible (to some degree) by children, by necessity: phonological forms of one sort or another make up the input they are receiving. This predicts, however, that relatively robust interactions of (early-accessible) phonological material with aspects of syntactic structure are possible.

Richards (2016, 2017) argues that this is exactly what happens: there are syntax-phonology interactions in the course of a syntactic derivation, quite to the contrary of standard assumptions. Richards (2016, 2017) provides an expansive theory (with predictive force) arguing that the presence/absence of syntactic movements (e.g. 'EPP' movement to Spec,TP and wh-movement to Spec,CP) is explained by the prosodic properties of languages: movement occurs to satisfy prosodic constraints. Richards (2017) removes some of the more stipulative aspects of Richards (2016), suggesting that prosodic structures are constructed alongside syntactic ones, i.e. along with the cyclic structure building of Merge. "In a sense, the proposal of Contiguity Theory is a very modest one; this kind of phonological operation, which applies to a syntactic tree, can apply, not after the syntactic derivation is completed, but while it is still under way. In fact, if there are any operations that the phonological derivation can perform before the syntactic derivation is complete, perhaps it makes sense for them to be performed as early as possible, if the goal is for the derivation to produce linguistic objects as quickly and efficiently as it can" (Richards, 2017, 24).

As noted by Ott (2017) in his review of Richards 2016,

perhaps the most interesting challenge posed by [Richards'] proposals concerns the place of morphophonology in the overall organization of the grammar. In his model, at least some syntactic operations apply in the service of constructing prosodic structure in tandem with the syntactic derivation. Phonological information such as the presence of metrical boundaries is directly accessed by the syntactic computation; the phonology does not merely impose output conditions on completed derivations" (Ott, 2017, 722). Yet, at the same time, "if [Richards] is right, the phonology is more

than an ancillary mapping relating the internal computational system to articulation and perception: it is an ‘active player’ in syntactic computation (*contra* long-standing claims by Chomsky). (Ott, 2017, 723)

Of course, this kind of outcome is precisely what is predicted by DMS: we expect phonological properties of language (at least, those acquired by children sufficiently early) to be not only be eligible to participate in syntactic operations, but perhaps also *likely* to participate/shape syntactic acquisition, since the syntactic observations children are making would be in the context of the phonology they have acquired (at each respective stage of acquisition).

We discuss Richards’ claims in some more depth in Chapter 5, as well as some initial evidence from the acquisition literature suggesting that the specific sequences of prosody-syntax interaction addressed by Richards are what appears to occur in acquisition as well. These are only initial discussions, however, but the initial evidence is highly promising (from our perspective).

Essentially, the DMS approach provides a litmus test for evaluating the legitimacy of counter-cyclic analyses. We can interpret “counter-cyclic” quite broadly, even outside the Merge-based cycle of structure building, to anything that doesn’t straightforwardly fit the standard assumptions of “early” vs. “late” in standard generative models. If DMS is on the right track, “early” grammatical processes in our models should correlate with “early” acquisition of those grammatical patterns. What we have seen so far suggests this is true, but there is a lot to be investigated.

6 Conclusions

This tl;dr summary is a highly truncated summary of a much longer work (which is [available on lingbuzz](#)). Our hope in writing this summary was to communicate the main ideas of the full manuscript, to make them accessible to a broader audience and to help people evaluate their interest in the full work. Essentially, this is intended as an extended abstract. As such, we would hope that any critics (either about the proposal, or of work we don’t engage here) would read the full manuscript before concluding we have missed something. We surely have missed many things, but establishing the particular details about which things we missed requires looking over the full manuscript first.

We believe these proposals to be on excellent footing, both empirically and theoretically; this is evident by the broad range of intersecting and overlapping ideas from literature on adult syntax and child language acquisition, across major frameworks as well (the Minimalist Program, generative approaches to language acquisition, and usage-based approaches to both child and adult grammars, i.e. Construction Grammar). Our original goal was to seek out extra-grammatical correlations of (relatively stable) Minimalist theoretical postulates: if this approach is in any way on the right track for explaining the research questions it investigates, it ought to intersect with research in other areas of language, if not other areas of cognition too. We have been pleasantly surprised to find so many convergent findings. And in the course of this research, we believe we have also found good reason to accept (and incorporate) counter-cyclic processes into adult grammatical analyses as is relevant, with the associated litmus test that counter-cyclic analyses of structures make specific predictions for acquisition timelines.

In Chapter 1 of the full manuscript we discuss some of the critiques of counter-cyclic analyses: quite recently both Sportiche (2019) and Chomsky (2019) offered fairly direct critiques of Late Merger (and other counter-cyclic operations). The critiques against Late Merger leveled by Chomsky (2019) and Sportiche (2019) are entirely reasonable, based on Chomsky’s Strong Minimalist Thesis (SMT): allowing counter-cyclic operations *does* expand the the generative capacity of UG, and the model resulting from DMS is not the simplest possible model of syntax as could be deduced based on the interface conditions of language (i.e. the need for syntax to interface with both sensory-motor articulation and interpretation). On this account, however, we recall Chomsky’s (2001) comments on this issue:

The strongest minimalist thesis SMT would hold that language is an optimal solution to such conditions. The SMT, or a weaker version, becomes an empirical thesis insofar as we are able to determine interface conditions and to clarify notions of “good design.” While the SMT cannot be seriously entertained, there is by now reason to believe that in nontrivial respects some such thesis holds, a surprising conclusion insofar as it is true, with broad implications for the study of language, and well beyond.

Tenable or not, the SMT sets an appropriate standard for true explanation: anything that falls short is to that extent descriptive, introducing mechanisms that would not be found in a “more perfect” system satisfying only legibility conditions. If empirical evidence requires mechanisms that are “imperfections,” they call for some independent account: perhaps path-dependent evolutionary history, properties of the brain, or some other source. It is worthwhile to keep this standard of explanation in mind whether or not some version of a minimalist thesis turns out to be valid. (Chomsky, 2001, 2)

It’s important to note that the SMT is a theoretical heuristic, a philosophical stance, and not (as Chomsky points out) a necessary truth. And we agree: given the depth of achievements of Minimalist syntax in providing a framework to discover, describe, and analyze grammatical constructions across the world’s languages, there is good reason to think that the framework developed (driven by the SMT) is a reasonable result. But what we claim in this work is that there is in fact a strong and reasonable independent account (as Chomsky insists upon) of the diversions from the SMT-inspired UG mechanisms of Merge and Agree (and perhaps phases) alone. These may well be “imperfections” from the perspective of an optimal mapping between the PF/SM and LF/CI interfaces, but by now it should be more than clear (on moral, biological, social, and cognitive bases) that humans are composed of little that is perfect. So we find it unsurprising that the SMT as articulated as a theoretical heuristic may not hold in its strictest sense in actual human cognition.

Now, we are not equipped to say whether these “imperfections” arose from the evolution of language or instead are simply a necessity of the maturational processes involved in cognitive development (we’re inclined to say the latter, though without any real evidence). But wherever they are found, the strong developmental correlates of countercyclic processes in adult grammars are sufficient, in our minds, to consider the viability of DMS on a much broader scale.

It’s worth pointing out, despite the expansive implications of these claims, that the empirical coverage of DMS is severely limited as well, leaving most of Language untouched. From one perspective, the claims here are so sweeping that we can’t even defend them thoroughly in

an entire monograph, we can only provide case studies arguing that they are reasonable. From another perspective, we have simply made a claim correlating a few key outcomes of Minimalist syntactic research to some key properties of language acquisition. Apart from all the subfields of linguistics that we have nothing to say about, even within syntax we have completely ignored central areas of generative syntactic research (e.g. binding, ellipsis, agreement, case: to name just a few extraordinarily large areas of research.) If it turns out that DMS is on the right track, it will surely offer some important perspectives on these areas of research. But we want to be completely clear that even if DMS turns out to be the correct way to relate adult grammars to syntactic acquisition, this would still only be a small piece of understanding the never-ending puzzle of human language.

References

- Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral Dissertation, Massachusetts Institute of Technology.
- Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Bjorkman, Bronwyn M., and Hedde Zeijlstra. 2019. Checking up on (ϕ -)Agree. *Linguistic Inquiry* 50:527–569.
- Bossi, Madeline. 2017. On the psychological reality of bottom-up structure building. Unpublished undergraduate thesis, Pomona College.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: M.I.T. Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. Cambridge: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in linguistics*, ed. Michael Kenstowicz, 1–52. Cambridge: MIT Press.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory*, ed. Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2019. Some puzzling foundational issues: The Reading program. *Catalan Journal of Linguistics* Special Issue:263–285.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford University Press.
- Clahsen, Harald. 1990/1991. Constraints on parameter setting: A grammatical analysis of some acquisition stages in German child language. *Language Acquisition* 1.
- Clahsen, Harald, Sonja Eisenbeiss, and Anne Vainikka. 1994. The seeds of structure: a syntactic analysis of the acquisition of case marking. In *Language acquisition studies in generative grammar*, ed. Teun Hoekstra and Bonnie D. Schwartz, 85–118. Amsterdam: John Benjamins Publishing Company.
- Clahsen, Harald, Martina Penke, and Teresa Parodi. 1993/1994. Functional Categories in Early Child German. *Language Acquisition* 3:395–429.
- ud Deen, Kamil, and Misha Becker. 2020. *Language acquisition and development: A generative introduction*. Cambridge, MA: MIT Press.

- Ernst, Thomas. 2014. The syntax of adverbs. In *The Routledge Handbook of Syntax*, ed. Andrew Carnie, Yosuke Sato, and Daniel Siddiqi, 108–130. Routledge.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford University Press.
- Grinstead, John. 2016. Root infinitives in child language and the structure of the clause. In *The Oxford handbook of developmental linguistics*, ed. Jeffrey Lidz, William Snyder, and Joe Pater, 341–366. New York, NY: Oxford University Press, 1 edition.
- Guasti, Maria Teresa. 2016. *Language acquisition: The growth of grammar*. Cambridge, MA: The MIT Press, 2 edition.
- Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20:137–94.
- Kallestinova, Elena. 2007. Three stages of root infinitive production in early child Russian. *First Language* 27:99–131.
- Lebeaux, David. 1988. Language acquisition and the form of the grammar. Doctoral Dissertation, University of Massachusetts, Amherst.
- Lebeaux, David. 2000. *Language acquisition and the form of the grammar*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Legate, Julie Anne, and Charles Yang. 2007. Morphosyntactic learning and the development of tense. *Language Acquisition* 14:315–344.
- Müller, Gereon. 2017. Structure removal: An argument for feature-driven Merge. *Glossa: a journal of general linguistics* 2:1–35. DOI: <https://doi.org/10.5334/gjgl.193>.
- Ott, Dennis. 2017. Contiguity theory by Norvin Richards (review). *Language* 93:720–723.
- Pesetsky, David. 2019. Exfoliation: towards a derivational theory of clause size. Ms, MIT. Available at <https://ling.auf.net/lingbuzz/004440>.
- Progovac, Ljiljana. 2015. *Evolutionary syntax*. Oxford University Press.
- Radford, Andrew. 1990. *Syntactic theory and the acquisition of English syntax*. Cambridge, Massachusetts: Basil Blackwell Ltd.
- Rakhlin, Natalia, and Ljiljana Progovac. 2017. Acquisition of clausal structure: the ‘Weakest Continuity’ view. In *Proceedings of the 19th Annual International Conference of the Japanese Society for Language Sciences (JLS)*, 54–57. Kyoto: Kyoto Women’s University.
- Rakhlin, Natalia, and Ljiljana Progovac. 2020. Hierarchical clause structure as a tool for cognitive advances in early childhood. *Language Sciences* available online.
- Ramchand, Gillian, and Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46:152–174.
- Richards, Norvin. 2016. *Contiguity theory*. Cambridge: MIT Press.
- Richards, Norvin. 2017. Deriving contiguity. Ms, MIT. Available at <https://ling.auf.net/lingbuzz/003289>.
- Ritter, Elizabeth, and Martina Wiltschko. 2014. The composition of INFL: An exploration of

- tense, tenseless languages, and tenseless constructions. *Natural Language & Linguistic Theory* 32:1331–1386.
- Rizzi, Luigi. 1994. Early null subjects and root null subjects. In *Language acquisition studies in generative grammar*, ed. Teun Hoekstra and Bonnie D. Schwartz, 151–176. Amsterdam: John Benjamins Publishing Company.
- Roberts, Ian. 2019. *Parameter hierarchies and Universal Grammar*. Oxford: Oxford University Press.
- Sportiche, Dominique. 2019. Somber prospects for Late Merger. *Linguistic Inquiry* 50:416–424.
- Takahashi, Shoichi, and Sarah Hulseley. 2009. Wholesale Late Merger: Beyond the A/A' Distinction. *Linguistic Inquiry* 40:387–426.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, Massachusetts and London, England: Harvard University Press.
- Vainikka, Anne, and Martha Young-Scholten. 2007. Minimalism vs. Organic Syntax. In *Phrasal and clausal architecture: Syntactic derivation and interpretation. in honor of Joseph E. Emonds*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, volume 101 of *Linguistics Today*. Amsterdam/Philadelphia: John Benjamins.
- Vainikka, Anne, and Martha Young-Scholten. 2011. *The acquisition of German: Introducing Organic Grammar*, volume 44 of *Studies on Language Acquisition [SOLA]*. Mouton De Gruyter.
- Vainikka, Anne, and Martha Young-Scholten. 2013. Stagelike development and Organic Grammar. In *The Cambridge Handbook of Second Language Acquisition*, ed. Julia Herschensohn and Martha Young-Scholten. Cambridge/New York: Cambridge University Press.
- Wiltschko, Martina. 2014. *The universal structure of categories: Towards a formal typology*. Cambridge University Press.
- Zeller, Jochen. 2015. Argument prominence and agreement: Explaining an unexpected object asymmetry in Zulu. *Lingua* 156:17–39.