

THEORETICAL ESSAY

# PROBLEMS WITH **VARIABLE** **PROPERTIES** IN SYNTAX

David W. LIGHTFOOT  

Department of Linguistics – Georgetown University

## ABSTRACT

Like those birds born to chirp, humans are born to parse; children are predisposed to assign linguistic structures to the amorphous externalization of the thoughts that we encounter. This yields a view of variable properties quite different from one based on parameters defined at Universal Grammar (UG). Our approach to language acquisition makes two contributions to Minimalist thinking. First, in accordance with general Minimalist goals, we minimize the pre-wired components of internal languages, dispensing with three separate, central entities: parameters, an evaluation metric for rating the generative capacity of grammars, and any independent parsing mechanism. Instead, children use their internal grammar to parse the ambient external language they experience. UG is “open,” consistent with what children learn through parsing. Second, our understanding of language acquisition yields a new view of variable properties, properties that occur only in certain languages. Under this open UG vision, specific elements of l-languages arise in response to new parses. Both external and internal languages play crucial, interacting roles: unstructured, amorphous external language is parsed and a structured internal language system results. My **Born to parse** (Lightfoot 2020) explores case studies that show innovative parses of external language shaping the history of languages. I discuss 1) how children learn through parsing, 2) the role of parsing at the two interfaces between syntactic structure and the externalization system (sound or sign) and logical form, 3) language change, and 4) variable linguistic properties seen through the lens of an open UG. This, in turn, yields a view of variable properties akin to that of evolutionary biologists working on Darwin’s finches; see section 7.



OPEN ACCESS

### EDITORS

– Miguel Oliveira, Jr. (UFAL)  
– René Almeida (UFS)

### REVIEWERS

– Janayna Carvalho (UFMG)  
– Cilene Rodrigues (PUC-Rio)

### DATES

– Received: 10/09/2020  
– Accepted: 10/12/2020  
– Published: 01/13/2021

### HOW TO CITE

LIGHTFOOT, David W. (2021). Problems With Variable Properties In Syntax. *Cadernos de Linguística*, v. 2, n. 1, p. 01-19.

## RESUMO

Como as aves canoras, que nascem para gorjear, os humanos nascem para processar informação linguística. As crianças apresentam predisposição para atribuição de estrutura às externalizações amorfas de pensamento que encontram no decorrer do desenvolvimento. Isso nos dá uma visão de propriedades variáveis bem diferente daquela baseada em parâmetros definida na Gramática Universal (GU). Nossa abordagem da aquisição da linguagem oferece duas contribuições para o pensamento minimalista. Em primeiro lugar, em consonância com as metas do Programa Minimalista, minimizamos o componente pré-programado da linguagem (Língua-I(nterna)), dispensando três entidades centrais separadas: parâmetros, métrica avaliativa da capacidade gerativa de gramáticas, mecanismos independentes de processamento. Contrariamente, as crianças usam a gramática interna para processar o ambiente linguístico externo que experienciam. Isto é, a GU é “aberta”, sendo consistente com o que as crianças aprendem por meio do *parser*. Em segundo lugar, nosso entendimento da aquisição da linguagem leva a uma nova visão sobre propriedades variáveis, propriedades particulares de certas línguas. Dentro dessa visão aberta da GU, elementos específicos da Língua-I emergem como respostas a processamentos inovadores. Desse modo, tanto a língua-I como a língua-E(xterna) desempenham papéis fundamentais e interativos: línguas externas amorfas, não estruturadas, são processadas, tornando-se um sistema linguístico interno estruturado. No livro **Born to parse** (Lightfoot 2020), exploramos estudos de caso indicadores de processamentos inovadores de línguas externas, que moldam a história das línguas. Discutimos 1) como as crianças aprendem por meio do processamento, 2) o papel do processamento nas interfaces entre estrutura sintática e sistema de externalização (som ou sinais) e forma lógica, 3) mudança linguística, e 4) propriedades linguísticas variáveis vistas através das lentes de uma GU aberta. Essa proposta nos fornece uma visão de propriedades variáveis similar à perspectiva da biologia evolutiva sobre os tentilhões de Darwin; ver seção 7.

## KEYWORDS

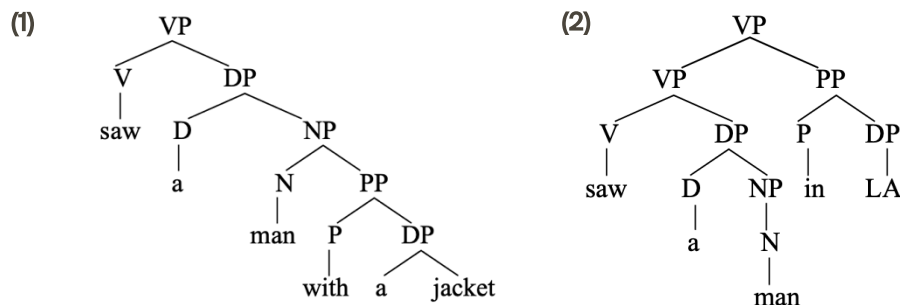
Open UG; Parsing; Parameters; Internal Languages. Minimalism.

## PALAVRAS-CHAVE

GU Aberta; Processamento; Parâmetros; Línguas Internas; Minimalismo.

## INVARIANT PRINCIPLES

For decades, generative syntacticians have been discovering invariant principles holding of all internal language systems and, for the last twenty-five years, proponents of the Minimalist Program have been seeking to minimize and naturalize those principles in ways that make them biologically plausible. One version appeals to invariant computational operations of Project and Merge, which build hierarchical structures bottom-up and combine two elements, a head and a complement (1) or a phrasal category and an adjunct phrase (2). These two skeletal structures suffice for all languages when supplemented by the results of parsing external language, as we shall see.



So in (1) the verb *saw* “projects” to a Verb Phrase also containing a direct object, the Determiner Phrase *a man with a jacket*. The Merge operation brings *saw* and *a man with a jacket* together, forming the VP *saw a man with a jacket*. Similarly, Merge unifies *with* and *a jacket*, creating the Preposition Phrase (PP) *with a jacket*. All internal systems draw on these operations, accounting for the binary branching structures that are everywhere. Merge applies recursively to yield complex structures.

Language systems typically have three recursive devices: relativization, complementation, and coordination, each of which may yield structures of indefinite complexity (3-5).

(3) **Relativization:** This is the cow that kicked the dog that chased the cat that killed the rat that caught the mouse that nibbled the cheese that lay in the house that Jack built on the street where Maria lives.

(4) **Complementation:** Ray said that Kay said that Jay thought that Fay said that Gay told ...

(5) **Coordination:** Ray and Kay went to the movie and Jay and Fay to the store, while Gay and May worked where Shay and Clay watched.

And, of course, these three options may all be used in one expression: *Gay and May said that the man who loves Maria also likes ice cream*. These simple, invariant options permit the generation of expressions of indefinite complexity, and in all language systems: English, Japanese, Quechua, and Nicaraguan Sign Language.

Let us explore some challenges for English internal language systems and begin to get a sense of what this biolinguistic enterprise consists in. European languages have interrogatives where the interrogative word is pronounced at the beginning of the expression and is understood in a wide range of positions marked with a strike-through in (6). In accordance with the Minimalist Program, building these expressions involves multiple re-applications of Merge (yielding relative clauses, complements, and coordinate structures), plus copying and deleting the *wh*-phrase.

- (6) a. Who did we see ~~who~~?  
 b. Who did we speak to ~~who~~?  
 c. Who did she say ~~who~~ left?  
 d. Who did she say that she saw ~~who~~?  
 e. Who did she say she saw ~~who~~?  
 f. Who did she see pictures of ~~who~~?

In all these expressions the interrogative word is pronounced in sentence-initial position and is understood in the strike-through position, so it is copied into the higher position and deleted from its understood position, expressing the thought “who is the person *x* such that we saw *x*?”

Things become interesting when we note a set of expressions that no speaker of English would say (7) (\* indicates an expression that does not exist).

- (7) a. \*Who did she say that ~~who~~ left?  
 b. \*Whose did she see ~~whose~~ pictures?  
 c. \*Who did she wonder whether ~~who~~ left?  
 d. \*Who did she wonder whether Bill visited ~~who~~?  
 e. \*Who did she meet the woman who knew ~~who~~?  
 f. \*Which did she read [~~which~~ books]?  
 g. \*How many did she see [~~how many~~ cars]?  
 h. Combien a-t-elle vu de voitures?

The key question here is what principles do the non-existent forms of (7) violate; put differently, what principles prevent English speaking children from using the non-existent forms of (7)? To sharpen matters, how is it that certain expressions are well-formed in French but the word-for-word translation in English is never said and why do English and French speaking children learn differently (7g,h)? There must be systematic differences that predict why French differs from English.

There are proposals that seem to have the right properties but, rather than giving readers those details, I will let you discuss the issues with well educated graduate students at your local university.

The Principles-and-Parameters model has been around for forty years, since Chomsky 1981, and postulates a set of invariant principles and a set of option points, parameters, defined in UG, allowing the child to select one of two parameter settings. (8) illustrates this with the ideas behind two principles (deletion and a locality restriction) and two binary parameters accounting for variable properties (head directionality and another locality restriction).

### (8) Principles

- something may be deleted, if it is (in) the complement of an adjacent, overt word.
- nothing may move across more than one bounding node.

### Parameters

- {YP, X}
- CP and/or IP are bounding nodes.

The principle governing deletion accounts for the undeletability of the strike-through *who* in (7a), because *who* is not the complement of the adjacent *that*. The parameters require that initial structures have VPs consisting of either V DP order, like English, or DP V order like Dutch, German, Japanese, or Korean.

Consider now VP ellipsis, another construction of English. Just as displaced interrogative elements may be understood in a wide range of positions (6), (9) illustrates the wide range of contexts where VPs may be deleted but the empty VP is always the complement of the overt, adjacent word to its left, as required by the deletion principle in (8), and always understood as “leave for Rio.” (9a) shows two conjoined clauses, (9b,c) shows a main clause and a subordinate clause in different orders, (9d) shows separate sentences, (9e) shows the ellipsed VP embedded within a very complex structure, and (9f) shows an ellipsed VP with no spoken antecedent at all; the syntactic condition is met; the ellipsed phrase is licensed by the overt, adjacent *don’t*, of which it acts as the complement, perhaps understood to mean “Don’t tickle me.”

- (9) a. Max left for Rio today and Kim did <sub>VP</sub>E as well.  
 b. Max left for Rio, although Kim didn't <sub>VP</sub>E.  
 c. Although Max couldn't <sub>VP</sub>E, Kim was able to leave for Rio.  
 d. Max went to Rio. Yes, but Kim didn't <sub>VP</sub>E.  
 e. The man who left for Rio knows the woman who didn't <sub>VP</sub>E.  
 f. Don't <sub>VP</sub>E.

Again we see a wide range of possibilities but there are limits and our invariant principle (8) explains why and where.

English (and some other languages) allows subject pronouns to occur with a quantificational word, *all* or *often*, either preceding or following it (10). Using VP ellipsis and a quantificational word shows interesting effects that follow from our analysis so far.

- (10) a. They all had read it.  
 b. They had all read it.  
 c. They often had seen it.  
 d. They had often seen it.

However, an ellipsed VP may only occur where licensed by an adjacent, overt word of which it serves as a complement (8). Hence the well-formed (11a, b) but not (11c, d).

- (11) a. They denied reading it, although they all had <sub>VP</sub>E.  
 b. They denied reading it, although they often had <sub>VP</sub>E.  
 c. \*They denied reading it, although they had all <sub>VP</sub>E.  
 d. \*They denied reading it, although they had often <sub>VP</sub>E.

Two fundamental properties of internal language systems are that, first, they embody much variation; indeed, it is possible that no two I-languages are identical. I have two daughters, both aged within 18 months of each other, raised under very similar circumstances, attending the same schools, etc, but I know within seconds which one is making the telephone call. The second fundamental property of I-languages is that they are acquired by children in very similar ways. Minimalists have had very little to say about both these fundamental properties. Lightfoot 2020 changes that in ways that I will sketch briefly here.

## 1. VARIATION AND PARAMETERS

Linguists, Minimalists or not, have no biologically coherent, general approach to variable properties that occur in some I-languages but not in others, nor about how they are acquired. Furthermore, we study them in silos. Some of us work on parameters, others on variable rules, and others on constraint re-ranking, and we don't talk to each other about possible commonalities or generalizations. Such things are largely ignored by Minimalists. For proponents of the Principles-and-Parameters approach, our success with parameters comes nowhere near what we have achieved with invariant principles. Beyond this, our two fundamental properties are related: Variable properties, being language-particular, must be learned, triggered by language particularities experienced by children. So it is no surprise that researchers who do not work on acquisition do not focus on variable properties, and vice versa. Nonetheless there has been much recent discussion on problems with parameters; among others, one thinks of work by Theresa Biberauer, Cedric Boeckx, Fritz Newmeyer, Marit Westergaard, and people working in the Cambridge, UK ReCos group (Re-thinking Comparative Syntax) under the leadership of Ian Roberts.

Discussion has focused on the fact that we have no generally agreed theory of parameters. It is sometimes suggested that the so-called "Borer-Chomsky conjecture" provides a basis for such a theory, stipulating that parameters are linked to functional categories. However, that simply transfers the problem and emphasizes the fact that we have no general theory of functional categories. Even worse, we have no theory of how parameters are set, except by the deeply flawed approach of grammar evaluation and input matching. Gibson & Wexler (1994) and Clark (1992) developed such approaches in the 1990's through their (respectively) Trigger Learning Algorithm and Fitness Metric. The TLA sought to identify the grammar with minimal "errors" in parameter setting or the fewest "violations," instances of misgeneration of structures not represented in the child's corpus of expressions generated by the most fit. The TLA encouraged children to seek a better fit between what current parameter settings generate and what children have experienced. One (huge) problem with such attempts is that they postulate that children can remember the totality of what they have experienced and perform elaborate calculations on the full set of possible grammars, each of which generates an infinite number of expressions; children rate the fitness of grammars by counting what those grammars can and cannot generate.

Another approach seeks to treat as triggers for elements of I-languages what those I-languages can generate but this introduces a vicious circularity and obligates investigators to distinguish what the child hears as distinct from negative data concerning what does not occur, which are generally treated as not available to language acquirers. This also entails

that children store what they have heard, a matter that raises huge feasibility issues for claims about childhood memory, to which we shall return.

However, for a Minimalist, seeking to minimize information postulated of the linguistic genotype (i.e. what linguists call “UG”), parameters constitute a more fundamental problem: if parameters are stated at UG (8), they violate the aspirations of the Minimalist Program. Those aspirations encourage us to find an alternative to UG-defined parameters, as I shall advocate below. These input-matching accounts are rendered unfeasible when one factors in the numbers of I-language systems to be rated, roughly a billion if there are about thirty independent, binary parameters, and roughly a trillion if there are forty parameters. Children are “batch learners” and need to know which expressions are in the batch of expressions generated by each I-language system; each of those systems includes an infinite number of expressions generated. The calculations required of children under this parametric view are vitiated by the vast numbers involved, including infinite numbers. This hardly looks feasible.

These difficulties with parameters encouraged Chomsky (2001) to imagine an approach where there are no variable properties, hence no parameters. He postulated the “Uniformity hypothesis;” ‘in the absence of compelling evidence to the contrary, assume languages to be uniform, **with variety restricted to easily detected properties of utterances**’ (my emphasis – DWL). So, we all speak Human and the task of investigators is to find an abstract level of representation where an utterance in Japanese has the same logical form as the equivalent utterance in Quechua. That may be possible one day but, until we have clear proposals along those lines, we need alternative approaches.

Berwick & Chomsky 2016 postulated that the “basic property” of I-language systems is that they have the Merge computational operation. A complementary basic property is to say that fish are born to swim, certain birds are born to chirp, and, in the same sense, the basic property of humans is that they are born to parse, born to assign linguistic structure to what they hear (Lightfoot 2020). Under this view, parsing is central and children invent new variable properties. This exploits a fundamental distinction drawn by Chomsky (1986): E-language is parsed and I-languages result from that parsing. Under that view, we adapt an approach of Colin Phillips (2003), whereby there is no independent parser but rather the I-language itself is the parsing mechanism and yields what is parsed: people parse by assigning linguistic structures made available by their emerging I-language, i.e. what is provided initially by UG and then also by the results of early learning, by the effects of Label and Project. So a child’s parsing capacity becomes richer as his or her I-language develops.

Children parse E-language and postulate specific I-language elements required of aspects of the parse. The aggregation of I-language elements constitutes the complete I-language. When E-language shifts, children may parse differently and thus attain/invent a new I-language, as revealed in work on syntactic change and as we shall discuss in a moment.



Children invent variable properties of their I-languages through parsing; there is no evaluation of I-languages and no parameters or distinct cognitive entity of a parser. UG is open but some things are learned.

Learning paths emerge: a child cannot determine whether an I-language has verb-object VPs until she has identified phrasal categories. Representations are elaborated step-by-step in the course of acquisition, and the structures needed become increasingly abstract and grammar-internal. The emerging learning path is part of linguistic theory, a function of the way in which the structures are stated, as shown by Elan Dresher (1999).

A child discovers the structures and categories needed for parses, using what their current I-language makes available to analyze their ambient E-language. Children learn irregular past tenses, plurals, and that  $_{VP}[V+past\ PP]$  is a structure. They do this by identifying contrasts and conducting a kind of distributional analysis familiar through the work on parsing initiated by Roman Jakobson (1941) and, more recently, by Janet Fodor (1998). Those contrasts enable a child to build its mature I-language and structures like (12) for an expression *The cat sat on the mat*.

(12)  $DP[_D\ the\ N\ cat]_{VP}[_V\ sit+past\ PP[_{on\ the\ mat}]]$

Individuals develop their own, private, internal language, building on genetically prescribed principles, with its elements triggered by the ambient external language, which may shift, as we shall see in the next four sections. I-languages are discrete, biological entities, finite but ranging over an infinitude of recursively enumerable structures, represented in people's brains, and generating expressions and their structures. In all likelihood, no two I-languages are identical. E-language, on the other hand, is a mass sociological notion, amorphous and not a system, in constant flux and appearing differently to different children, and not recursively enumerable.

Under this view, new E-language may yield new parses. There is no restructuring, just new acquisition. Let us examine four new parses that affected the history of English I-language systems.

## 2. FIRST NEW PARSE: MODAL AUXILIARIES

First, in early English the words *can*, *could*, *must*, *may*, *might*, *will*, *would*, *shall*, *should*, *do*, (and sometimes *dare* and *need*) behaved like verbs and were parsed (categorized) as verbs projecting to a VP (13).

(13) Kim<sub>VP</sub>[<sub>VP</sub>can<sub>VP</sub>[visit London]]

(14) Kim<sub>InflP</sub>[<sub>IP</sub>can<sub>VP</sub>[visit London]]

However, this changed and these verbs came to be parsed differently, categorized as Inflection elements and projecting to an Inflection Phrase (14). We know this, because certain expressions ceased to occur in the language, the b forms of (15-19), and indeed they cannot be generated by systems with (14) as the basic structures. For example, if *could* is an Inflection element and if Infl elements occur above a VP and only once per clause, then (15a) can be generated and (15b) cannot. Similarly, if (17a) can be generated with *to* as an Infl element, then (17b) cannot be generated if *can* is also an Infl element.

(15) a. He has seen stars.

b. \*He has could see stars.

(16) a. Seeing stars, she looked for planets.

b. \*Canning see stars, she looked for planets.

(17) a. She wanted to see stars.

b. \*She wanted to can see stars.

(18) a. She will try to see stars.

b. \*She will can see stars.

(19) a. He understands music.

b. \*He can music.

No new relevant constructions were introduced into the language but if words like *can* and *must* ceased to be categorized as verbs, they ceased to have the distributional syntax of verbs. The obsolescence of the b forms in (15-19) is the evidence of the new parse. The fact that those forms dropped out of the language at the same time (on the death of Sir Thomas More) suggests that there was a single change in the internal system. That explains how the new parse was structured but not why the change took place.

Under the approach we have adopted, there can only be one explanation for the new acquisition, namely emergence of new external language: the new E-language entailed the new parse, hence the new I-language.

The major change in the language heard by young speakers, E-language, was the loss of the rich morphology of Old English verbs: verbs had many different forms depending on the tense or the person of the subject DP and the conjugational class of the verb, as is typical of highly inflected languages (20).

### (20) Verb morphology

<b>Present:</b>	fremme, fremst, fremþ, fremmaþ
<b>Past:</b>	fremed, fremedest, fremede, fremedon
<b>Present:</b>	sēo, siehst, siehþ, sēoþ
<b>Present:</b>	rīde, rītst, rītt, rīdaþ
<b>Past:</b>	rād, ride, rād, ridon

Many of these inflectional endings were lost, apparently first under the influence of the Scandinavians living in the north east of England, often in bilingual homes with Norwegian fathers and English mothers. The result was that verb endings became used less and less. However, the antecedents of the modern English modal auxiliaries, sometimes called the “premodals,” never had the third person *-(e)th* ending typical of verbs in early English. In very early English that was one distinction among very many but after centuries of Scandinavian influence, it became distinctive, characterizing verbs: that’s where verbs occurred, with third person singular inflections. Lexical elements without these third person endings were not verbs but Infl elements, projecting not to a VP but to a IP.

## 3. SECOND NEW PARSE: ENGLISH VERBS CEASE TO RAISE

A similar change, also reflecting the loss of inflectional morphology, is another new parse that affected English systems but not those of most other European systems, English taking its own path forward. Early English verbs used to raise to higher positions in negative and interrogative constructions, as in most other European languages like French, German, French, Spanish, Italian, etc. We found expressions like (21a-d) up until the eighteenth century and even beyond.

- (21) a. Sees Kim stars?  
 b. Kim sees not stars.  
 c. Kim sees always stars.

- d. I like not that.
- e. Does Kim see stars?
- f. Kim does not see stars.

These expressions indicated that the structures of (22) were part of early English I-language systems, where the finite verb moves up to the Infl position, as is standard in most European internal language systems (22a). In later English, the morphological elements lower into the VP, as in (22b).

- (22)a. Kim <sub>InflP</sub>[<sub>Infl</sub>[<sub>V</sub>saw] <sub>VP</sub>[saw stars]]
- b. Kim <sub>InflP</sub>[<sub>Infl</sub>~~past~~ <sub>VP</sub>[<sub>V</sub>see+past stars]]

Early English systems, along with the systems of Dutch, Spanish, and French, allowed verbs to raise to the Infl position, <sub>Infl</sub>V. Unlike with our first new parse, new expressions entered the language that had not occurred in earlier generations, namely new forms with the “periphrastic” *do* (21e,f), spreading from the south west, under the influence of Cornish according to John McWhorter (2009). As a result, for every periphrastic *do* form that occurred in the texts, a verb raised to the higher Infl position might also have occurred, yielding expressions like (21).

## 4. THIRD NEW PARSE: PSYCH-VERBS

The most striking change that lends itself to our parsing approach concerns a complex structural shift involving the syntax and meaning of so-called psych-verbs. Early English I-language systems contained psych-verbs like *chance*, *must*, *grieve*, *irk*, *dream*, *need*, *repent*, *think*, and forty or so others, which typically occurred with an initial dative experiencer followed by a nominative theme acting as the subject of the verb (23).

- (23) Gode ne licode na heora geleafleat
- God [dat.] not liked their faithlessness [nom.]
- ‘Their faithlessness did not please God.’

Middle English dictionaries show *licode* meaning to “please” but that changed in ways that we can now understand. As the case endings of Old and Middle English disappeared, expressions like (23) showed up not only without case endings but also with the clause-initial

DP occurring as the subject of the clause and the clause-final DP serving as experiencer and *licode* meaning “enjoy” instead of “please.” Given the changes in the syntax and morphology of the expression, (23) could only mean what it meant for early English speakers if the verb meant “enjoy” instead of “please.”

The before and after analyses of expressions like (23) were quite different. But once the case endings were lost, it is easy to understand why the new parse was adopted. Given the complexity of the changes, it is hard to imagine how one might describe the relevant binary parameters that might characterize the changes in morphology, syntax, and semantics.<sup>1</sup>

## 5. FOURTH NEW PARSE: ATOMIC BE

Another change that challenges the parameter based vision of variable properties is an innovation identified by Anthony Warner. Warner identified some ways in which the speech of Jane Austen differs from Present-Day English. He noted, for example, that the verb *be* showed some surprising behavior: the past tense for regular verbs like *sleep* behaves quite differently from that of past forms like *was*. For example, (24a) shows normal behavior, where the gapped verb to the right of *will* is understood to be *sleep*. However, the past tense of *be*, *was*, seems not to be amenable to a comparable analysis: (24b) does not exist and the gapped verb following *will* may not be understood as *was*. In Present-Day English, the only way to externalize the relevant thought is (24d), with *be* and no gapped verb.

- (24)a. Kim slept well and Jim will, too.  
 b. \*Kim was here and Jim will, too.  
 c. I wish our opinions were the same. But in time they will.  
 (1816 Jane Austen, *Emma*)  
 d. Kim was here and Jim will be, too.

This suggests that the logical form of *slept* in (24a) is the bi-morphemic  $v$ [sleep+past], where the antecedent for the understood verb “sleep” in the right conjunct is the bi-morphemic form indicated, containing “sleep” in the left conjunct. This suggests that in Present-Day English parses *slept* as  $v$ [sleep+past] but *was* is not treated bimorphemically as be+past; rather, *was* is monomorphemic or “atomic.” For

<sup>1</sup> For a similar approach, see Allen 1995, Roberts 2007: ch 2.

expressions comparable to (24a), instead of a gapped verb in the right conjunct, the overt *be* is needed: *Kim was here and Jim will be, too*. In Jane Austen's informal speech, *was* was parsed like *slept*, as *be+past*, and the gapped verb of the right conjunct had an antecedent, capturing the well-formedness of (24c).

So *was* was formerly parsed as two morphemes, *be+past*, but came to be treated atomically. This casts light on another new property of forms of *be*: individual morphological forms developed their own syntactic subcategorization frames: *been* is the only form that may be followed by a directional Preposition Phrase, *She has been to Paris* but not *\*She was being to Paris* nor *\*She was to Paris*. Likewise, only finite forms may be followed by a *to* infinitive to express obligation, *She was to visit Monet* but not *\*She has been to visit Monet*.

Our vision of variable properties is not that children select from a modest number of parameter settings. Rather, variable properties are less disciplined than parameter enthusiasts visualize and we understand that children parse what they hear and invent elements of I-language that will generate, including categorizing words like *will* and *would* as Inflection elements rather than as verbs, like their translations in other European languages. We do not expect internal languages to fall into narrow classes defined by parameters provided at UG. Indeed, we are not surprised to see internal languages being less disciplined, falling into a wider kind of variation and with more unusual properties, for example particular morphological forms with idiosyncratic syntactic properties, as just discussed. So English has developed complex expressions like  $DP[DP[The\ man\ from\ LA]]_D's[NPspeech\ to\ us]]$  and other Germanic languages have not, in the same way that English has developed "stranded" prepositions, unlike other European internal language systems, *The author was spoken to* but not *L'auteur a été parlé à*.

## 7. CONCLUSION

Children use those structures that are expressed by the external language they hear, i.e., required for the analysis of the expressions experienced. The full set of structures used constitutes the mature I-language. Meanwhile, UG is open; we have over-theorized limits on variable properties through binary parameters. Rather, there are no binary, UG-defined parameters and no global evaluation of grammars. Children parse their ambient E-language and invent I-language elements, using what is provided by UG and early learning. English I-language systems have developed idiosyncratic properties; we need an approach to variation that makes this understandable; E-language shifts, leading to new parses, new I-languages. UG is open and some things are learned by children through parsing.

One concern that has motivated some linguists is that the Principles-and-Parameters approach to variable properties is biologically implausible. However, work by evolutionary biologists now suggests that our postulation of an “open UG” may enable us to link arms with some biologists.

Charles Darwin lamented a number of times that neither he nor anybody else had ever witnessed the evolution of a new species and he regarded that as a major failure of his theory. However, many people have pointed to the work of Rosemary and Peter Grant on what became known as Darwin’s finches (Grant & Grant *Evolutionary dynamics of a natural population*, 1989). The Grants identified thirteen species of finches, living on the various islands of the Galápagos archipelago, and differing in the shape of their beaks. Some had large beaks suitable for gathering the large seeds of the islands where they lived; others had beaks suitable for gathering different shaped and different sized seeds from other islands; some ate tree bark and had beaks suitable for gathering soft bark; vampire finches peck the wings and tails of their victims, wounding them and sipping their blood, taking advantage of their sharp beaks. Galápagos finches typically have one of the thirteen beaks the Grants identified, and the specific beak shape is the one suitable for picking up the seeds of the island they inhabit. This specialization developed over time: initially the finches’ genetic material was neutral or “open” with respect to beak size and shape, but natural selection led to further specifications such that the Grants’ correlations between beak characteristics and feeding patterns emerged, reflecting new genetic information. The variation we have seen in the syntax of different languages and in different historical stages of languages is typical of the kind of variation that inspired Darwin and the Grants. It is not the kind of variation that is subject to genetically defined limitations characterized by syntactic parameters. Rather, it reflects the openness of genetic information, the way in which the environment might enhance genetic properties. That enables us to see at least twelve new species of finch evolving in the relevant environments.

Of course, the enhancements we see in Darwin’s finches are different from those we see in three-year-old children: the finch species have selected particular beak shapes, and that selection is inherited by their offspring, whereas the three-year-old child selecting the I-language of some form of English has selected new I-language elements, and each child has to discover their I-language anew. There is no comparable inherited change.

So variable properties across the I-languages of the world may be seen as similar in nature to the variable properties that we see elsewhere in the biological world. And in all these cases, external factors have internal effects, whether on genetic makeup or on emerging I-languages. Variation familiar to biologists is not fundamentally different from what comparative linguists observe. Seeing the similarities may enhance communication between linguists and evolutionary biologists and between different kinds of linguists who have become used to working in their isolating silos. We view UG as open, with its effects

complemented by the very specific effects of parsing. This is analogous to biologists seeing the genetics underlying variation in beak shapes as open enough to be enhanced by the effects of natural selection. This takes us into the world of complex adaptive systems, self-organization, and variation stemming from apparently minor fluctuations and varying initial conditions in evolutionary and cell biology, statistical biophysics, and other factors.

UG keeps languages similar to each other in conforming to invariant properties that are part of our biological endowment. But UG is open, open enough to allow languages to vary as parsing requirements demand, when children discover new contrasts and select new I-language structures accordingly. Evolutionary biologists have found that same kind of variation in the beaks of Darwin's finches and we expect that the parsing-based analysis we have developed and the approach to learning that it entails will lead to a better understanding of language variation than the Principles-and-Parameters vision has yielded, one where information provided by UG is supplemented by information that emerges through learning through parsing.

## 7. QUESTIONS

This paper is a written version of a lecture delivered from my study at home to a worldwide audience by Zoom technology. People attending that lecture posed some interesting questions and observations, which I will address here.

**Stephanne da Cruz Santiago** asked whether lexical items are playing a more important role in theorizing these days. Lexical items have always played an important role in the generative enterprise. Remember Chomsky's **Aspects of the theory of syntax**. Virtually all the new substantive, technical proposals of that book had to do with the nature of the lexicon; certainly in the 1960s theories of the lexicon were hugely important and much was written about the expressive power of transformations and what the balance of work was between transformations and lexical operations. That matter became the major focus of the beginnings of the so-called linguistic wars of that period and went on to establish its own research paradigm, known now as Distributed Morphology.

**Janayna Carvalho** asked how, if we have no parameters, we could account for the fact that several languages show similar properties with respect to verb-subject order, null referential subjects, and null expletives. With or without parameters, if a generativist is to compare hypotheses, the hypotheses will need to be explicit and if the hypotheses concern the acquisition of new systems, researchers will need to identify the trigger experience, what it takes for a child to acquire the mature system identified and, for parametrists, that will entail sketching what children need in order to set the parameters postulated. Janayna asks about similarities between internal languages, particularly those similarities captured by



parameters. I am skeptical about productive generalizations captured by parameters. We have been postulating parameters for forty years but the search for actual parameters has been something of a wild goose chase and we have very little making up the beginnings of a general theory of parameters. Indeed, when one looks at careful examination of putative parameters, one sees much variation within the alleged parameter. For example, work on null subjects in Brazilian (see work by Acrísio Pires) shows a great deal of variation that does not fall under the null subject parameter, different phenomena in different contexts. This is the kind of thing that has led some syntacticians to postulate “micro-parameters.”

My friend from Georgetown’s Psychology Department, **Fathali Moghaddam**, asks the fundamental question, to what degree may I-languages differ from each other? My answer is that they may differ within the limits given by the invariant principles of UG but one has to be careful. Counting grammars makes sense only as part of an effort to compare the generative capacity of I-languages, seeking a system that matches the input, (i.e., generates the expressions found in the primary linguistic data). If the Principles-and-Parameters vision were along the right lines, variable properties being captured by binary parameters and being independent of each other and there being perhaps thirty parameters, there would be just over a billion I-languages; if there were forty such parameters, there would be over a trillion grammars, each generating an infinite number of expressions. As a child compares what number of expressions each grammar might generate, she would perform calculations over astronomical numbers, all of which would need to be stored in the memories of these “batch learners,” which does not look feasible. Children need to remember everything they have been exposed to and what batch it belonged to, that is which grammar generated each expression. Particular expressions do not wear the flags of the I-language that generated them. These are some of the grounds for trying another approach. Under the approach explored here, we might ask how many structural entities might need to be identified and parsed and we would have no reason that I can see to hazard a single number that might constitute a limit. Under our approach, children might vary in the complexity of the mature I-language they invent.

**Sayantani Bamerjee** asks what parameters could be in play universally for nominative morphological case markers but parameters inherently deal with variable properties and principles deal with invariant properties. Careful examination of nominative case markers across I-languages will distinguish variable and invariant properties, where invariant properties follow from UG principles while variable properties follow from aspects of I-languages learned through the mediation of parsing.

**Anderson Silva** asks what linguists should change under the new paradigm advocated here. Linguists should stop waving their hands at vague elements of UG that “explain” universal or parameterized properties of I-languages. Children can learn things by the kind of distributional analysis presupposed by parsing and linguists need to identify what children

learn in their I-languages and what expressions they parse to come up with the necessary elements of I-languages. I believe that much can be learned about language acquisition by examining historical changes that have taken place where we can identify the new parses that have arisen, as I indicated here in sections 3-6. Whether one thinks in terms of UG-defined parameters or in terms of I-language structures resulting from parsing of E-language the clustering of properties will be quite different, as indicated by sections 3-6.

**Jairo Nunes** has made many important contributions to Minimalist analyses and I see the work reported here as offering significant support for the aspirations of the Minimalist Program. Minimalists seek to minimize the genetic information postulated, partly in order to give a plausible account of how that language faculty might have evolved in the species, as discussed by Berwick & Chomsky 2016 and by Ian Tattersall. In **Born to parse** I argue against parameters, an evaluation metric, and an independent parser, and I expect that these arguments will be welcomed by Minimalists. In addition to arguing against these entities, I show how children can learn through parsing what others have attributed to a dangerously enriched, non-Minimalist UG.

Children acquire their internal languages under quite different circumstances and our children invent the elements of their I-languages aided by the invariant principles provided by UG and by what they learn about their I-language through parsing their external language, as we have illustrated in our sketch of new parses. This is true of children learning a new language like a creole or even a pidgin; there is nothing exceptional about such circumstances, as **Michel De Graff** (MIT) and **Enoch Aboh** (University of Amsterdam) have argued for many years. Indeed, we can learn a great deal about the acquisition of new languages, including those that emerged many years ago, like Middle English or what some have proposed to call Anglo-Norse. We have been privileged to live through the emergence of Nicaraguan Sign Language over the last few decades and learned a great deal about the role of biology in the emergence of this new language under unusual circumstances. On the other hand, the notion of parameters, specifically, has not been particularly useful in understanding the acquisition of such new languages and I agree with **Cilene Rodrigues'** skepticism about the usefulness of parameters in understanding the development of so-called partial null subjects in the history of languages like Brazilian Portuguese. Parameters make predictions about how phenomena cluster in acquisition and history and those predictions have not been as fruitful as was hoped in the early days of parameters. That is the principal reason why I have advocated that we need a new research paradigm. From the 1960s onwards, several linguists thought in terms of language acquirers idealized as living in homogeneous speech communities, ignoring the variation that is more often observed. On the contrary, we can learn a great deal about normal language acquisition by studying carefully unusual acquisition, where children are exposed to unusual triggering experiences.

---

## REFERENCES

- ALLEN, Cynthia. *Case marking and reanalysis: Grammatical relations from Old to Early Modern English*. Oxford: Oxford University Press. 2015.
- BERWICK, Robert C. & Noam CHOMSKY *Why only us: Language and evolution*. Cambridge, MA: MIT Press. 2016.
- CHOMSKY, Noam 1981 "Principles and parameters in syntactic theory". *In: Hornstein, Norbert, & David W. Lightfoot* *Explanation in linguistics: The logical problem of language acquisition*. London: Longman, 1981, p. 32-75.
- CHOMSKY, Noam *Knowledge of language: its nature, origin, and use*. New York: Praeger. 1986.
- CHOMSKY, Noam "Derivation by phase" *In: KENSTOWICZ, Michael* *Ken Hale: A life in language* Cambridge, MA: MIT Press, 2001, p. 1-52.
- CLARK, Robin "The selection of syntactic knowledge." *Language Acquisition* 2.1, 1992. p. 83-149.
- DRESHER, B. Elan "Charting the learning path: Cues to parameter setting". *Linguistic Inquiry* 30.1, 1999, p. 27-67.
- FODOR, Janet D. "Parsing to learn". *Journal of Psycholinguistic Research* 27.3, 1998, p. 339-374.
- GIBSON, Edward & Kenneth WEXLER "Triggers". *Linguistic Inquiry* 25.3, 1994, p. 407-454.
- GRANT, Rosemary & Peter GRANT *Evolutionary dynamics of a natural population*. Princeton, NJ: Princeton University Press. 1989.
- JAKOBSON, Roman *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala: Uppsala Universitets Årsskrift. 1941.
- LIGHTFOOT, David W. *Born to parse: How children select their languages*. Cambridge, MA: MIT Press. 2020.
- McWHORTER, John H. "What else happened to English A brief for the Celtic hypothesis." *English Language and Linguistics* 13.1: 2009, p. 163-191.
- PHILLIPS, Colin "Linear order or constituency". *Linguistic Inquiry* 34.1, 2003, p. 37-90.
- ROBERTS, Ian G. *Diachronic syntax*. Oxford: Oxford University Press. 2007.
- WARNER, Anthony. "Predicting the progressive passive: Parametric change within a lexicalist framework". *Language* 71, 1995, p. 533-557.