Chapter 45

LFG and Simpler Syntax

Giuseppe Varaschin
Humboldt-Universität zu Berlin

The theories of Lexical Functional Grammar (LFG) (Kaplan & Bresnan 1982) and Simpler Syntax (SiSx) (Culicover & Jackendoff 2005) both emerged out of a dissatisfaction with the conceptual and formal assumptions of Mainstream Generative Grammar (MGG) (Chomsky 1957, 1965, 1981, 1995). Due to their similar origins, LFG and SiSx have a lot in common: the reduced role of phrase-structure in the explanation of linguistic phenomena, the adoption of constraint-based formalisms and the recognition of autonomous representations for grammatical functions. But there are also crucial differences between the two approaches that relate to some of the most lively issues in linguistics: e.g. the nature of the lexicon and the role of formal grammar in explaining linguistic judgments. The goal of this chapter is to compare these two alternatives to MGG, highlighting their differences and similarities with respect to theoretical and empirical issues.

1 Introduction

The goal of this chapter is to provide a comparison between Lexical Functional Grammar (LFG) and Simpler Syntax (SiSx). Historically, both theories were born out of a dissatisfaction with the conceptual and formal assumptions of Mainstream Generative Grammar (MGG) (Chomsky 1957, 1965, 1981, 1995). Due to their similar origins, LFG and SiSx have a lot in common: the reduced role of phrase-structure in the explanation of linguistic phenomena, the adoption of constraint-based formalisms and the recognition of autonomous representations for grammatical functions, to name a few. But there are also crucial differences that relate to some of the most lively issues in linguistics: e.g. the nature of the lexicon and the role of grammar in explaining linguistic judgments.

Giuseppe Varaschin. LFG and Simpler Syntax. To appear in: Dalrymple, Mary (ed.). Handbook of Lexical Functional Grammar. Berlin: Language Science Press. [preliminary page numbering]

In Section 2, I offer a short summary of the Simpler Syntax Hypothesis (SSH). In Section 3, I lay out some goals and architectural assumptions that SiSx and LFG share, as well some important theoretical differences between the two approaches. Section 4 deals with the motivations for the constructional lexicon assumed in SiSx, which does not adhere to LFG's Lexical Integrity Principle (Bresnan & Mchombo 1995). Section 5 examines the role of constraints that are not part of the grammar, comparing SiSx with an LFG alternative. Section 6 wraps up discussing what LFG and SiSx can learn from each other.

Throughout this chapter, I will assume basic familiarity with the LFG side of the comparison and focus mainly on explaining the SiSx approach. The basic source for the latter is Culicover & Jackendoff (2005), but I will also draw freely on Jackendoff (2002, 2010), Jackendoff & Audring (2019) and Culicover (2009, 2013b, 2021).

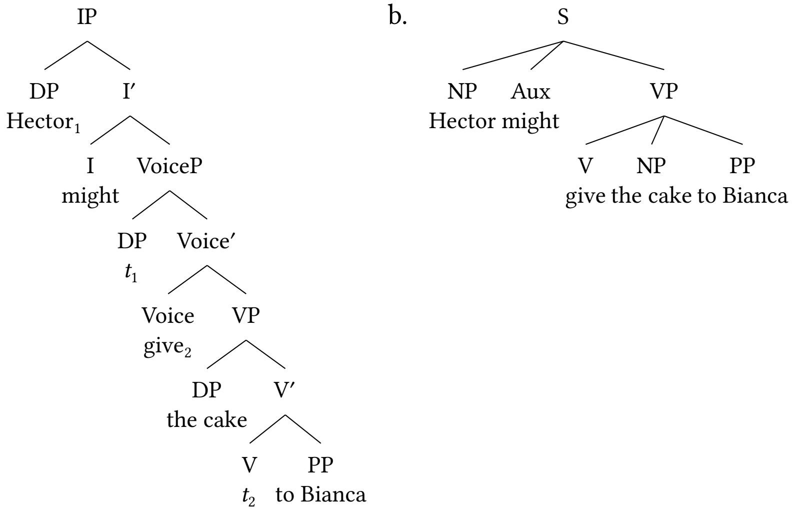
2 The Simpler Syntax Hypothesis

Like other syntactic theories, SiSx is an attempt to describe and explain the language user's ability to establish a correspondence between meaning and sound or gesture. What defines it is the claim that this correspondence should be as minimal as possible – i.e. that syntax should *only* be invoked when other factors (e.g. semantics, prosody, processing) are insufficient to explain the phenomena at hand. This claim is embodied in the Simpler Syntax Hypothesis (Culicover & Jackendoff 2005: 5):

(1) The Simpler Syntax Hypothesis (SSH)
The most explanatory syntactic theory is one that imputes the minimum structure necessary to mediate between phonology and meaning.

Assuming Chomsky's (1965) notions of descriptive and explanatory adequacy, what the SSH says is that, given a set of descriptively adequate grammars of a language *L*, the one the theorist should choose (i.e. the more explanatory one) is the one that assigns less structure to the expressions of *L*. The SSH favors, thus, representational economy (Chomsky 1991; Trotzke & Zwart 2014) over other notions of simplicity, such as minimizing the class of possible grammars or the number of principles in particular grammars. The latter two goals are the main driving forces of MGG since the advent of the Principles and Parameters framework (Chomsky 1973, 1981, 1995).

As an example, contrast the relatively flat constituent structure SiSx assigns to the English sentence *Hector might give the cake to Bianca* in (2b) with the MGG variant in (2a), which is based on the widely adopted VP-shell analysis (Larson 1988; Kratzer 1996; Hale & Keyser 1993; Chomsky 1995):

(2) a. 

MGG opts for structures like (2a) because the grammar that generates them involves *fewer* principles (and is allegedly *more restrictive*) than the one that yields (2b).[1] The idea is that (2a) follows a universal blueprint for structure-building that is virtually *invariant* across languages – one that imposes strict binary branching, endocentricity and a rigid order among heads. Moreover, the hierarchical organization of phrases in (2a) is semantically transparent, reflecting a universal THEMATIC HIERARCHY, in which AGENTS are higher than THEMES, THEMES are higher than GOALS and GOALS are higher than MODIFIERS (see Baker 1997).

The structure itself, however, is clearly much simpler in (2b): (2b) has fewer degrees of embedding (just two), no empty functional projections (e.g. VoiceP) and no phonetically null elements (traces or deleted copies). Given a suitably flexible interface, (2a) can also be placed in correspondence with a level of SEMANTIC STRUCTURE (Jackendoff 1990). The semantic properties that (2a) purports to reflect can be more naturally represented in this level, which is independently required to explain inferences that go well beyond what narrow syntax can express.[2] Thus, between representations (2a) and (2b) – the former illustrating simplicity of principles and the latter simplicity of structure – SSH recommends (2b).

[1]The suggestion that (2a) implies a more restrictive grammatical formalism is probably not true. As Kornai & Pullum (1990) show, as soon as empty elements are introduced, X′-theory becomes equivalent to an arbitrary context-free grammar that can generate structures like (2b). Similar considerations apply to minimalist descendants of X′-theory (cf. Chomsky 1995).

[2]Even the rich structure in (2a) fails to encode the inference that *Hector* is the Source of *the cake* (in addition to the Agent of *give*), or that cakes are artifacts typically used for eating. The latter influences the interpretation of evaluative adjectives: a *good cake* is a cake that is good to *eat* (Pustejovsky 1995). The phrase-structure formalism has no natural way to represent this.

A theoretical reason for pursuing the SSH (as opposed to other measures of simplicity) is that it approximates syntactic structures to what is directly inferable from input, thereby reducing the task of the language learner (cf. Culicover 1998, 1999; Jackendoff 2011a). The child has no direct evidence for the traces and empty elements assumed in (2a). As Chomsky (1982: 19) notes, this raises poverty-of-stimulus issues, which call for the invocation of a richer Universal Grammar (UG). Insofar as SiSx posits more concrete structures, it contributes to the minimalist project of a leaner UG (cf. Chomsky 2005; Hornstein 2009).

Aside from being more *explanatory*, the option for simpler structures is also more *descriptively adequate* than accounts based on rich uniform representations like (2a). Classic constituency tests, for example, only provide motivation for the major constituent divisions shown in (2b): VPs, PPs, NPs, etc. The empirical virtues of the SSH also manifest in accounts of specific linguistic phenomena (some of which will be mentioned in Sections 4 and 5). Most arguments for SiSx analyses have the following form:

> [G]iven some phenomenon that has provided putative evidence for elaborate syntactic structure, there nevertheless exist numerous examples which demonstrably involve semantic or pragmatic factors, and in which such factors are [...] impossible to code uniformly into a reasonable syntactic level [...]. Generality thus suggests that, given a suitable account of the syntax–semantics interface, all cases of the phenomenon in question are accounted for in terms of the relevant properties of semantics/pragmatics; hence no complications are necessary in syntax. (Culicover & Jackendoff 2005: 5)

As this makes clear, the SSH eschews any kind of covert structure that is motivated exclusively in order to provide a uniform mapping onto semantics. This means that SiSx rejects the syntactocentric architecture of MGG – i.e. the view that syntax is solely responsible for the combinatorial richness of language (Culicover & Jackendoff 2005: 17) –, as well as the assumption of Interface Uniformity – i.e. the view that the interface between syntax and semantics is perfectly transparent (Culicover & Jackendoff 2005: 47).

As an alternative, SiSx adopts the Parallel Architecture of Jackendoff (2002), according to which linguistic structure is determined by (at least) three independent formal systems: phonology, syntax and semantics. In addition, SiSx borrows from LFG the idea of a separate syntactic layer for representing grammatical functions: the GF-tier (Culicover & Jackendoff 2005: chapter 6). Each one of these systems is defined by its own characteristic primitives and formation rules and is connected to the others by means of more or less "messy" interfaces:

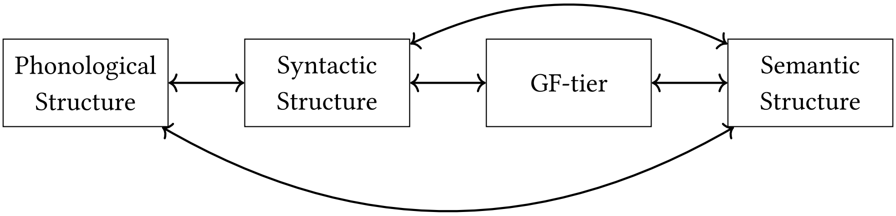


Figure 1: The Parallel Architecture of SiSx

A well-formed sentence must be well-formed in each level, in addition to having well-formed links among the interfaces.[3] A toy example is shown in (3), where natural numbers indicate interface links between the components:[4]

(3)
$$\left[\begin{array}{ll} \text{PHON} & \text{mɛəri}_1\text{\#kɪs}_2\text{+d}_3\text{\#ʤɒn}_4 \\ \text{SYN} & [_{\text{S}}\ \text{NP}_1\ [_{\text{VP}}\ \text{V}_2\ \text{- past}_3\ \text{NP}_4]] \\ \text{GF} & [_{\textsc{pred}}\ \text{GF}_1 > \text{GF}_4]_2 \\ \text{SEM} & \mathbf{past}'_3(\mathbf{kiss}'_2(\textsc{agent}\text{:mary}_1, \textsc{patient}\text{:john}_4)) \end{array}\right]$$

The structure in (3) represents the sentence *Mary kissed John.* The most opaque aspect of the formalism is likely the GF-tier. The basic units of this level are PREDS (short for syntactic predicates), which contain a sequence of ranked positions for syntactic arguments (excluding adjuncts). These positions are not explicitly labeled with grammatical function names, like SUBJECT or OBJECT. For reasons that will become clear in Section 4, these notions are relationally defined as *first GF* of PRED, *second GF* of PRED, etc. The ranking of GFs is determined according the FUNCTIONAL HIERARCHY, which has its roots in Relational Grammar (Perlmutter & Postal 1977, 1983) and Keenan and Comrie's (1977) work.

Note, furthermore, that there is nothing in SYN that signals that NP_1 in (3) corresponds to the string *Mary* – this information is phonological, and, as such, it is only represented in PHON. The terminal strings in a tree like (2b) are, thus, not strictly speaking part of the syntactic structure. A similar division between

[3]An interface link is well-formed *iff* it instantiates some lexeme or construction in the grammar: e.g. the links indicated by subscript 1 in (3) conform to what is stipulated by the lexical entry of *Mary.* The way SiSx represents lexemes and constructions is discussed in Section 4.

[4]Throughout this chapter, I will use the AVM notation adopted in Culicover (2021) for representing linguistic objects and the constraints that such objects must satisfy. For convenience, the formalism for SEM will be a simplified version of Montague's (1974) PTQ appended with an (implicit) event semantics. The thematic predicates (AGENT, PATIENT, etc.) are abbreviations for relations between individuals and the events they partake in, as in Parsons (1990). The SEM tier in (3) is, thus, equivalent to $\exists e[\mathbf{kiss}'(e)\&\ \mathbf{Agent}'(e, \text{mary})\&\ \mathbf{Patient}'(e, \text{john})\&\mathbf{past}'(e)]$.

phonological, syntactic and semantic forms is anticipated in Distributed Morphology (Halle & Marantz 1994; Marantz 1997) as well as in variants of Categorial Grammar that build on Curry's (1963) PHENOGRAMMAR *vs.* TECTOGRAMMAR distinction (e.g. Oehrle 1994; Mihaliček & Pollard 2012).

In order to capture the inner workings of the subsystems of language as well as how these systems interact with each other, SiSx abandons the formal device of derivations in favor of CONSTRAINTS (or, in the terminology of Jackendoff & Audring (2019), SCHEMAS). This and many of the other points mentioned above are shared with LFG, as we will see in the next section. SiSx also draws a lot from HPSG (Przepiórkowski forthcoming [this volume]), as will become particularly clear in Section 4.

3 Goals and assumptions

Among all non-transformational syntactic theories, SiSx and LFG are probably the most closely related ones as far as programmatic aspirations and architectural assumptions are concerned. Most of these stem from the adherence to what Jackendoff (2007b: chapter 2) identifies as two founding themes of Generative Grammar: MENTALISM and COMBINATORIALITY.

MENTALISM is the view that language is a product of the mind/brain of individual speakers. SiSx and LFG are committed to a particularly strong version of this, which Bresnan & Kaplan (1982) and Kaplan & Bresnan (1982), following Chomsky (1965: 9), dub the COMPETENCE HYPOTHESIS. This is the suggestion that the *same* body of knowledge underlies *every* type of language-related behavior (e.g. speaking, reading, learning). In this approach, the linguist's theoretical constructs are not only *psychologically real* in an abstract sense, but must be integrated to an account of how language is actually processed and acquired by real speakers.

The second founding theme of Generative Grammar shared by LFG and SiSx is COMBINATORIALITY: i.e. the view that knowledge of language is instantiated as a finite system of *rules* that define (or "generate") an unbounded array of structured expressions. The linguist's explicit formulation of these rules (i.e. the grammar) must, ideally, entail well-formedness for all sentences judged acceptable by speakers – making no principled distinction between pure manifestations of "core grammar" and "peripheral data" (Culicover 1999).

In line with these commitments, LFG and SiSx seek to characterize the human language capacity in a way that is: (i) PSYCHOLOGICALLY PLAUSIBLE, seeking a graceful integration of linguistic theory with what is known about the structure and function of mind/brain (Bresnan 1978; Jackendoff 2011b); and (ii) FORMALLY and descriptively adequate, representing generalizations of varying granularities with sufficient precision. Different aspects of these objectives are emphasized by LFG and SiSx (e.g., LFG is much more preoccupied with the formal underpinnings and SiSx with the psychological and biological foundations). The remainder of this section summarizes some of the ways the theories converge and diverge in implementing these goals.

3.1 The structure of the grammar

The commitments to mentalism and combinatoriality lead SiSx and LFG to similar conclusions regarding the overall structure of grammar. Compare Figure 1 above, which contains the architecture of SiSx, with the LFG architecture below:

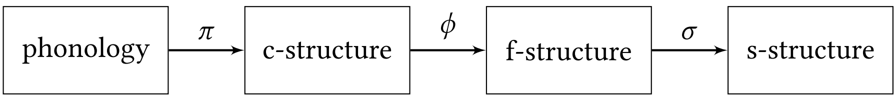


Figure 2: LFG Architecture

The most striking similarity between the two architectures above is that they abide by Representational Modularity, as defined by Jackendoff (1997):[5]

> The overall idea is that the mind/brain encodes information in some finite number of distinct representational formats or "languages of the mind." Each of these "languages" is a formal system with its own proprietary set of primitives and principles of combination, so that it defines an infinite set of expressions along familiar generative lines. For each of these formats, there is a module of mind/brain responsible for it. (Jackendoff 1997: 41)

In both theories, the primitives of phonology are things like segments (or featural decompositions thereof) and syllables. Constituent structure in syntax is built from syntactic categories (e.g. V, N, VP, and Aux) and their dominance and precedence relationships, as in a context-free grammar. The basic units of the GF-tier and f-structure are syntactic predicates and their arguments. Semantics is composed of entities, events, properties and relations (at least). These modules

[5]There are actually different versions of LFG's general architecture going back to Kaplan (1987) (Asudeh 2006; Findlay 2016; Dalrymple & Findlay 2019, among others), but all agree on the essentials of Figure 2. The most striking omission from Figure 2 is the separate component for a-structure proposed in Butt et al. (1997) and subsequently adopted by most researchers within LFG.

are connected to one another via systematic correspondences. In this sense, the architectures in Figures 1-2 can be called CORRESPONDENCE ARCHITECTURES.

The correspondence architecture sets LFG and SiSx apart from sign-based theories like HPSG and SBCG (Przepiórkowski forthcoming [this volume]). The latter use the same kind of data structure to model all aspects of linguistic objects: i.e. typed features organized in AVMs. Different types of information are not related by means of modular correspondences, but in virtue of being values assigned to different attributes of the same sign. The design of HPSG/SBCG does not make it clear that phonology, syntax and semantics are autonomous combinatorial systems. Combinatoriality only exists at the level of signs as a whole (e.g. in features like DTRS, which take lists of *signs* as values, instead of syntactic nodes).

Even though SiSx follows HPSG/SBCG in using AVMs to represent all aspects of linguistic objects, its basic ontology is much closer to LFG's: each linguistic level is conceptualized as an autonomous formal system in its own right. Just as in LFG, this requires positing correspondence principles to link the objects independently defined by each of these systems.

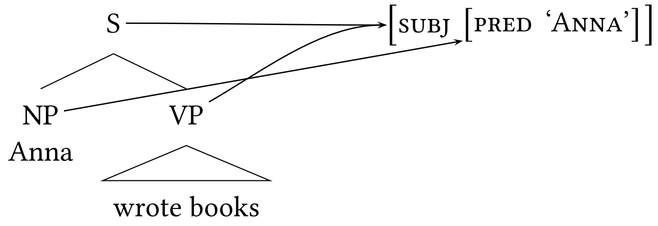
However, LFG and SiSx construe these correspondences in different ways. In LFG, structures of different types are related to each other in virtue of the projection functions π, ϕ and σ of Figure 2. This sort of mapping allows descriptions of elements in the range of a function to be defined in terms of elements in its domain. For instance, the function ϕ – whose domain and range are, respectively, c-structure nodes and f-structures – allows properties of f-structures to be "read off" from c-structure configurations.

This is crucially exploited in LFG's annotated phrase-structure rules. An example is given in (4), where "$*$" stands for the node that matches the element above it in the rule and M is the *mother-of* function (Kaplan 1995: 18):

(4) $$\text{S} \rightarrow \underset{(\phi(M(*))\textsc{subj}) = \phi(*)}{\text{NP}} \quad \underset{\phi(M(*)) = \phi(*)}{\text{VP}}$$

This rule allows one to deduce from the c-structure of *Anna wrote books* (assuming the annotations on lexical entry of *Anna*) the correspondences in (5):

(5)



Since ϕ is a (total) function, it requires that *all* elements in its domain be mapped into elements in its range. This entails that every c-structure node – even nodes corresponding to adjuncts – must be assigned a particular f-structure.

In SiSx, on the other hand, correspondences between structures of different types are not functional, but merely relational. Therefore, there is no sense in which the properties of any level are "projected" from properties of any other, like f-structure is projected from c-structure in LFG. From the point of view of SiSx, this looks like a residue of MGG's syntactocentrism. Consider the SiSx equivalent to LFG's annotated phrase-structure rule in (4) (italics indicate that the element is a variable and not a concrete member of its respective category):

(6) $$\begin{bmatrix} \text{SYN} & [_{S}\ \mathit{NP}_1\ \ \mathit{VP}_2]_3 \\ \text{GF} & [_{\text{PRED}}\mathit{GF}_1\ >\ldots]_{2,3} \end{bmatrix}$$

Like (4), (6) expresses the information that the sister of VP corresponds to a SUBJECT (i.e. the highest ranked GF in a PRED). But, unlike (4), (6) is not a phrase-structure rule: it is a CORRESPONDENCE RULE, which is defined over independently well-formed representations on SYN and the GF-tier. No level has primacy over the others, as suggested by the symmetry of the coindexing notation. Since levels of structure are allowed more independence, the mapping between them can also be seen as only PARTIAL. This avoids the implication that all nodes in SYN must correspond to units on the GF-tier. I will come back to some positive consequences of this looser requirement below.

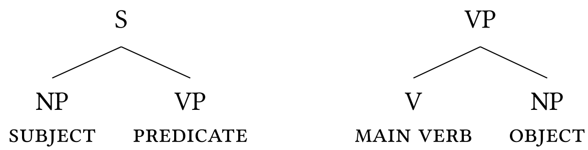
Regardless of these differences, LFG and SiSx both benefit from the general advantages of correspondence architectures, which are better suited for integration with theories of other cognitive faculties than syntactocentric models (this point is hinted at by Bresnan (1993: 45), but see Jackendoff (2007a, 2011b) for full versions of the argument). It is a given that the mind includes relations between non-linguistic representations. For instance, visual and haptic information relate to a modality-independent understanding of the spatial structure of objects (Marr 1982). This spatial structure, in turn, relates to language in a way that allows us to talk about what we perceive (Jackendoff 1987; Landau & Jackendoff 1993). Actions are also spatially guided, requiring an interface between spatial structure and schemas encoding action patterns. It does not make any sense to think of any of these representations as being *algorithmically derived* from any other – they are, rather, related in virtue of modular correspondences.

In this sense, the correspondence architectures of LFG and SiSx see the internal components of language as "connected to each other in the same way as language is connected with the rest of the mind, and in the same way as other faculties of mind are connected to each other" (Jackendoff & Audring 2019: 8). Though many details about how such connections work remain unknown, LFG and SiSx seem better suited for fruitful cross-disciplinary dialogue with cognitive science than MGG, which opts for a syntactocentric derivational design.

3.2 The role of grammatical functions

In any theory, grammatical functions (GFs) serve as abstract "relators" between a class of surface syntactic properties (e.g. linear order, case marking) and semantic roles. MGG assumes that these abstract GFs are represented in the same format as syntactic groupings – i.e. GFs are treated as epiphenomena of constituent structure configurations. An early statement of the MGG view is found in Chomsky (1965: 68–74), who claims that notions like subject and object are universally definable in terms of the structural positions in (7):

(7)



LFG and SiSx both reject this configurational design of UG for similar reasons. Consider what it implies for the English sentence in (8):

(8) Brad seems to like Janet.

In (8), *Brad* behaves like the subject of two predicates: the one headed by *seem* (where it establishes agreement) and the one headed by *like* (where it gets interpreted semantically). The configurational design requires that each of these GFs be realized in different positions, which *Brad* has to occupy simultaneously. This, however, is technically impossible in a typical phrase-structure system, since it entails multi-dominance. The alternative is to posit a sequence of phrase-markers in which these positions are occupied at separate stages, as in (9):

(9) seems [$_S$ Brad [$_{VP}$ to like Janet]] ⇒ [$_S$ Brad$_i$ [$_{VP}$ seems [$_S$ t_i to like Janet]]]

The configurational design thus calls for operations that map phrase-markers onto phrase-markers – i.e. syntactic transformations (Chomsky 1957: 44). Note, however, that these mappings are simply a way to encode the effects of multidominance in a system that does not naturally allow for it.

Though this might seem plausible for English (where subjects typically correspond to the configuration in (7)), it is less appealing for languages like Russian, where word order is freer and GFs are signaled mainly by case endings on nouns. A derivation for the Russian OVS sentence (10) would have to look like (11):

(10) Russian
Vaz-u razbila Olj-a (Kallestinova 2007: 30)
vase-ACC broke Olya-NOM
'Olya broke the vase'

(11) $[_S \text{ Olja } [_{VP} \text{ razbila vazu}]] \Rightarrow [_{S'} [_{VP} \text{ razbila vazu}]_i \ [_S \text{ Olja } t_i]]$
$\Rightarrow [_{S''} \text{ vazu}_k \ [_{S'} \ [_{VP} \text{ razbila } t_k]_i \ [_S \text{ Olja } t_i]]]$

The SUBJECT and OBJECT in (11) are base-generated in the positions signaled in (7) and then scrambled to where they are actually pronounced via roll-up movements (cf. Bailyn 2003). The resulting structure is a representation of "several types of information that seem quite dissimilar in nature" (Kaplan & Zaenen 1989: 137): on the one hand, GFs like SUBJECT and OBJECT and, on the other, linear order, dominance relations and syntactic categories.

LFG and SiSx reject this on the grounds of REPRESENTATIONAL MODULARITY. Dominance, order and syntactic categories are naturally represented in a phrase-structure system but the organization of GFs has different formal properties (e.g. multi-dominance) that justify positing a separate component. This is the GF-tier in SiSx and f-structure in LFG. A SiSx analysis of (8) is sketched in (12) (from now on, tenses will be ignored and PHON will be simplified as orthography):

(12)
$$\begin{bmatrix} \text{PHON} & \text{Brad}_1 \text{ seems}_2 \text{ to like}_3 \text{ Janet}_4 \\ \text{SYN} & [_S \text{ NP}_1 \text{ V}_2 \ [_{VP} \text{ V}_3 \text{ NP}_4]] \\ \text{GF} & [_{\text{PRED}} \text{ GF}_1]_2 [_{\text{PRED}} \text{ GF}_1 > \text{GF}_4]_3 \\ \text{SEM} & \textbf{seem}'_2(\textbf{like}'_3(\text{EXPERIENCER:brad}_1, \text{THEME:janet}_4)) \end{bmatrix}$$

In the GF-tier, GF_1 (which corresponds to *Brad*) is doubly dominated by the PRED linked to *seem* and the one linked to *like*. This direct encoding of multi-dominance – which is also central to LFG's functional control analysis of raising (see Bresnan 1982a) – makes transformations like (9) unnecessary.

Likewise, the autonomy of GFs in SiSx and LFG also makes it possible to state mappings between GFs and SYN without specifying syntactic configuration or linear order. So, for dependent-marking languages like Russian, GFs can be linked directly to Ns with the appropriate case morphology, as in (13) (Culicover 2009: 154).

(13) a. $\begin{bmatrix} \text{SYN} & [_S \ldots N\text{-NOM}_1 \ldots]_2 \\ \text{GF} & [_{\text{PRED}} \ GF_1 > \ldots]_2 \end{bmatrix}$ b. $\begin{bmatrix} \text{SYN} & [_S \ldots N\text{-ACC}_3 \ldots]_4 \\ \text{GF} & [_{\text{PRED}} \ GF > GF_3 \ldots]_4 \end{bmatrix}$

This proposal avoids abstract *ad hoc* MGG derivations like (11), opening the possibility of licensing flat structures. A SiSx analysis for (10) in this spirit could be something like (14). Note that configuration does not play a role in determining GFs in this case. (This does not mean that it cannot play a role in defining information structure properties, which are not being represented in (14).)

(14)
$$\begin{bmatrix} \text{PHON} & \text{Vaz-u}_3\ \text{razbila}_2\ \text{Olj-a}_1 \\ \text{SYN} & [_{\text{S}}\ \text{N-ACC}_3\ \text{V}_2\ \text{N-NOM}_1] \\ \text{GF} & [_{\text{PRED}}\ \text{GF}_1 > \text{GF}_3]_2 \\ \text{SEM} & \textbf{break}'_2(\text{AGENT:olya}_1, \text{PATIENT:the-vase}_3) \end{bmatrix}$$

The idea that word parts can carry information about GFs bypassing syntax is shared with LFG (Bresnan 2001). The proposal sketched in (13–14) bears a particularly close resemblance to Nordlinger's (1998) CONSTRUCTIVE CASE theory.

Notwithstanding their similar motivations, LFG's f-structures and the GF-tier in SiSx have very different formal properties. The most striking of these is the fact that GFs in SiSx are UNLABELED; hence, notions like SUBJECT and OBJECT are not primitives of the theory. They are defined RELATIONALLY in terms of a hierarchy of arguments, as in Relational Grammar (Perlmutter & Postal 1977, 1983) – the most direct inspiration for the GF-tier, according to Jackendoff (personal communication). A motivation for this will be given in Section 4.[6]

Another peculiarity of the GF-tier is that it lacks the unlimited embedding found in LFG's f-structures. Each PRED in the GF-tier is represented as a self-contained unit. There is no sense in which the PRED that corresponds to *like* in (12) is embedded under the one that corresponds to *seem*. The f-structure LFG assigns to the same sentence, on the other hand, virtually mirrors the hierarchical organization of the c-structure from which it is projected:

(15)
$$f_1\!:\!\begin{bmatrix} \text{PRED} & \text{`SEEM}\langle\text{XCOMP}\rangle\ \text{SUBJ'} \\ \text{SUBJ} & f_2\!:\![\text{PRED}\ \ \text{`BRAD'}] \\ \text{XCOMP} & f_3\!:\!\begin{bmatrix} \text{PRED} & \text{`LIKE}\langle\text{SUBJ, OBJ}\rangle\text{'} \\ \text{SUBJ} & f_2 \\ \text{OBJ} & f_4\!:\,[\text{PRED}\ \ \text{`JANET'}] \end{bmatrix} \end{bmatrix}$$

6 Patejuk & Przepiórkowski (2016) argue that a similar move is advantageous for LFG as well. Following Alsina (1996), they show that most GF labels redundantly represent information already available in morphosyntax and s-structure. Borrowing ideas from HPSG (Przepiórkowski forthcoming [this volume]), they propose to replace GF attributes by a single ordered DEPS list which looks a lot like SiSx's GF-tier. This also allows a direct encoding of the functional hierarchy, which is used in LFG analyses of binding (Falk 2001) and control (Bresnan 1982a).

Moreover, since SiSx is not committed to an exhaustive mapping from SYN nodes to the GF-tier, the inventory of GFs can be much smaller than in LFG. Only elements whose morphosyntactic forms are unrevealing about their semantic roles – e.g. direct NP or CP arguments – actually need a representation on the GF-tier (Culicover 2021: chapter 6). This is not the case for adjuncts and (most) obliques, whose θ-roles are transparent in the morphology or choice of preposition. In English, for instance, PPs headed by *near* and *under* are always LOCATIONS while those headed by *during* and *after* are invariably interpreted as TIMES. Correspondence rules for these elements can, thus, be stated directly as relations between SYN and SEM, circumventing the GF-tier (as anticipated in Figure 1).

The GF-tier in SiSx is, therefore, restricted to LFG's CORE GFs (Bresnan 2001: 96): SUBJ, OBJ and OBJ2 (relations 1, 2 and 3 in Relational Grammar). These are the GFs that most strongly justify a tier for GFs in the first place, because they are the typical targets for phenomena like agreement, raising, passive, and structural case-marking– none of which can be stated in terms of direct correspondences between SEM and SYN (Culicover & Jackendoff 2005: 188–189). LFG's NON-CORE functions (e.g. ADJ, OBL_θ, COMP, XCOMP) are not necessary in SiSx.

What this shows is that, all in all, most of the richness that is present in SYN and SEM is absent from the GF-tier, which ends up being a much *simpler* level than LFG's f-structure. This derives from the fact that SiSx builds upon a more radical version of representational economy than the one LFG assumes – one that applies not only to phrase structure, but to ALL LEVELS OF GRAMMAR. If some correspondences *can* be stated as direct relations between SYN and SEM, SiSx can do this without invoking an intermediate mapping through the GF-tier.

This, however, is only possible because SiSx also abandons the assumption of INTERFACE UNIFORMITY (discussed in Section 2), which is pervasive in MGG and survives – albeit in a much lighter fashion – in LFG's version of the correspondence architecture in Figure 2. It is the idea that the mapping to semantics is established uniformly on the basis of GFs that forces LFG to populate f-structure with semantically relevant c-structure information.

SiSx's more sparing use of GFs is partly motivated by the commitment to what Jackendoff (2011a) calls the EVOLUTIONARY CONSTRAINT – namely, the idea that the architecture of grammar should be compatible with a plausible evolutionary scenario. Proponents of SiSx concur with mainstream evolutionary psychologists in assuming that the emergence of human language was gradual, involving a series of incremental steps (protolanguages), each of which offered some adaptive advantage over the previous one (Pinker & Bloom 1990; Corballis 2017; Dennett 2017; Fitch 2017; Boeckx 2017; Martins & Boeckx 2019; de Boer et al. 2020).

Given the absence of a fossil record, one of the main ways to investigate the particular stages of this incremental process is reverse-engineering: i.e. asking what components of language are advantageous without the whole system in place (Jackendoff 1999; Jackendoff & Pinker 2005; Progovac 2016). In this spirit, Jackendoff (2002: 261) speculates that the GF-tier is probably "the latest developing part of the architecture", since its properties are asymmetrically dependent upon the existence of articulated systems of constituent structure and semantics – i.e. the latter two components can exist without the GF-tier, but not vice-versa. It is hard to reconcile the LFG architecture – where f-structures are essential to the mapping between c-structure and semantics – with these considerations.

Regardless of these differences, the point remains that autonomous levels for GFs (as we see in LFG and SiSx) contribute to the overall simplification of the grammar. Insofar as these levels liberate syntax from encoding GFs configurationally, constituent structure can become more concrete. The next section shows that this is an advantage for theories that take psychological plausibility as a goal.

3.3 Surface-oriented and model-theoretic grammars

Like HPSG (Przepiórkowski forthcoming [this volume]) and Construction Grammar (Matsumoto forthcoming [this volume]), LFG and SiSx are SURFACE-ORIENTED. A model of grammar is SURFACE-ORIENTED if it posits syntactic structures that are directly associated with observable word strings, with a minimum of empty elements and degrees of embedding. In LFG and SiSx, this WYSIWYG flavor is a consequence of the correspondence architecture – which provides *other* levels for encoding GFs and semantic relations – along with principles that enforce representational economy on phrase-structure representations: Economy of Expression in LFG (Bresnan 2001: 91) and the SSH in SiSx.

Surface-orientation is driven by matters of psychological plausibility. Empty elements are not easily detectable from linguistic input. This raises the question of how they come to be learned (as discussed above in connection to the SSH) and inferred in real-time language processing (see Sag & Wasow 2011). The common conclusion is that they are *not learned*, but constitute part of UG. Though this move does solve the learnability problem (albeit by raising the more difficult question of how these elements evolved in humans), it hardly addresses the concern over language processing.

However, learnability and processing issues do not arise if empty elements *can* be inferred on the basis of language-internal evidence. This is arguably the case in situations where invisible structure systematically alternates with visible material, such as gaps in unbounded dependency constructions (see Kluender & Kutas 1993; Clark & Lappin 2011).[7] In these cases LFG and SiSx *do* allow them as a kind of "last resort" to maintain the generality of the mapping between form and meaning (Bresnan 2001: 193; Culicover & Jackendoff 2005: 304).

The status of empty elements in LFG and SiSx is very different from their status in MGG: they are not leftovers of transformations, but directly licensed by CONSTRAINTS. This distinction reflects the contrast between the PROOF-THEORETIC design of MGG and the MODEL-THEORETIC flavor of SiSx, LFG and many other syntactic theories (Pullum & Scholz 2001; Pullum 2013). A PROOF-THEORETIC GRAMMAR (PTG) relies on the technology of stepwise algorithmic derivations to recursively enumerate the infinite set of grammatical expressions in a language. A MODEL-THEORETIC GRAMMAR (MTG), on the other hand, formulates its basic statements as declarative constraints. The objects that satisfy the constraints (i.e. their models, in the logician's sense) are the expressions licensed by the grammar.

The manner of characterizing expressions in PTGs invites the dynamic and procedural metaphors that are routinely employed in the MGG literature. The problem with such locutions is that it is unclear what they should mean in terms of real-time processing. The practical consequence of this has been a gradual stiffening of the competence/performance distinction through the history of MGG.

The MTG formalism avoids all such problems, lending itself to a much more direct relation to processing models (Sag & Wasow 2011; Jackendoff 2007a, 2011b). Since constraints have no inherent directionality, they can be invoked in *any* order. Starting with a fragment of phonology, one can pass through its mappings to syntax and semantics and do the same the other way around. This accounts for the fact that the processor is "opportunistic" and uses diverse types of information as soon as they become available (Acuña-Fariña 2016). It also makes MTGs neutral with respect to production (which goes from semantics to phonology) and comprehension (which goes from phonology to semantics).

Moreover, constraints also yield a monotonic mapping from form to meaning – i.e. there are no destructive operations that throw out information inferable from parts of a structure. This makes MTGs suitable to deal with the grammaticality of linguistic fragments and with the incremental nature of parsing – yet another

[7] For most of these scenarios, it can also be shown that grammars *with* empty elements are extensionally equivalent to grammars *without* them. This effectively reduces empty elements to notational devices for stating generalizations more directly and reducing the overall complexity of the grammar (see Müller (2018: chapter 19) for discussion). If one assumes a simplicity-based evaluation metric like the one in Chomsky (1951), this notational choice actually has empirical consequences for language acquisition (see Chomsky (1965: 45) for a similar point).

desirable property in light of psychological adequacy (Cahill & Way forthcoming [this volume]).

SiSx and LFG can both be naturally stated as MTGs (cf. Blackburn & Gardent 1995; Pullum 2019 for some caveats). This has practical consequences for the general architecture of the grammar. As we will see below, in a full-blown MTG, it is no longer necessary to uphold a rigid distinction between the lexicon and the grammar, because BOTH can be stated in the same format: i.e. as CONSTRAINTS.

4 The structure of the lexicon

Up to now, I have talked mostly about how SiSx and LFG represent the structure of LINGUISTIC OBJECTS. This section turns to the kinds of CONSTRAINTS that are responsible for licensing these objects. A widespread assumption is that these constraints fall into two radically different classes, depending on whether they apply to WORDS and their internal parts or to larger PHRASAL UNITS. This view is famously expressed in LFG's LEXICAL INTEGRITY PRINCIPLE (LIP):

(16) THE LEXICAL INTEGRITY PRINCIPLE (Bresnan & Mchombo 1995: 181): Words are built out of different structural elements and by different principles of composition than syntactic phrases.

LFG enforces LIP by separating the LEXICON from the RULES OF (PHRASAL) GRAMMAR. The latter are responsible for the organization of novel phrases while the former is supposed to register idiosyncrasies as well as capture some partial regularities among stored items (in the form of LEXICAL RULES).[8]

SiSx argues that there is much to be gained by abandoning this distinction. The first step of the argument involves asking WHAT THE LEXICON IS. Due to the MENTALIST commitment, SiSx frames this issue in essentially psycholinguistic terms, taking the lexicon to be whatever the language user has to learn and store

[8] In its contemporary form, this distinction dates back to Chomsky's (1970) LEXICALIST HYPOTHESIS. In that framework, however, the divide between LEXICAL RULES and RULES OF GRAMMAR overlapped with the distinction between CONSTRAINTS and ALGORITHMS. In a MTG – where ALL rules are stated as constraints – these two kinds of rules can only be distinguished by the types of variables they contain: variables on lexical constraints range over word-like elements and the ones on grammatical constraints range over phrases. LIP is, then, a requirement that constraints containing different types of variables involve fundamentally different relations (i.e. "different principles of combination"): e.g. constraints on word formation should not mention long-distance relationships between items, like the ones found in phrasal grammar. Though this is requirement is formulable in a MTG setting, it is not clear whether it can be empirically justified. See Bruening (2018) for some relevant discussion.

in long-term memory. The argument then goes on to show that a lexicon thus conceived must contain entries of such variety that a sharp distinction between lexical items and grammatical rules becomes artificial (see Jackendoff 1997; Culicover et al. 2017; Jackendoff & Audring 2019, among others). The slippery slope from words to rules of grammar prompts SiSx to view the latter as PART OF the lexicon, as in Construction Grammar (Goldberg 1995; Sag 2012). This looks natural under an MTG design, where lexicon and grammar are equally stated as CONSTRAINTS.

A typical instance of a lexical item is an individual word like *cow*. SiSx, following the Parallel Architecture in Figure 1, treats this as an interface rule, linking a small piece of phonology, a syntactic category and a meaning, as in (17):

(17) $$\begin{bmatrix} \text{PHON} & \text{cow}_1 \\ \text{SYN} & \text{N}_1 \\ \text{SEM} & \lambda x[\mathbf{cow}'_1(x)] \end{bmatrix}$$

The same format can be used to represent items with idiosyncratic subcategorization properties that do not follow from general linking rules. The verb *depend*, for example, subcategorizes for an NP within a PP headed by *on*, as in (18):

(18) $$\begin{bmatrix} \text{PHON} & \text{depend}_1\ \text{on}_2\ \varphi_3 \\ \text{SYN} & [_{\text{VP}}\ \text{V}_1\ [_{\text{PP}}\ \text{P}_2\ NP_3]] \\ \text{SEM} & \lambda y[\lambda x[\mathbf{depend}'_1(\textsc{experiencer}{:}x, \textsc{theme}{:}y)]](\sigma_3) \end{bmatrix}$$

Italicized elements and Greek letters represent typed variables that must be contextually instantiated in order for the item to be licensed (Culicover 2021). They are what give lexical items their combinatoric potential.

Productive morphology receives a similar treatment. Since regular forms *can* be computed online – and MUST be so computed in agglutinative languages like Turkish (Hankamer 1989) – we cannot require every one of them to be stored in the lexicon (Jackendoff 1997, 2002). Therefore, regular affixes must have their own lexical entries with variables specifying the phonology, category and semantics of their putative roots – as was also assumed in American Structuralist models of immediate constituent analysis (Bloomfield 1933). (19) is an entry for the English past suffix.

(19) $$\begin{bmatrix} \text{PHON} & \varphi_2\text{-ed}_1 \\ \text{SYN} & [_V V_2\text{-}\textsc{past}_1] \\ \text{SEM} & \mathbf{past}_1(\sigma_2) \end{bmatrix}$$

Note that, as far as SiSx is concerned, there is no deep formal distinction between the *syntactic* combinatoriality of the verb in (18) and the *morphological* combinatoriality of the affix in (19). The only difference has to do with the nature of the variable in SYN: NP_3 in (18) is phrasal and V_2 in (19) is not. So SiSx, unlike LFG (see Sadler & Spencer 2004), has no separate MORPHOLOGICAL COMPONENT.

A lexicon conceived in these terms should also contain a variety of multiword entries (Culicover et al. 2017). Among these are idioms with fully specified material on all tiers, such as *kick the bucket*. In SiSx, these expressions can be stored as whole phonological/syntactic units, linked to noncompositional semantics, as in (20). We know that this particular idiom instantiates the canonical syntactic structure of an English VP because *kick* inflects just like an ordinary verb (e.g. *John kick**ed** the bucket*, *John **will** kick the bucket*, etc.).

(20) $\begin{bmatrix} \text{PHON} & \text{kick}_1\ \text{the}_2\ \text{bucket}_3 \\ \text{SYN} & [_{\text{VP}}\ \text{V}_1\ [_{\text{NP}}\ \text{Det}_2\ \text{N}_3]]_4 \\ \text{SEM} & \lambda x[\textbf{die}'_4(\textsc{patient}{:}x)] \end{bmatrix}$

Like the verb in (18) and the affix in (19), some idioms have variables that grant them combinatorial potential of their own. These are cases like *stab NP in the back*, *put NP on ice* and *catch NP's eye*. Here is a lexical entry for this last one:

(21) $\begin{bmatrix} \text{PHON} & \text{catch}_1\ \varphi_2\text{'s}_3\ \text{eye}_4 \\ \text{SYN} & [_{\text{VP}}\ \text{V}_1\ [_{\text{NP}}\ NP_2\text{-GENITIVE}_3\ \text{N}_4]] \\ \text{SEM} & \lambda x[\textbf{notice}'(\text{EXPERIENCER}{:}\sigma_2, \text{THEME}{:}x)] \end{bmatrix}$

The entries in (20) and (21) pose a kind of ordering paradox for theories that assume a radical separation between grammar and lexicon, as prescribed by the LIP. The information that *kick the bucket* and *catch NP's eye* are VPs has to be stated *in* the lexicon, because their semantics is idiosyncratic. However, the phrase-structure rule that generates VPs can only apply *outside* the lexicon.

In addition to these cases, the lexicon also has to include a class of CONSTRUCTIONAL IDIOMS that use normal syntax to unusual (i.e. noncompositional) semantic ends (Jackendoff 1997, 2002). An example is the SOUND+MOTION CONSTRUCTION in (22) (Levin & Rappaport Hovav 1995; Goldberg & Jackendoff 2004):

(22) The car [$_{\text{VP}}$ rumbled past Sue].
go′(THEME:the-car, PATH:past-Sue, EFFECT:**rumble′**(the-car))

Syntactically, the VP in (22) is merely a sequence of a verb followed by a PP. Its semantics is unusual because the verb is not interpreted as a functor over the PP, but as specifying the EFFECT of a motion that is not codified by any of the words in the sentence. The effect of the motion, is, moreover, predicated of whoever is interpreted as the THEME (i.e. the entity undergoing the motion). A lexical entry with these properties is sketched in (23).

(23) SOUND+MOTION CONSTRUCTION (adapted from Culicover 2013b: 42):

$$\begin{bmatrix} \text{SYN} & [_{VP}\, V_1\; PP_2] \\ \text{SEM} & \lambda x[\mathbf{go'}(\textsc{theme}{:}x, \textsc{path}{:}\sigma_2, \textsc{effect}{:}\sigma_1(x))] \end{bmatrix}$$

What is peculiar about constructional idioms is that the SYN tier in their lexical entries consists *entirely* of variables that are completely unlinked to phonology.[9] This makes them much more rule-like than word-like.[10] However, since their interpretation does not follow from general principles, they have to be explicitly learned and stored just like words are (see Culicover 1999).

Two other examples of constructional idioms along with the relevant lexical entries proposed in the SiSx literature are given below: (24) represents the DITRANSITIVE CONSTRUCTION (Jackendoff 1990; Goldberg 1995; Asudeh et al. 2014); and (25) represents the PROXY CONSTRUCTION (Nunberg 1979, Jackendoff 1997, Varaschin 2020), wherein the meaning of NP is coerced into a proxy of its literal denotation.

(24) DITRANSITIVE CONSTRUCTION (adapted from Culicover 2021: 40):

a. Brad kicked Janet the ball.

b. $$\begin{bmatrix} \text{PHON} & \varphi_1\; \varphi_2\; \varphi_3 \\ \text{SYN} & [_{VP}\, V_1\; NP_2\; NP_3] \\ \text{SEM} & \lambda x.\mathbf{transfer'}(\textsc{source}{:}x, \textsc{goal}{:}\sigma_2, \textsc{theme}{:}\sigma_3, \textsc{means}{:}\sigma_1(x)) \end{bmatrix}$$

(25) PROXY CONSTRUCTION (adapted from Varaschin 2020: 11):

[9] The existence of "defective" lexical items lacking terms in some level is not surprising in a correspondence architecture. Jackendoff (1997: 94) notes that there are words with phonology, syntax and no meaning (e.g. expletives), others with meaning, phonology and no syntax (*hello, ouch, yes*) and even sequences with nothing but phonology (*e-i-e-i-o, inka-dinka-doo, tra-la-la*). All of these are clearly stored in long-term memory and recognized in the same way typical words are. Moreover, they fit into the phonotactic and stress patterns of English. This indicates that, though some of them have no syntax, they are still part of language. The only reason for excluding them from the lexicon is syntactocentrism – which is abandoned in SiSx and LFG.

[10] This is what drives Asudeh et al. (2013) to propose that idioms like (23) are not derived from lexical entries, but from phrase-structure rules annotated with templates. Other idioms, like the WAY CONSTRUCTION (e.g. *Sue laughed her way out of the restaurant*), would be lexically encoded by individual words (in that case, by *way*). However, it is not clear how this account extends to idioms like (20–21), which are specified by *discontinuous* portions of morphosyntax. Space prevents me from exploring further details of LFG's template-based accounts of constructions.

a. I put ⟨ *book-by* ⟩ Keynes on the top shelf.

b. $\begin{bmatrix}\text{SYN} & \mathit{NP}_1 \\ \text{SEM} & \mathbf{proxy'}(\sigma_1)\end{bmatrix}$

Language turns out to be full of constructional idioms like these (see Goldberg 1995; Jackendoff 1997, 2008; Culicover 1999). However, insofar as recognizing their existence commits us to syntactically complex lexical items without phonology, nothing stops us from seeing general syntactic and interface rules – usually thought of as part of the GRAMMAR – in the same way. The context-free rule for a transitive VP can be construed as a declarative schema for licensing a particular configuration of labeled nodes, as in (26):

(26) TRANSITIVE VP CONSTRUCTION (adapted from Jackendoff 2002: 180):
$\begin{bmatrix}\text{SYN} & [_{\mathit{VP}}\ V\ \mathit{NP}\ ]\end{bmatrix}$

As far as SiSx is concerned, this is simply one of the possibilities allowed by the system: a lexical item with no idiosyncratic phonology or semantics, just syntactic category variables arranged in a particular configuration. In this respect, SiSx deviates from variants of Construction Grammar which require every syntactic configuration to be paired with a meaning (e.g. Goldberg 1995).

Default principles of compositional type-driven interpretation can also be represented as lexical items which license a maximally general correspondence between syntactic variables and meaning variables of the appropriate type. (27) represents the two possible scenarios of Heim & Kratzer's (1998) Functional Application rule (where X, Y and Z are variables over syntactic categories).[11]

(27) COMPOSITIONALITY CONSTRUCTIONS:
a. $\begin{bmatrix}\text{SYN} & [_X\ Y_1\ Z_2] \\ \text{SEM} & \sigma_1(\sigma_2)\end{bmatrix}$ b. $\begin{bmatrix}\text{SYN} & [_X\ Y_1\ Z_2\ ] \\ \text{SEM} & \sigma_2(\sigma_1)\end{bmatrix}$

Likewise, the main intuition guiding linking hierarchies – such as the one in LFG's Lexical Mapping Theory (Bresnan & Kanerva 1989) – can also be formalized, within SiSx, as constructions that establish a correspondence between GF

[11] These general constraints on form and interpretation do not need to be instantiated by ALL grammatical expressions in a language. Many of them are not satisfied in idioms, for instance. For a linguistic object to be licensed in SiSx, it suffices that each of its terms and correspondences fully instantiate SOME constraint (Culicover 2021). This entails that a linguistic object can fail to satisfy a given constraint and still be grammatical AS LONG AS there is some other constraint in the grammar which it satisfies. For instance, the idiom in (22) fails to meet the compositional constructions in (27). Since there is another (more specific) construction which it satisfies (the SOUND+MOTION CONSTRUCTION in (23)), SiSx predicts that (22) is grammatical.

variables and SEM variables. (28) represents the rule that says that the highest thematic argument maps to the first GF.

(28) LINKING CONSTRUCTION (adapted from Culicover & Jackendoff 2005: 185):

$$\begin{bmatrix} \text{GF} & [_{\text{PRED}}\ GF_1\ (> \dots)]_2 \\ \text{SEM} & \sigma_2(\theta{:}\sigma_1, \dots) \end{bmatrix}$$

Correspondences between GFs and SYN – which are accomplished by functional annotations in LFG – can be stated as abstract lexical items as well. The canonical correspondence for SUBJECTS and (transitive) OBJECTS in English are (29a) and (29b), respectively:

(29) ARGUMENT STRUCTURE CONSTRUCTIONS:

a. $$\begin{bmatrix} \text{SYN} & [_S\ NP_1\ VP_2]_3 \\ \text{GF} & [_{\text{PRED}}\ GF_1(> \dots)]_{2,3} \end{bmatrix}$$

b. $$\begin{bmatrix} \text{SYN} & [_{VP}\ V_2 NP_1]_3 \\ \text{GF} & [_{\text{PRED}} GF > GF_1]_{2,3} \end{bmatrix}$$

In this set-up, the PASSIVE can be seen as a more complex strategy for linking the GF-tier to SYN, as in (30) below. The same applies to relation-changing constructions in other languages (e.g. applicatives, anti-passives) (Culicover 2009).

(30) PASSIVE CONSTRUCTION (adapted from Culicover & Jackendoff 2005: 203):

$$\begin{bmatrix} \text{PHON} & \varphi_1\ (\text{by}_2\ \varphi_3) \\ \text{SYN} & [\dots\ V\text{-PASSIVE}_1\ ([_{PP}\ P_2\ NP_3])]_4 \\ \text{GF} & [_{\text{PRED}} GF_3 > [_{\text{PRED}}\ GF]_{1,4}] \end{bmatrix}$$

The construction in (30) looks very much like a non-derivational version of the Relational Grammar account of passivization (Perlmutter & Postal 1977). It expresses two fundamental intuitions: (i) that the first GF (i.e. the "logical subject") is "demoted" to an optional *by*-phrase (without disrupting the link between this GF and its θ-role, as defined by (28)); and (ii) that the second GF gets mapped to SYN like a typical SUBJECT would in virtue of (29a). This last result is accomplished by adding a second pair of brackets around the second GF.[12] A concrete example of a linguistic object which instantiates (30) is given in (31):

[12] This also happens to be the main technical reason why GFs in SiSx are unlabeled. If GFs were defined in terms of substantive roles (e.g. SUBJ, OBJ), as in LFG, a constructional account of relation-changing rules like PASSIVE would involve replacing one function name by another. This would violate monotonicity and Kaplan & Bresnan's (1982) DIRECT SYNTACTIC ENCODING principle. LFG avoids this problem by stating PASSIVE as a LEXICAL RULE (Bresnan 1982c). For evidence that lexical accounts of argument structure (like the one found in LFG) are superior to the SiSx constructional account sketched here, see Müller (2013, 2018). For a lexical account of PASSIVE in SiSx (which resembles the LFG one), see Culicover (2021).

(31) $\begin{bmatrix} \text{PHON} & \text{The-cake}_1 \text{ was-eaten}_2 \text{ by}_3 \text{ Hector}_4 \\ \text{SYN} & [_{\text{S}}\text{NP}_1[_{\text{VP}}\text{V-}\textsc{passive}_2[_{\text{PP}}\text{P}_3\text{NP}_4]]_5] \\ \text{GF} & [_{\textsc{pred}}\ \text{GF}_4 > [_{\textsc{pred}}\ \text{GF}_1]_{2,5}] \\ \text{SEM} & \textbf{eat}'_{2,5}(\textsc{agent}\text{:hector}_4, \textsc{theme}\text{:the-cake}_1) \end{bmatrix}$

SiSx's rule-like lexical entries can play two roles in the grammar: a GENERATIVE ROLE, where they are used in on-line processing to derive novel structures *via* UNIFICATION with other lexical entries; and a RELATIONAL ROLE, where they function like nodes in an inheritance hierarchy, "lending" their structure to other independently stored items (Jackendoff & Audring 2019).

The relational role of lexical entries can be defined in terms of ENTAILMENT between separate constraints stored in the lexicon. A lexical entry α entails an entry β *iff* every linguistic object which is a model of α is a model of β. When a specific lexical entry α entails a more general entry β we can say that α inherits structure from β. In this sense, the *kick the bucket* idiom in (20) inherits structure from the more general VP construction in (26), which, in turn, inherits from a more abstract HEAD-COMPLEMENT CONSTRUCTION, akin to the head-complement schema of HPSG (Pollard & Sag 1994, Przepiórkowski forthcoming [this volume]).

Likewise, if particular passive or past tense verbs happen to be overtly stored due to high frequency, they will inherit from the past tense and passive schemas in (19) and (30). These relational links can be represented in an inheritance hierarchy, where the more dominated nodes entail the less dominated ones. SiSx assumes that, other things being equal, a lexical item with relational links should be easier to store and learn than one without such links (see Jackendoff 1975).

There is an obvious connection between this relational function of lexical entries and the use of templates in LFG and constructions in HPSG/SBCG (Sag et al. 2003; Dalrymple et al. 2004; Asudeh et al. 2013). These devices all do the work of lexical rules in earlier approaches going back to Chomsky (1970). But there is a difference: since many of SiSx's abstract entries can *also* be used generatively, unmarked lexical properties (e.g. regular morphology, subcategorization) can, *in principle*, be kept out of individual lexemes. There is no need to list separately the active, passive and regular past tense forms for ALL verbs. These forms can be "built" by unification with abstract items like (29b), (30) and (19) (respectively) (Culicover & Jackendoff 2005: 188). In LFG terms, it is as if schemas like (29b), (30) and (19) were, at once, templates that can be invoked in particular lexical entries and rules to license novel structures that are not in the lexicon.

The SiSx view, is, in sum, that rules of grammar ARE lexical items. There is a *continuum* from stereotypical words, which specify fully linked phonology, syntax, and semantics (cf. (17)), through idioms with a few variables (cf. (21)), constructional idioms with *nothing but* variables (cf. (23–25)) to fully general rules (cf. (26–30)), from which many constructions can inherit structure. All of these things are stated in the same format: as declarative schemas, either licensing structures at a single level (e.g. (26)) or establishing correspondences between various levels (e.g. (17)). Theories like LFG, which adopt a rigid lexicon/grammar distinction, must draw an artificial line somewhere in this *continuum*.

5 Constraints outside of the grammar

If language is indeed integrated into the larger ecology of the mind, it is expected that grammatical constraints are not all there is to explain the (un)acceptability of sentences. Since Miller & Chomsky (1963), the influence of EXTRA-GRAMMATICAL factors on linguistic judgments has been a major topic of investigation – one that is very much relevant to the pursuit of the SSH. In this section, I explore this issue in connection with the phenomena of UNBOUNDED DEPENDENCIES (UDs).

The hallmark of UDs is the presence of a GAP, by means of which a constituent in a non-canonical position (i.e. a FILLER) acquires its semantic role. In SiSx – as in HPSG (Pollard & Sag 1994: 161) – the effect of a gap can be reproduced by a lexical item that establishes a correspondence between an arbitrary phonological sequence containing the empty string (ε), a constituent containing an XP and a property which results from λ-abstraction over whatever semantics the XP would have (see Muskens 2003 for a similar proposal in Categorial Grammar):

(32) GAP CONSTRUCTION: (adapted from Culicover 2021: chap.7)

$$\begin{bmatrix} \text{PHON} & /\ldots\ \varepsilon\ \ldots\ /_2 \\ \text{SYN} & [\ldots\ \text{XP}\ \ldots]_2 \\ \text{SEM} & \lambda z[\sigma_2(z)] \end{bmatrix}$$

SiSx also needs a phrase-structure construction akin to (26) in order to license fillers in the left-periphery of clauses. (33) accomplishes this effect:

(33) FILLER CONSTRUCTION:

$$\begin{bmatrix} \text{SYN}\ [_{S'}\ YP\ S\ ] \end{bmatrix}$$

Consider how this works in the simple case of topicalization in (34) (I ignore the GF-tier and the information structure status of topics). The construction in (32) licenses an empty NP as the complement of *Janet kissed*, which, in turn, gets interpreted as a property (i.e. $\lambda z[\textbf{kiss}'(\text{AGENT:janet, THEME:}z)]$). (33) licenses a filler (i.e. *Brad*) in sentence-initial position. In virtue of the COMPOSITIONAL CONSTRUCTION in (27b), the property attained by (32) is applied to the semantics of the filler, yielding the right interpretation.

(34) $\begin{bmatrix} \text{PHON} & \text{Brad}_1\text{, /Janet}_2 \text{ kissed}_3\ \varepsilon/_4 \\ \text{SYN} & [_{S'}\ \text{NP}_1\ [_S\ \text{NP}_1\ [_{VP}\text{V}_3\ \text{NP}]]_4] \\ \text{SEM} & \lambda z[\mathbf{kiss}'_{3,4}(\textsc{agent}\text{:janet}_2, \textsc{theme}\text{:}z)](\text{brad}_1) \end{bmatrix}$

A similar structure is ascribed to the *wh*-question in (35). I follow Culicover (2021) in positing a quantifier-like entry for the *wh*-word, as in (36).

(35) [What$_i$ [did [Sue say Don bought t_i]]]?

(36) $\begin{bmatrix} \text{PHON} & \text{what}_1 \\ \text{SYN} & NP_1 \\ \text{SEM} & \lambda P[\mathbf{WH}x_1(P(x))] \end{bmatrix}$

The GAP CONSTRUCTION licenses a property interpretation for the portion of (35) which excludes the *wh*-phrase (*Sue say Don bought ε*). This property, in turn, is fed as an argument to the **WH** quantifier (licensed in initial positon by (33)), which ends up binding a variable corresponding to the gap. (37) illustrates the β-reductions in the SEM tier of (35):

(37) $\lambda P[\mathbf{WH}x(P(x))](\lambda z[\mathbf{say}'(\textsc{agent}\text{:sue}, \textsc{theme}\text{:}\mathbf{buy}'(\textsc{agent}\text{:don}, \textsc{theme}\text{:}z))])$
$\rightarrow \mathbf{WH}x(\lambda z[\mathbf{say}'(\textsc{agent}\text{:sue}, \textsc{theme}\text{:}\mathbf{buy}'(\textsc{agent}\text{:don}, \textsc{theme}\text{:}z))](x))$
$\rightarrow \mathbf{WH}x(\mathbf{say}'(\textsc{agent}\text{:sue}, \textsc{theme}\text{:}\mathbf{buy}'(\textsc{agent}\text{:don}, \textsc{theme}\text{:}x)))$

The constructions (32–33) and standard principles of type-driven interpretation are all SiSx needs to model the syntactic and semantic effects of UDs.[13] The dependency between the filler and the gap is represented as variable-binding, while a null XP in SYN guarantees that the subcategorization requirements of the head that licenses the filler are locally satisfied.

However, since this mechanism assumes that gaps can be freely introduced into representations, it does not explain why sentences like (38) are bad:

(38) * Who$_i$ does that Brad admires t_i disturb Janet?

It is entirely possible to derive a perfectly well-formed structure for (38) given the principles laid out so far. Most approaches to UDs take this "overgeneration" to be a flaw and attempt to encode into the grammar restrictions that prevent gaps from occurring in ISLAND environments like (38) (Ross 1967).

[13] Note incidentally that the type-driven rules in (27) make the presence of subject gaps in sentences like *Who sang?* unnecessary. In those cases, the **WH** quantifier can combine directly with the bare property semantics of the VP, with no need to invoke the GAP CONSTRUCTION.

Kaplan & Zaenen's (1989) LFG account of island constraints exemplifies this tendency. Their proposal represents UDs in terms of functional identity in f-structure (Kaplan & Zaenen forthcoming [this volume]). So, for the sentence (35) above, the identification between the focalized *wh*-word and the OBJ of *buy* is accomplished by the equivalence (f FOCUS)=(f COMP OBJ). This expression is an instantiation of a more general functional uncertainty equation which is annotated to the phrase-structure rule that introduces discourse functions (namely TOPIC or FOCUS). The particular equation Kaplan & Zaenen (1989: 153) suggest for English is (39).

(39) (f DF) = (f {COMP, XCOMP}* GF−COMP)

What (39) says is that the f-structure for any discourse function (DF) will be identical to a subordinate f-structure somewhere along a (possibly empty) path of COMP and XCOMP functions, as long as that path terminates in a GF function which is not a COMP. The specifications on the BODY (i.e. the middle) and on the BOTTOM of uncertainty paths like (39) are how LFG records restrictions on UDs.

For example, an identification between the filler and the gap in (38) requires passing through SUBJ, which is not specified as a possible attribute in the body of (39). This accounts for SUBJECT ISLAND violations in general. Likewise, COMPLEX NP ISLANDS like (40) are also covered, because RELMOD (the GF Kaplan & Zaenen (1989) assign to relative clauses) is not designated on the body of (39) either.

(40) * What castle$_i$ does Janet know the strange man [who owns t_i]?

From the point of view of SiSx, the functional uncertainty formalism is unobjectionable as a device to model UDs. However, it is not clear whether it should really embody substantive restrictions to account for the unacceptability of UDs in syntactic terms. Upon closer examination, there does not seem to be a purely grammatical characterization of precisely the contexts in which certain patterns of UDs are ruled out by speakers. The explanation for most (if not all) island constraints must, therefore, lie outside of the grammar, in pragmatics, discourse structure or in processing complexity. A growing body of literature points to this conclusion (Hofmeister et al. 2007; Hofmeister & Sag 2010; Hofmeister et al. 2013; Kluender 1991, 1992, 2004; Kluender & Kutas 1993; Sag et al. 2007; Chaves 2013; Chaves & Dery 2014, 2019; Culicover 2013a,b). In what follows, I briefly summarize some of the empirical evidence against grammatical theories of islands. Space limitations prevent me from getting into the details of particular performance-based alternatives.

The suspicion that something is amiss in purely grammatical accounts of island phenomena comes from the observation that concrete proposals tend to be both too weak and too strong. The constraint in (39), for example, is too weak because it fails to explain real contrasts like (41–42), originally due to Erteschik-Shir (1973: 84).

(41) a. What$_i$ did Janet claim that veganism can do t_i for you?
b. ?? What$_j$ did Janet transcribe that veganism can do t_j for you?

(42) a. What$_i$ did Frank say that Brad would like t_i for lunch?
b. ?? What$_j$ did Frank snarl that Brad would like t_j for lunch?

The equation in (39) predicts the b-cases to be just as good as the a-cases since, in both of them, the value for the FOCUS attribute is identified with the value of OBJ through a path consisting of a single COMP – exactly as in (35). That is, the a-cases and b-cases both contain (f FOCUS)=(f COMP OBJ) in their f-descriptions.

It is, of course, possible to assign different GFs to the complement of *transcribe* and *snarl* other than COMP (something like ISLANDCOMP). In this case (41b) and (42b) would be excluded due to the body constraint in (39). But this move is simply a stipulation – one that is hard to imagine how a child could learn. The ultimate explanation might be related to the lexical semantics of the verbs (i.e. UDs are impossible with verbs that specify *manner* of speaking) or simply to frequency (*claim* and *say* are more frequent than *transcribe* and *snarl*). Whatever the ultimate truth is, no apparent syntactic difference – in f-structure or otherwise – can be identified for pairs such as (41–42).

There are also cases in which grammatical principles that purport to account for island phenomena are too strong – i.e. they exclude sentences that are actually acceptable. I observed above that (39) derives the effects of SUBJECT ISLANDS and COMPLEX NP ISLANDS. However, UDs whose gaps are contained within Subjects and Complex NPs are reasonably acceptable under suitable conditions (Kluender 2004; Sag et al. 2007; Chaves 2013), as the b-cases in (43–44) show:

(43) a. * Who$_j$ does [that you baked ginger cookies for t_j] irritate you?
b. Who$_i$ does [baking ginger cookies for t_i] irritate you?

(44) a. * Who$_i$ did Phyllis hear the claim [that Bob is dating t_i]?
b. Who$_j$ did Phyllis make the claim [that Bob is dating t_j]?

The equation in (39) rightly excludes (43a) and (44a). The problem is that, by the same token, it also bars (43b) and (44b). Since the a-b pairs are functionally indistinguishable – the bracketed strings map to the same GFs (SUBJ in (43) and RELMOD in (44)) – the real explanation for the contrasts must lie elsewhere.

Kluender (2004) argues that the contrast in (43) is due to a difference in the amount of discourse referential processing. In (43a), the SUBJECT is a finite clause, which introduces the reference to a temporal event. This reference is absent for the non-finite form in (43b), which makes the sentence in question less complex in processing terms (see Gibson (2000) for a similar account).

For (44), Culicover & Jackendoff (2005) suggest an explanation along the lines of Kroch (1998): (44a) presupposes the existence of *the claim* while (44b) doesn't. The unacceptability of (44a) follows from a general principle which says that a gap cannot be referentially dependent on an operator if its reference is part of a presupposition in the discourse. This principle extends to contrasts like (45), which are also hard to account for in purely syntactic terms.

(45) a. * Who$_i$ did he buy that picture of t_i? (presupposes there is a picture)
b. Who did he buy a picture of t_i? (no presupposition)

The debate on whether all island constraints reduce to extra-grammatical factors is still very much ongoing (see Newmeyer (2016) for a useful survey). What this section meant to illustrate is that the SiSx view – which might seem too unconstrained at first glance – could turn out to be just what the data requires. If there is no grammatically coherent characterization of when UDs are unacceptable, then island constraints should not be built into the rules that license UDs (in SiSx terms, they should not be registered as conditions on the GAP CONSTRUCTION). On this view, sentences that incur island violations are not technically ungrammatical, but merely unacceptable for performance-related reasons.[14]

The overall view SiSx ends up with is this: Explanations about our intuitions regarding which structures are possible divide between grammatical constraints (as recorded in the lexicon) and extra-grammatical factors (pragmatics, processing, etc.). The former tend to correlate with sharp judgments, while the latter

[14]Extra-grammatical accounts of island constraints have a long history in SiSx. They go as far back as Jackendoff & Culicover (1972), In this early paper, the authors propose that "perceptual strategy constraints on acceptability" explain otherwise puzzling contrasts like (i):

(i) a. Who$_i$ did John give a book to t_i?
b. *? Who$_j$ did John give t_j a book?

Note that (i) is also not explained by Kaplan & Zaenen (1989), since the equation required to establish the dependency in (ib) – i.e. (f FOCUS) = (f OBJ) – satisfies the constraint in (39).

tend to show more variability and dependence on contextual factors (see Culicover 2013c). Sources of universals are mostly confined to extra-grammatical factors and to the pressure to reduce constructional complexity (Culicover 2013b). These correspond to the THIRD FACTOR properties of Chomsky (2005).

This leads to a very minimalist conception of UG – as it happens, one that conforms (in an unorthodox way) to what Baker (2008: 353) calls the BORER-CHOMSKY CONJECTURE: the hypothesis according to which all parameters of variation among languages are attributed to individual properties of lexical items. In this respect, SiSx is closer to MGG than to LFG. But the difference between SiSx and MGG is that, as discussed in Section 4, lexical items are highly structured and include what are traditionally thought of as rules of grammar. The result is that most aspects of speakers' knowledge of language end up being potentially subject to variation.

6 What can SiSx and LFG learn from each other?

The purpose of this chapter was to survey the theoretical landscape of SiSx and compare it to LFG. This exercise revealed that both approaches seek to reconcile formal theories of grammar and psychological reality – a common goal that leads them to adopt similar architectures and analyses for particular phenomena.

However, despite these programmatic and architectural similarities, the two theories differ in important respects. Many of these differences stem from SiSx's radical commitment to representational economy, which is sustained even when this entails messier and less systematic interfaces. Another source of discrepancies is the explicit recognition, on the part of SiSx, of extra-grammatical influences on linguistic judgments, as discussed in Section 5.

Insofar as SiSx posits fewer constraints and fewer representational devices, less knowledge about abstract linguistic structure (of all kinds) is attributed to learners. This reduces the impulse to posit rich principles of UG, which, in turn, alleviates some of the burden on evolutionary accounts of the language faculty (Jackendoff 1999, 2002; Jackendoff & Pinker 2005). A similar concern with evolutionary adequacy drives current Minimalist work in MGG (Hornstein 2009; Berwick & Chomsky 2015). This does not seem to be much of a worry in LFG, which is more preoccupied with providing a formally precise and computationally tractable framework.

There is sometimes a trade-off between formal refinement and the general goal of unification with other sciences. As we saw in Section 3, the fact that the mapping from form to meaning can bypass the GF-tier in SiSx helps integrating the theory into gradualist scenarios of language evolution, given that it is implausible that stages of protolanguage had anything like abstract GFs (Jackendoff 1999, 2002; Progovac 2016). Since LFG makes the mapping to semantics critically dependent on f-structure, it is hard to imagine a story of how these simpler sound-meaning pairings could have existed in the evolutionary antecedents of language. On the other hand, LFG's rich conception of f-structure lends itself to a much more complete and computation-friendly formalization, which makes the theory more easily testable.

SiSx and LFG can, therefore, learn a lot from each other. LFG can profit from SiSx's more ambitious aspiration of connecting linguistics to human biology. This implies seeking theories of language which are not only descriptively and explanatorily adequate, but which also offer the prospect of integration with plausible evolutionary scenarios. Simpler Syntax, in turn, can benefit from a number of the virtues found in LFG, such as: (i) the development of a formally precise and fully explicit architecture which can feed computational applications and simulations; (ii) the great variety of typologically oriented work which constantly submits the theory's formal assumptions to the test of descriptive adequacy (Part VI).

Once SiSx and LFG assimilate each other's merits, some of the differences between them might diminish and some others might become even sharper. Regardless of the outcome, the process of cross-theoretical comparison is a fruitful one, as it often leads to formal innovations and surprising discoveries about the foundations of linguistic theory.

Acknowledgements

I would like to express my gratitude to Peter Culicover, Ray Jackendoff and Rafaela Miliorini for the constant dialogue and for the numerous comments on previous drafts of this chapter. I would also like to thank Bob Levine, Vitor Augusto Nóbrega and Symon Stevens-Guille for discussions about some of these topics, as well as Mary Dalrymple for the help with LaTeX. This research was partially funded by CAPES with a PROEX scholarship 88882.344838/2019-01, a PDSE scholarship 88881.189650/2018-01. and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1412, 416591334.

References

Acuña-Fariña, J. Carlos. 2016. Opportunistic processing of language. *Language Sciences* 57. 34–48. DOI: 10.1016/j.langsci.2016.05.003.

Alsina, Alex. 1996. *The role of argument structure in grammar: Evidence from Romance* (CSLI Lecture Notes). Stanford: CSLI Publications.

Asudeh, Ash. 2006. Direct compositionality and the architecture of LFG. In Miriam Butt, Mary Dalrymple & Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*, 363–387. Stanford: CSLI Publications.

Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2013. Constructions with Lexical Integrity. *Journal of Language Modelling* 1(1). 1–54. DOI: 10.15398/jlm.v1i1.56.

Asudeh, Ash, Gianluca Giorgolo & Ida Toivonen. 2014. Meaning and valency. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '14 conference*, 68–88. Stanford: CSLI Publications.

Bailyn, John Frederick. 2003. A (purely) derivational approach to Russian scrambling. In Wayles Browne, Ji-Yung Kim, Barbara H. Partee & Robert A. Rothstein (eds.), *Formal approaches to Slavic linguistics*, vol. 11, 41–62. Ann Arbor: University of Michigan Press.

Baker, Mark C. 1997. Thematic roles and syntactic structure. In Liliane Haegeman (ed.), *Elements of grammar: Handbook of generative syntax*, 73–137. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-011-5420-8_2.

Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 351–374. Philadelphia: John Benjamins. DOI: 10.1075/la.132.16bak.

Berwick, Robert C. & Noam Chomsky. 2015. *Why only us: Language and evolution.* Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/9780262034241.001.0001.

Blackburn, Patrick & Claire Gardent. 1995. A specification language for Lexical-Functional Grammars. In *Proceedings of the 7th conference of the European chapter of the ACL (EACL 1995)*, 39–44. Association for Computational Linguistics. DOI: 10.3115/976973.976980.

Bloomfield, Leonard. 1933. *Language.* New York: Henry Holt.

Boeckx, Cedric. 2017. Not only us. *Inference: International Review of Science* 3(1). DOI: 10.37282/991819.17.7.

Bresnan, Joan. 1978. A realistic transformational grammar. In Morris Halle, Joan Bresnan & George A. Miller (eds.), *Linguistic theory and psychological reality*, 1–59. Cambridge, MA: The MIT Press.

Bresnan, Joan. 1982a. Control and complementation. *Linguistic Inquiry* 13(3). Reprinted in Bresnan 1982b, 343–434.

Bresnan, Joan (ed.). 1982b. *The mental representation of grammatical relations.* Cambridge, MA: The MIT Press.

Bresnan, Joan. 1982c. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 3–86. Cambridge, MA: The MIT Press.

Bresnan, Joan. 1993. Interaction between grammar and discourse in Chicheŵa (Bantu). In William A. Foley (ed.), *The role of theory in language description*, 45–60. Berlin: Mouton de Gruyter.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chicheŵa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.

Bresnan, Joan & Ronald M. Kaplan. 1982. Introduction: Grammars as mental representations of language. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, xvii–lii. Cambridge, MA: The MIT Press.

Bresnan, Joan & Sam A. Mchombo. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language & Linguistic Theory* 13(2). 181–254. DOI: 10.1007/bf00992782.

Bruening, Benjamin. 2018. The lexicalist hypothesis: Both wrong and superfluous. *Language* 94(1). 1–42. DOI: 10.1353/lan.2018.0000.

Butt, Miriam, Mary Dalrymple & Anette Frank. 1997. An architecture for linking theory in LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '97 conference*, 1–16. Stanford: CSLI Publications.

Cahill, Aoife & Andy Way. Forthcoming. Treebank-driven Parsing, Translation and Grammar Induction using LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Chaves, Rui P. 2013. An expectation-based account of subject islands and parasitism. *Journal of Linguistics* 49(2). 285–327.

Chaves, Rui P. & Jeruen E. Dery. 2014. Which subject islands will the acceptability of improve with repeated exposure? In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 96–106. Somerville, MA: Cascadilla Proceedings Project.

Chaves, Rui P. & Jeruen E. Dery. 2019. Frequency effects in subject islands. *Journal of Linguistics* 55(3). 475–521.

Chomsky, Noam. 1951. *Morphophonemics of Modern Hebrew*. Philadelphia: University of Pennsylvania. (MA thesis).

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton. DOI: 10.1515/9783110218329.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. New York: Holt, Reinhart & Winston.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications. DOI: 10.1515/9783110884166.

Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding* (Linguistic Inquiry Monographs). Cambridge, MA: The MIT Press.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and parameters in comparative grammar*, 417–454. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/9780262527347.003.0002.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/9780262527347.001.0001.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI: 10.1162/0024389052993655.

Clark, Alexander & Shalom Lappin. 2011. *Linguistic nativism and the poverty of the stimulus*. Oxford: Blackwell. DOI: 10.1002/9781444390568.

Corballis, Michael C. 2017. The evolution of language. In Josep Call, Gordon M. Burghardt, Irene M. Pepperberg, Charles T. Snowdon & Thomas R. Zentall (eds.), *APA handbook of comparative psychology: Basic concepts, methods, neural substrate, and behavior*, 273–297. Washington, DC: American Psychological Association. DOI: 10.1037/0000011-014.

Culicover, Peter W. 1998. The Minimalist impulse. In Peter W. Culicover & Louise McNally (eds.), *The limits of syntax* (Syntax and Semantics 29), 47–77. San Diego, CA: Academic Press. DOI: 10.1163/9789004373167_004.

Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory, and language acquisition*. Oxford: Oxford University Press.

Culicover, Peter W. 2009. *Natural language syntax*. Oxford: Oxford University Press.

Culicover, Peter W. 2013a. English (zero-)relatives and the competence-performance distinction. *International Review of Pragmatics* 5. 253–270. DOI: 10.1163/18773109-13050206.

Culicover, Peter W. 2013b. *Grammar and complexity: Language at the intersection of competence and performance*. Oxford: Oxford University Press.

Culicover, Peter W. 2013c. Simpler Syntax and explanation. In Stefan Müller (ed.), *Proceedings of the 20th international conference on Head-Driven Phrase Structure Grammar*, 263–283. Stanford: CSLI Publications.

Culicover, Peter W. 2021. *Language change, variation and universals – A constructional approach*. Oxford: Oxford University Press.

Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199271092.001.0001.

Culicover, Peter W., Ray Jackendoff & Jenny Audring. 2017. Multiword constructions in the grammar. *Topics in Cognitive Science* 9(3). 552–568. DOI: 10.1111/tops.12255.

Curry, Haskell B. 1963. Some logical aspects of grammatical structure. In Roman Jakobson (ed.), *Structure of language and its mathematical aspects: Proceedings of the twelfth Symposium in Applied Mathematics*, 56–68. American Mathematical Society.

Dalrymple, Mary & Jamie Y. Findlay. 2019. Lexical Functional Grammar. In András Kertész, Edith Moravcsik & Csilla Rákosi (eds.), *Current approaches to syntax: A comparative handbook*, 123–154. Berlin: De Gruyter. DOI: 10.1515/9783110540253-005.

Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2004. Linguistic generalizations over descriptions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '04 conference*, 199–208. Stanford: CSLI Publications.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.). 1995. *Formal issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

de Boer, Bart, Bill Thompson, Andrea Ravignani & Cedric Boeckx. 2020. Evolutionary dynamics do not motivate a single-mutant theory of human language. *Scientific Reports* 10(1). 1–9. DOI: 10.1038/s41598-019-57235-8.

Dennett, Daniel C. 2017. *From bacteria to Bach and back: The evolution of minds*. New York: WW Norton & Company.

Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: Massachussetts Institute of Technology. (Doctoral dissertation).

Falk, Yehuda N. 2001. *Lexical-Functional Grammar: An introduction to parallel constraint-based syntax*. Stanford: CSLI Publications.

Findlay, Jamie Y. 2016. Mapping theory without argument structure. *Journal of Language Modelling* 4(2). 293–338. DOI: 10.15398/jlm.v4i2.171.

Fitch, W. Tecumseh. 2017. Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review* 24(1). 3–33. DOI: 10.3758/s13423-017-1236-5.

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita & Wayne O'Neil (eds.), *Image, language, brain*, 95–126. Cambridge, MA: The MIT Press.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. & Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80(3). 532–568. DOI: 10.1353/lan.2004.0129.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 53–110. Cambridge, MA: The MIT Press.

Halle, Morris & Alec Marantz. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics* 21. 275–288.

Hankamer, Jorge. 1989. Morphological parsing and the lexicon. In William Marslen-Wilson (ed.), *Lexical representation and process*, 392–408. Cambridge, MA: The MIT Press.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.

Hofmeister, Philip, T. Florian Jaeger, Ivan A. Sag, Inbal Arnon & Neal Snider. 2007. Locality and accessibility in *Wh*-questions. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in search of its evidential base*, 185–206. Berlin: Mouton de Gruyter.

Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86. 366–415. DOI: 10.1353/lan.0.0223.

Hofmeister, Philip, Laura Staum Casasanto & Ivan A. Sag. 2013. Islands in the grammar? Standards of evidence. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and the islands debate*, 42–63. Cambridge, UK: Cambridge University Press.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and universal grammar*. Cambridge, UK: Cambridge University Press.

Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language* 51(3). 639–671. DOI: 10.2307/412891.

Jackendoff, Ray. 1987. On beyond zebra: The relation of linguistic and visual information. *Cognition* 26(2). 89–114. DOI: 10.1016/0010-0277(87)90026-6.

Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: The MIT Press.

Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge, MA: The MIT Press.

Jackendoff, Ray. 1999. Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences* 3(7). 272–279. DOI: 10.1016/s1364-6613(99)01333-9.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780198270126.001.0001.

Jackendoff, Ray. 2007a. A parallel architecture perspective on language processing. *Brain Research* 1146. 2–22. DOI: 10.1016/j.brainres.2006.08.111.

Jackendoff, Ray. 2007b. *Language, consciousness, culture: Essays on mental structure*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/4111.001.0001.

Jackendoff, Ray. 2008. *Construction After Construction* and its theoretical challenges. *Language* 84(1). 8–28. DOI: 10.1353/lan.2008.0058.

Jackendoff, Ray. 2010. *Meaning and the lexicon*. Oxford: Oxford University Press.

Jackendoff, Ray. 2011a. Alternative minimalist visions of language. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar*, 268–296. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444395037.ch7.

Jackendoff, Ray. 2011b. What is the human language faculty? Two views. *Language* 87(3). 586–624. DOI: 10.1353/lan.2011.0063.

Jackendoff, Ray & Jenny Audring. 2019. *The texture of the lexicon: Relational morphology and the parallel architecture*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198827900.001.0001.

Jackendoff, Ray & Peter W. Culicover. 1972. A reconsideration of dative movement. *Foundations of Language* 7(3). 397–412.

Jackendoff, Ray & Steven Pinker. 2005. The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition* 97(2). 211–225. DOI: 10.1016/j.cognition.2005.04.006.

Kallestinova, Elena Dmitrievna. 2007. *Aspects of word order in Russian*. Iowa City: University of Iowa. (Doctoral dissertation).

Kaplan, Ronald M. 1987. Three seductions of computational psycholinguistics. In Peter Whitelock, Mary McGee Wood, Harold L. Somers, Rod Johnson & Paul Bennett (eds.), *Linguistic theory and computer applications*, 149–188. London: Academic Press. Reprinted in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 339–367).

Kaplan, Ronald M. 1989. The formal architecture of Lexical-Functional Grammar. *Journal of Information Science and Engineering* 5. 305–322. Revised version published in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 7–27).

Kaplan, Ronald M. 1995. The formal architecture of Lexical-Functional Grammar. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.), *Formal issues in Lexical-Functional Grammar*, 7–27. Stanford: CSLI Publications. Earlier version published as Kaplan (1989).

Kaplan, Ronald M. & Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 173–281. Cambridge, MA: The MIT Press. Reprinted in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 29–130).

Kaplan, Ronald M. & Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In Mark Baltin & Anthony Kroch (eds.), *Alternative conceptions of phrase structure*, 17–42. Chicago: University of Chicago Press. Reprinted in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 137–165).

Kaplan, Ronald M. & Annie Zaenen. Forthcoming. Long-distance dependencies. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.

Kluender, Robert. 1991. *Cognitive constraints on variables in syntax*. La Jolla, CA: University of California, San Diego. (Doctoral dissertation).

Kluender, Robert. 1992. Deriving island constraints from principles of predication. In Helen Goodluck & Michael Rochemont (eds.), *Island constraints: Theory, acquisition and processing*, 223–258. Dordrecht: Springer. DOI: 10.1007/978-94-017-1980-3_8.

Kluender, Robert. 2004. Are subject islands subject to a processing account? In *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, 475–499. Somerville, MA: Cascadilla Press.

Kluender, Robert & Marta Kutas. 1993. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5(2). 196–214. DOI: 10.1162/jocn.1993.5.2.196.

Kornai, András & Geoffrey K. Pullum. 1990. The X-bar theory of phrase structure. *Language* 66(1). 24–50. DOI: 10.1353/lan.1990.0015.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 109–137. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-015-8617-7_5.

Kroch, Anthony. 1998. Amount quantification, referentiality, and long wh-movement. In Alexis Dimitriadis, Hikyoung Lee, Christine Moisset & Alexander Williams (eds.), *Penn Working Papers in Linguistics*, vol. 5(2). Philadelphia: University of Pennsylvania.

Landau, Barbara & Ray Jackendoff. 1993. “What” and “where” in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16(2). 217–238.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19(3). 335–392.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity*. Cambridge, MA: The MIT Press.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2).

Marr, David. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman & Company. DOI: 10.7551/mitpress/9780262514620.001.0001.

Martins, Pedro Tiago & Cedric Boeckx. 2019. Language evolution and complexity considerations: The no half-merge fallacy. *PLoS Biology* 17(11). DOI: 10.1371/journal.pbio.3000389.

Matsumoto, Yo. Forthcoming. LFG and Cognitive and Constructional Theories. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Mihaliček, Vedrana & Carl Pollard. 2012. Distinguishing phenogrammar from tectogrammar simplifies the analysis of interrogatives. In Philippe De Groote & Mark-Jan Nederhof (eds.), *Formal grammar*, 130–145. Berlin/Heidelberg: Springer. DOI: 10.1007/978-3-642-32024-8_9.

Miller, George & Noam Chomsky. 1963. Finitary models of language users. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of mathematical psychology*, vol. 2, 419–492. New York: Wiley.

Montague, Richard. 1974. *Formal philosophy*. New Haven: Yale University Press.

Müller, Stefan. 2013. Unifying everything. *Language* 89(4). 920–950.

Müller, Stefan. 2018. *Grammatical theory: From transformational grammar to constraint-based approaches*. 2nd edn. Berlin: Language Science Press. DOI: 10.5281/zenodo.1193241.

Muskens, Reinhard. 2003. Language, lambdas, and logic. In Geert-Jan Kruijff & Richard T. Oehrle (eds.), *Resource-sensitivity, binding and anaphora*, 23–54. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-010-0037-6_2.

Newmeyer, Frederick J. 2016. Nonsyntactic explanations of island constraints. *Annual Review of Linguistics* 2. 187–210. DOI: 10.1146/annurev-linguistics-011415-040707.

Nordlinger, Rachel. 1998. *Constructive case: Evidence from Australian languages*. Stanford: CSLI Publications.

Nunberg, Geoffrey. 1979. The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy* 3(2). 143–184. DOI: 10.1007/BF00126509.

Oehrle, Richard T. 1994. Term-labeled categorial type systems. *Linguistics and Philosophy* 17(6). 633–678. DOI: 10.1007/bf00985321.

Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics* (Current Studies in Linguistics 19). Cambridge, MA: The MIT Press.

Patejuk, Agnieszka & Adam Przepiórkowski. 2016. Reducing grammatical functions in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-Driven Phrase Structure Grammar and Lexical Functional Grammar*, 541–559. Stanford: CSLI Publications.

Perlmutter, David M. & Paul M. Postal. 1977. Toward a universal characterization of passivization. In Kenneth Whistler, Robert Van Valin Jr., Chris Chiarello, Jeri J. Jaeger, Miriam Petruck, Henry Thompson, Ronya Javkin & Anthony Woodbury (eds.), *Proceedings of the 3rd annual meeting of the Berkeley Linguistics Society*, 394–417. Berkeley: Berkeley Linguistics Society. DOI: 10.3765/bls.v3i0.2266.

Perlmutter, David M. & Paul M. Postal. 1983. Some proposed laws of basic clause structure. In David M. Perlmutter (ed.), *Studies in Relational Grammar 1*, 81–128. Chicago: University of Chicago Press.

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13(4). 707–727. DOI: 10.1017/s0140525x00081061.

Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press & CSLI Publications.

Progovac, Ljiljana. 2016. A gradualist scenario for language evolution: Precise linguistic reconstruction of early human (and Neandertal) grammars. *Frontiers in Psychology* 7. 1714. DOI: 10.3389/fpsyg.2016.01714.

Przepiórkowski, Adam. Forthcoming. LFG and HPSG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Pullum, Geoffrey K. 2013. The central question in comparative syntactic metatheory. *Mind & Language* 28(4). 492–521. DOI: 10.1111/mila.12029.

Pullum, Geoffrey K. 2019. What grammars are, or ought to be. In Stefan Müller & Petya Osenova (eds.), *Proceedings of the 26th international conference on Head-Driven Phrase Structure Grammar*, 58–79. Stanford: CSLI Publications.

Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retoré (eds.), *Logical Aspects of Computational Linguistics: 4th International Conference* (Lecture Notes in Artificial Intelligence 2099), 17–43. Berlin: Springer. DOI: 10.1007/3-540-48199-0_2.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/3225.001.0001.

Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Sadler, Louisa & Andrew Spencer (eds.). 2004. *Projecting morphology*. Stanford: CSLI Publications.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar*, 61–197. Stanford: CSLI Publications.

Sag, Ivan A., Philip Hofmeister & Neal Snider. 2007. Processing complexity in subjacency violations: The complex noun phrase constraint. In Malcolm Elliott, James Kirby, Osamu Sawada, Eleni Staraki & Suwon Yoon (eds.), *Proceedings from the 43rd annual meeting of the Chicago Linguistic Society*, 215–229.

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar*, 359–377. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444395037.ch10.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. Stanford: CSLI Publications.

Trotzke, Andreas & Jan-Wouter Zwart. 2014. The complexity of narrow syntax: Minimalism, representational economy, and simplest merge. In Frederick J. Newmeyer & Laurel B. Preston (eds.), *Measuring grammatical complexity*, 128–147. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199685301.003.0007.

Varaschin, Giuseppe. 2020. Anti-reflexivity and logophoricity: An account of unexpected reflexivization contrasts. *Glossa: A Journal of General Linguistics* 5(1). DOI: 10.5334/gjgl.974.