Derivations or constraints? Core aspects of syntax and morphology in competing grammatical frameworks

Jochen Zeller, Linguistics programme, University of KwaZulu-Natal, Durban, SA
zeller@ukzn.ac.za

1 Introduction

This chapter has two objectives. The first is to provide an overview of some of the main theoretical assumptions and ideas that characterize contemporary versions of the so-called Principles-and-Parameters theory (Chomsky 1981), also referred to – usually by its opponents – as "mainstream generative grammar" (MGG). I focus specifically on the current version of the MGG, known as the Minimalist Program (MP) (Chomsky 1995), and its associated morphological framework, Distributed Morphology (DM) (Halle & Marantz 1993). The second objective is to compare current MGG treatments of core syntactic and morphological phenomena to the analyses offered by alternative, constraint-based models of generative grammar, specifically Construction Grammar (CxG) (Fillmore 1988; Fillmore & Kay 1993; Goldberg 1995, 2006), HPSG (Pollard & Sag 1987, 1994; S. Müller 2015, 2016), and Simpler Syntax (Culicover & Jackendoff 2005).[1]

It is not possible, within the limits of one book chapter, to provide an overview of the MP and DM as well as a complete summary of alternative syntactic and morphological theories (but see Kiss & Alexiadou (2015, Vol. II) and S. Müller (2016) for excellent recent book-

[1] CxG and HPSG belong to what is sometimes called "West Coast" linguistics, as their conceptual foundations were established at the University of California, Berkeley, and the Hewlett-Packard Research Laboratories in Palo Alto. I had to be selective regarding the choice of theories I picked as representatives of West Coast linguistics in this chapter. My choice of CxG and HPSG was determined by my familiarity with these frameworks, but also by the fact that there is some overlap between these theories. For example, a recent version of CxG, "Sign-based Construction Grammar" (Sag 2010, 2012; Michaelis 2013), combines ideas of both Berkeley CxG and HPSG. I have added Culicover & Jackendoff's (2005) Simpler Syntax model to my discussion of these theories, since it is based on the "tripartite parallel architecture" model of grammar developed in Jackendoff (1997, 2002), which incorporates many insights of CxG (see Jackendoff 2011, 2013). Another West Coast framework which shares some of the core assumptions about the organization of grammar with CxG, HPSG and Simpler Syntax is LFG (Bresnan 2001), but unfortunately, space constraints prevented me from including a discussion of LFG and other alternative theories in this chapter.

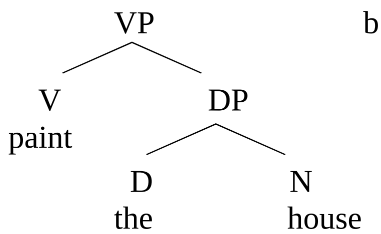
length overviews of competing grammatical frameworks). Given these limitations, I have organized the chapter around five core grammatical themes, discussed in sections 2-6: structure building, nonlocal dependencies, argument structure, linear order, and morphology. For each theme, I first provide a discussion of representative treatments offered by MGG, and then contrast these derivational analyses with constraint-based approaches.

Given my own research expertise, the focus of this chapter is on MGG. But my intention is also to provide readers with a general overview of how MGG differs from alternative, non-transformational frameworks, and to show where there are parallels. I believe that even broad descriptions of these alternatives serve to highlight the distinguishing features of MGG (including some of its problematic aspects) in ways which would not be possible in an overview without a comparative focus.

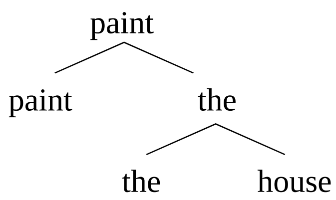
2 Building structure

The Minimalist Program (MP) advocates a *derivational* theory of syntax. The basic syntactic operation is *Merge*, a recursive, binary operation which takes two syntactic objects and combines them to form a new one. The outcome of Merge of α and β can be represented as the set {α, β}.[2] For this newly formed syntactic object to be interpretable at the interfaces with sound and meaning, it needs a *label*, which determines which of the two merged elements *projects*. According to the labeling algorithm of Chomsky (2013), when a head merges with phrase, the head is the label. For example, the syntactic object *paint the house* is derived by merging the verb *paint* and the determiner phrase (DP) *the house*. The verb is the label of the newly derived object and projects, and the resulting object is a VP:[3]

(1) a.



b.



[2] This applies to the merger of heads, complements, and specifiers. The syntactic object created by adjunction of α to β is not a set, but an ordered pair <α, β> (Chomsky 2004).

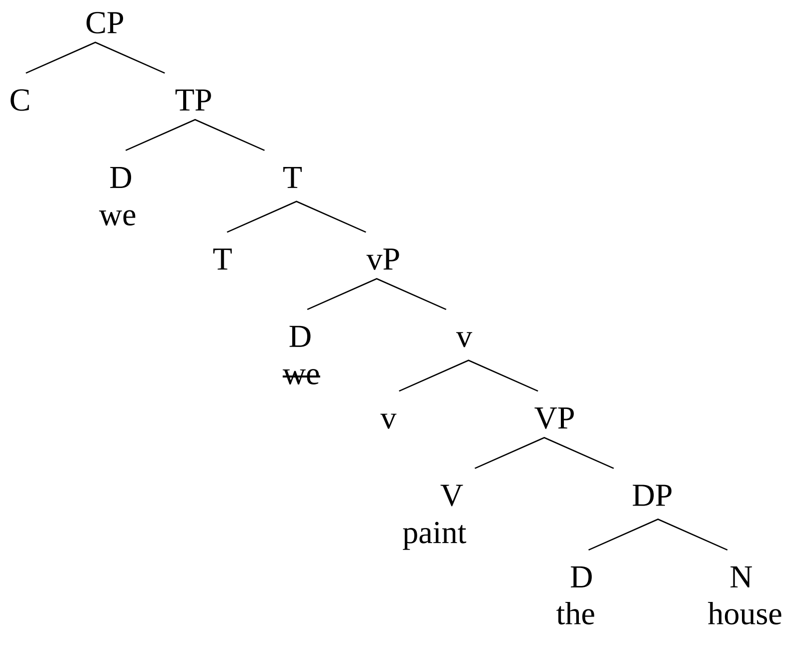
[3] The structures in (1) raise the question of how the head/label of the DP is determined, given that both D and N are terminal nodes. The answer provided in Chomsky (2013) is that N is in fact a phrase consisting of a root and a nominalizing head, as assumed in Distributed Morphology (see section 6).

(1a) shows the more traditional MGG tree notation, which is still commonly used in most studies, but minimalist considerations about structure building are more faithfully captured by the tree diagram in (1b), which is the "Bare Phrase Structure" (BPS) representation of the VP *paint the house*. According to BPS, there are no syntactic nodes and no projections independent of the lexical items that combine via Merge (Richards 2015: 813). A complement is the syntactic object that is first merged with a head; further applications of Merge may produce specifiers (potentially more than one), but X-bar theoretical principles have been abandoned in the MP. Since every complex syntactic object is the outcome of binary Merge, there are no unary (non-branching) or n-ary (with n > 2) projections in BPS.

Two types of Merge are distinguished, depending on whether an element α that is merged with β originates inside β (*internal* Merge) or whether it is taken from an external workspace or the lexicon (*external* Merge). Internal Merge corresponds to movement in earlier forms of MGG (see section 3). When α is a part of β and then re-merged at the root of β, it leaves an identical copy inside β (the copy theory of movement).[4]

(2) shows the outcome of the derivation of the transitive sentence *we paint the house*:

(2)



The syntax of this sentence is derived bottom-up in a step-by-step fashion via successive applications of external and internal Merge. The VP in (2) is externally merged with a light verb v (also known as "little v"), which projects vP and which hosts the external argument in its specifier ([Spec, v]) (the "VP-internal-subject hypothesis"); see section 4. The complex vP

[4] A third type of Merge, called *parallel* Merge, is proposed in Citko (2005), and briefly discussed in section 3.

is merged with the functional head T (tense), and in a next step, the subject pronoun is internally merged with TP, i.e. the subject moves from [Spec, v] to [Spec, T]. To complete the derivation, TP merges with a functional head C (complementizer), which projects a CP.

Not shown in (2) is *head movement* of the verb to v. The status of head movement in MGG is controversial, and it is sometimes suggested that it may not be part of narrow syntax (e.g. Boeckx & Stjepanovic 2001; Chomsky 1995, 2001). However, Matushansky (2006) and Roberts (2010) provide evidence that head movement is a syntactic operation; this is also the assumption adopted in the DM-framework discussed in section 6 below.

The representation in (2) includes various abstract elements, i.e. material with no phonetic content. The functional heads v, T and C in (2) are not pronounced, and neither is the copy of the moved subject (the silent copy is represented by ~~strikethrough~~). The postulation of invisible material is a prominent characteristic of the MGG, especially in so-called "cartographic" approaches, which assume large amounts of mostly abstract functional structure (cf. Cinque & Rizzi 2010).

According to the MP, derivations proceed in cycles, with each cycle corresponding to a subsection of the derivation called a *phase* (Chomsky 2000, 2001, 2008). At specific points in the derivation, the syntactic object generated by Merge is transferred to the interfaces with the sensory-motor and the conceptual-intentional system (the PF- and LF components) for interpretation. Dedicated functional heads are classified as phase heads; when they are merged, their complements undergo the operation *Transfer* and are "evaluated by the interfaces" (Richards 2015: 829). Material that has undergone Transfer is therefore no longer accessible for operations triggered by material outside the phase (the *Phase Impenetrability Condition* PIC; Chomsky 2000, 2001). According to Chomsky (2000, 2001, 2007), little v, C and possibly D are phase heads (but see Bošković (2014) and the references cited therein for a more dynamic view of phasehood; see also section 6).

The derivational, or "proof-theoretic", model of the MP stands in sharp contrast to the alternative computational formalism of *constraint-based* grammars, which is characteristic of monostratal grammatical frameworks such as Lexical-Functional Grammar (LFG; Bresnan 2001), Head-Driven Phrase Structure Grammar (HPSG; Pollard & Sag 1994), Construction Grammar (CxG; Fillmore 1988, Fillmore & Kay 1993; Goldberg 1995, 2006), or Simpler Syntax (Culicover & Jackendoff 2005; Jackendoff 2011). In these frameworks, structures are not derived by grammatical rules, but *licensed* by static grammatical constraints. A linguistic structure is well-formed if all applicable constraints are satisfied.

In HPSG, for example, constraints are expressed through *feature descriptions*, which specify the properties of linguistic objects in the form of attribute-value matrices (AVM). HPSG is a lexical theory, that is, most constraints are formulated as feature descriptions in the lexical entries of words and roots. The lexical entry of a head provides information about its phonological form, its meaning and its syntactic features, including detailed descriptions of the arguments with which the head can combine. For example, the abbreviated AVM in (3) represents (part of) the lexical entry of the finite verb *sees* (Pollard & Sag 1994: 29):

(3) Feature description for the verb *sees*:

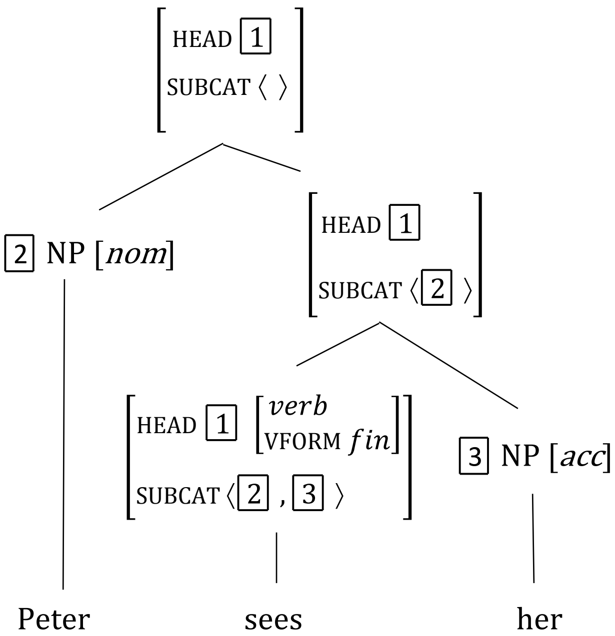
$$\textit{sees} \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb}[\textit{fin}] \\ \text{SUBCAT} & \langle \text{NP}[\textit{nom}]_{\boxed{1}[\textit{3rd, sing}]}, \text{NP}[\textit{acc}]_{\boxed{2}} \rangle \end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix} \text{RELN} & \textit{see} \\ \text{SEER} & \boxed{1} \\ \text{SEEN} & \boxed{2} \end{bmatrix} \end{bmatrix}$$

The syntactic properties of the head *sees* (its category and its subcategorization frame) are expressed as values of the feature CAT(EGORY); semantic information is expressed under CONTENT.[5] The semantic roles assigned by the verb are identified with its two syntactic arguments via *structure sharing*, represented by the boxed number tags.

If the feature description of an item is compatible with the feature description of an element in the valence list of a head, the two linguistic objects can be *unified* to form a phrase. The properties of the phrase are determined by the feature descriptions of the head and its argument, in combination with two additional constraints, the *Head Feature Principle* and the *Head-Argument Schema*. The Head Feature Principle (HFP) ensures that the properties of a head *project*, by requiring that the HEAD-value of any headed phrase be identical to (= structure-shared with) the HEAD-value of the head daughter (see Pollard & Sag 1994: 34; S. Müller 2016: 266). The Head-Argument Schema licenses a phrase whose nonhead daughter satisfies a subcategorization requirement of its head and which inherits the head's remaining valence list (S. Müller 2016: 263). (4) illustrates how these constraints operate in the analysis of the sentence *Peter sees her*:

[5] The simplified feature descriptions in (3) and (4) use the SUBCAT-feature of early versions of HPSG. In most current versions, the SUBCAT feature is split into three separate valence features, namely COMPLEMENTS (COMPS), SPECIFIER (SPR) and SUBJECT (SUBJ) (see S. Müller 2015; Pollard & Sag 1994).

(4) HPSG-analysis for *Peter sees her*:



Since *her* is accusative, it can be unified with *sees* in (3), and its feature description is structure-shared with the first argument from the verb's SUBCAT-list. Because of the Head-Argument Schema, the resulting phrase inherits the verb's valence list minus the argument saturated by the accusative NP. In addition, the HFP ensures that the head features of the phrase are structure-shared with the head features of the verb, making it a verb phrase. When the verb phrase is combined with the nominative NP, the result is a clause (a fully saturated verbal projection).

As in MGG, tree diagrams in HPSG visualize dominance relations and projection. However, (4) includes fewer nodes than the Minimalist representation in (2) above; in HPSG, the sentence is a projection of the verb, not of a functional head like C, as in the MP. Note that the tree in (4) is binary branching, because the Head-Argument Schema can be formulated in such a way that it licenses phrases with exactly one head daughter and one nonhead daughter (S. Müller 2016: 263). However, binary branching is not the only possibility in HPSG, and analyses with flat structures exist as well.

According to Construction Grammar (CxG), the basic units of language are *constructions*, which are defined as conventionalized, lexically listed associations between form and meaning. Therefore, all linguistic structures are licensed through constructional schemata in CxG. For example, Kay & Fillmore (1999: 7) postulate a *Head plus Complements*

Construction (HC), a fully schematic constraint which licenses the combination of a lexical head and one or more local "fillers" (i.e. non-extracted phrasal arguments):

(5) The Head plus Complements Construction (HC)

role head lex +	role filler loc +	+

The constructional schema in (5) is a general constraint on all types of phrasal head-argument constructs. More specific phrasal constructions inherit the general properties of the HC. For example, the VP-Construction only lists those specific properties of a VP which are not already specified by (5), i.e. that its head is verbal and that none of its daughters has the grammatical function of subject (Kay & Fillmore 1999: 8):

(6) The VP-Construction

INHERIT HC
cat v

	gf ¬subj	+

(5) and (6) illustrate that schemata in CxG are organized in *inheritance hierarchies*, in which more general schemata dominate more specific ones. Since constructions lower in the hierarchy inherit all the properties of schemata from a higher level, the system can express generalizations about shared properties of constructions. Properties inherited from higher schemata are predictable, and hence redundant; only the lexical information specific to the form and meaning of the lower schema produces additional informational cost. Inheritance hierarchies are also used in HPSG, to express generalizations across word classes or types of phrasal structures.

It is important to emphasize the ontological difference between the structure-building rule Merge of the MP and the licensing schemata and unification operations of constraint-based frameworks such as HPSG and CxG. In the MP, there is a categorical distinction between grammar and the lexicon. Grammatical rules such as Merge *apply* to lexical items such as words, roots, and functional heads. In contrast, HPSG's feature descriptions and CxG's constructions model both lexical items and phrasal structures, so the boundary between grammar and the lexicon is more fluid. Work in CxG often emphasizes that general constructional schemata such as (5) and (6) represent one endpoint of a continuum of linguistic phenomena with varying degrees of specificity (see e.g. Fried 2015; Goldberg 1995; see also Culicover & Jackendoff 2005). On the other end of the spectrum are entirely idiosyncratic constructions, i.e. words and phrasal idioms such as *kick the bucket*. Between the schematic constructions and fully specified lexical items are linguistic expressions that are neither entirely regular nor entirely fixed, for example phrasal constructions that contribute special meaning beyond the meaning of their parts (e.g. resultative or ditransitive constructions; see section 4), or "constructional idioms", which are lexical schemata with some fixed and some open parts, such as the [V X*'s way* PP]-construction (e.g. *he elbowed his way to the front*). But regardless of the degree of specificity, all constructions are lexically stored; there is no strict separation between the lexicon and grammar.

The nature of Merge and the labeling algorithm of Chomsky (2013, 2015) imply that all phrases are projections of heads/labels. However, some linguistic objects seem to behave differently. One example is the so-called NPN-construction (*day by day*, *picture after picture*, *face to face* etc.), where it is not clear which of its constituent parts (if any) would count as the head (Jackendoff 2008). Another example are coordinate structures, which inherit the categorial features of the conjuncts, and not of the conjunction (but see Chomsky (2013) for an MP-analysis of coordination within the theory of labeling; see S. Müller (2013) for arguments against Chomsky's analysis). Constraint-based frameworks can straightforwardly account for these cases, since their exceptional properties can be captured by allowing unheaded structures to be licensed by specific constraints (see e.g. Jackendoff's (2008) analysis of NPN-constructions as constructional idioms, or Pollard & Sag's (1994: 202) Coordination Principle, which explicitly stipulates that the category of a coordinate structure is determined by its conjunct daughters).

3 Nonlocal dependencies

In the MP, as in HPSG or CxG, linguistic objects are modelled as bundles of grammatical features. The most basic feature of a lexical item is its *edge feature* (EF), which indicates that it can be merged with another syntactic object (Chomsky 2007). Phenomena such as agreement and concord are analyzed in terms of the syntactic operation *Agree* (Chomsky 2000, 2001), which applies when the uninterpretable grammatical features of a functional head F (the *Probe*) find matching valued features (the *Goal*) of another functional category in F's c-command domain. Agree ensures that the feature values of the Goal are transferred to the Probe. When an uninterpretable feature is valued, it can be erased before it reaches the LF-interface, avoiding a "crash". Agree is subject to *Locality*, which stipulates that a probing head must agree with the *closest* Goal in its domain (with closeness defined in terms of c-command).

In the MP, nonlocal (unbounded; "long distance") dependencies are analysed in terms of movement, i.e. internal Merge. Internal Merge is also feature-driven and (in earlier Minimalist studies) linked to the Agree-operation. For example, the prevailing analysis of movement of the subject-DP from [Spec, v] to [Spec, T] (shown in (2) in section 2) postulates a so-called EPP-feature on T, which forces the projection of a specifier. The EPP-feature is associated with the probing agreement features of T (so-called φ-features), and attracts the Goal-bearing element that agrees with T (i.e. the subject-DP). Similarly, in earlier Minimalist studies, movement of wh-phrases to [Spec, C] in constituent questions was assumed to be triggered by an EPP-feature linked to an uninterpretable wh-agreement feature of C (see e.g. Chomsky 1995, 2001; Pesetsky & Torrego 2001). Since Agree is subject to Locality, Agree-based movement accounts can explain island effects such as the superiority violation in (8b):

(7) a. Who read what?
    b. *What did who read?

According to an Agree-based analysis, wh-movement is a consequence of agreement between a Probe in C and the wh-phrase. Since C's feature must agree with the feature of the closest wh-phrase in C's domain (Locality), it follows that only the subject wh-phrase in (7) can be attracted and moved to [Spec, C] (see Chomsky 2000: 128).

However, more recent MP-analyses do not necessarily view internal Merge as a correlate of agreement between a feature of the moved element and an uninterpretable feature of the target. According to Chomsky (2007, 2008), A-bar movement operations such as topicalization (*This book, I like*), which have information-structural effects, are solely triggered by EFs of phase heads. The EF of C can therefore attract any phrase in its domain, and the particular interpretation is established *as a result of* movement. While it is not clear if this revised account can still explain superiority effects of the sort shown in (7b), it is an attempt to find the motivation for internal Merge solely in the conditions imposed by the LF-interface.

Furthermore, Chomsky (2013, 2015) offers an analysis of subject movement to [Spec, T] and an explanation for the "mysterious property EPP" (Chomsky 2008: 156) in terms of labeling. As explained in section 2, the label of a syntactic object derived by Merge determines which of the two merged elements projects; when one of the merged objects is a head, the head is the label of the newly derived object. However, when the subject-DP merges with a projection of little v, the resulting set {DP, vP} cannot be labeled, because neither of its two members is a head. To solve this problem, Chomsky (2013) suggests that, when the subject moves to [Spec, T], its lower copy is no longer visible for the labeling algorithm, and vP can therefore be identified as the label of {DP, vP}. At the same time, Chomsky (2015) argues that T in English is too weak to act as a label. However, since T agrees with the subject, movement of the subject to [Spec, T] establishes the shared φ-features of T and the subject as the label of the set {DP, TP}. Labeling hence provides an independent motivation for why a subject-DP must raise to [Spec, T], which does not require the postulation of an EPP-feature.

As noted in section 2, derivations proceed in phases. When a phase head is merged, its complement undergoes Transfer, and becomes syntactically opaque (the PIC). Consequently, it should be impossible for a phrase that originates inside the complement of a phase head to be internally merged in a position outside the phase head's domain. However, wh-phrases can clearly be extracted across several phase heads. In (8a), *which book* has moved to [Spec, C] across the phase head v; in (8b), *what* has crossed embedded v and C as well as matrix v:

(8) a. Which book will you buy ~~which book~~?
  b. What did you say Peter likes ~~what~~?

The standard minimalist analysis of these long distance dependencies therefore postulates that a wh-phrase moves *successive-cyclically* through the specifier position of every phase head (the *edge* of the phase) that intervenes between the base position of the wh-phrase and its surface position. The derivation of (8b) is shown schematically in (9):

(9) [$_{CP}$ *what* [did you [$_{vP}$ ~~what~~ [say [$_{CP}$ ~~what~~ [Peter [$_{vP}$ ~~what~~ [likes ~~what~~]]]]]]]]?

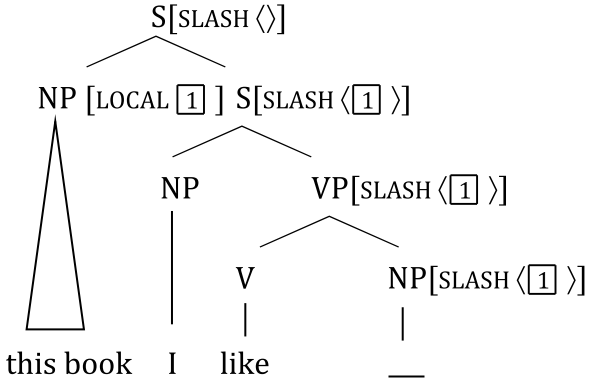
Each movement step takes the wh-phrase to the edge of the next higher phase head, where it remains accessible when the phase head's complement is transferred. Data from many different languages support the view that information about a nonlocal dependency is encoded locally at intermediate points along the dependency path (see e.g. Boeckx 2008 for a summary of the empirical evidence).

In a constraint-based generative grammar such as HPSG, nonlocal dependencies are not analyzed in terms of movement transformations. Instead, the HPSG-analysis of topicalization constructions such as *This book I like* in Pollard & Sag (1994) assumes that the missing object of *like* is represented by a special lexical item (called a *trace* or a *gap*) the purpose of which is to record information about the selectional conditions imposed locally by the verb. (11) shows the (simplified) feature description of the phonologically empty trace (Pollard & Sag 1994: 161):

(10) $$\begin{bmatrix} \text{PHON} \ \langle\rangle \\ \text{SYNSEM} \begin{bmatrix} \text{LOCAL} \ \boxed{1} \\ \text{NONLOC} \mid \text{INHERITED} \mid \text{SLASH} \ \langle \boxed{1} \rangle \end{bmatrix} \end{bmatrix}$$

The feature descriptions of lexical items in HPSG distinguish between information that is locally relevant (LOCAL) and information that plays a role in nonlocal dependencies (NONLOCAL). The LOCAL value of the trace is initially unconstrained, but when a verb is unified with the trace, the feature descriptions from the verb's SUBCAT-list (cf. the entry for *sees* in (3) in section 2) are imposed onto the LOCAL-feature of the trace. As (10) shows, the LOCAL-value of the trace is identified via structure-sharing with the value of its so-called SLASH-feature, which appears under the NONLOCAL path in (10). The *Nonlocal Feature Principle* determines that the nonlocal features of daughter nodes are inherited by the mother node (Pollard & Sag 1994: 162). Therefore, the value of the SLASH-feature, which encodes the selectional constraints specified by the verb, is passed up in the tree to the node that combines with the topicalized NP (the *filler*):

(11)



The SLASH-value is eventually bound off by combining the finite S-node with the topicalized filler NP. For this, HPSG postulates the *Head-Filler Schema* (Pollard & Sag 1994: 162), which licenses the combination of the filler and the finite S-node containing the trace. The Head-Filler Schema identifies the local feature description of the filler with the value of the SLASH feature of the S-node via structure-sharing. Since the SLASH-value is also identical to the LOCAL-value of the trace, the LOCAL-values of the filler and the trace are also identical with each other, and the verb's selectional conditions can apply to the topicalized NP as if the two had combined locally. The trace-based HPSG-analysis therefore produces the same outcome as the cyclic movement (internal Merge) analysis of the MP, albeit with a considerable amount of formal machinery.[6]

As in HPSG, CxG-analyses of nonlocal dependencies (including those proposed in Culicover & Jackendoff's 2005 Simpler Syntax) are non-transformational. Since there are no mechanisms in CxG that can derive one construction from another (Fried 2015: 983), a particular instance of a nonlocal dependency is licensed by the combination of different constructions (Goldberg 2006, 2013; Sag 2010[7]). For example, for a wh-question such as (8a) to be licensed, it must simultaneously satisfy a number of constraints captured by, among others, the VP- and the NP-construction, the subject-predicate construction, the subject-auxiliary inversion construction and, importantly, an unbounded dependency construction which links the wh-phrase to a gap in the argument position of the verb. (Examples of unbounded dependency constructions are Sag's (2010: 518) *Nonsubject Wh-interrogative Construction*, or Kay & Fillmore's (1999: 16) more general *Left Isolation Construction*.) Similarly, in Culicover & Jackendoff's (2005: 310) analysis of wh-movement, the position of the wh-phrase is licensed by a wh-question construction which links a question operator and a bound variable in the semantic representation to a syntactic chain consisting of the wh-phrase and a trace. Non-subject questions are well-formed if they simultaneously satisfy the wh-question construction and the subject-auxiliary inversion construction, which specifies that

---

[6] In the discussion, and in the representations in (11) and (12), I have ignored additional technical details of the HPSG-analysis. Most importantly, in order to prevent the SLASH feature from continuing to percolate after the filler has been added, it is necessary to add another nonlocal feature description (TO-BIND) as a value of NONLOCAL. For details of the analysis, as well as definitions of the Nonlocal Feature Principle and the Head-Filler Schema, see Pollard & Sag (1994) and S. Müller (2015, 2016). Note that HPSG-analyses without traces also exist (see S. Müller (2015, 2016) for references).

[7] Sag's (2010) proposal is formulated in Sign-based CxG, which is a version of HPSG.

inversion takes place in specific semantic contexts, including questions (see Culicover & Jackendoff 2005: 173).

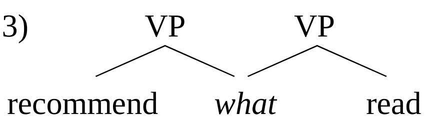
In constraint-based grammars, island effects are typically not analyzed in terms of syntactic Locality principles, but as violations of semantic/pragmatic constraints imposed by the relevant nonlocal dependency constructions, or as the result of processing difficulties (see Goldberg 2006 and Hofmeister & Sag 2010 for discussion). However, Culicover & Jackendoff (2005) assume that both syntactic and semantic constraints are responsible for island effects.

Sag (2010) discusses various nonlocal dependency constructions whose properties cannot easily be explained by transformational accounts based on internal Merge. For example, according to Sag (2010), "Across-the-board" (ATB) wh-constructions such as (12) are problematic for the MGG, because a single extracted wh-phrase corresponds to two gaps:

(12) What did John recommend __ and Mary read __?

However, cases such as (12) have been analyzed in the MP in terms of accounts that allow for *multidominance* relations to be established in the syntax. Citko (2005) proposes a third type of Merge, called *parallel Merge*, which allows for an external element β to merge with an element γ which is internal to α. To explain an ATB-construction such as (12), Citko assumes that the verb *read* first merges with the wh-phrase to form a VP (α). Subsequently, the verb *recommend* (β) also merges with the wh-phrase (γ):

(13)



Each of the VPs in (13) undergoes further merge operations and will eventually be dominated by its own TP. Both TPs are then coordinated, and the coordinated structure merges with C. When the wh-phrase is moved to [Spec, C], the sentence can be linearized as in (12) (see Citko 2005 for details).

4 Argument Structure

In the MGG, argument structure information has traditionally been encoded in the lexical entries of theta role assigners, with the projection of arguments in the syntax established by general linking rules (Jackendoff 1990; Levin & Rappaport-Hovav 1995), or by mapping principles such as Baker's (1988) *Uniformity of Theta Assignment Hypothesis* (UTAH). According to these accounts, a lexical category such as the verb specifies information about the number and syntactic category of its arguments, and the syntax has to be built in a way that reflects these properties (e.g. the UTAH requires the thematic hierarchy to be mirrored by the structural relationships between arguments in the syntax).

In the MP, arguments can also be introduced into the syntax by functional categories. A by now generally accepted view is that the external argument is not an argument of the verb, but of the functional head v (or Voice) (Chomsky 1995; Hale & Keyser 1993; Kratzer 1996), and first merged in the syntax as v's specifier (see section 2). Pylkkänen (2008) argues that the source/recipient argument of double-object constructions and the causer argument of causatives are also introduced by functional heads (Appl and Cause respectively), which are overtly realized by applicative or causative morphology in many languages.

The most radical version of the view that argument structure is determined by functional categories is Borer's (2005a, b, 2013) *exo-skeletal* approach. According to Borer, lexical items do not have any grammatical properties,[8] and do not determine the syntactic frame in which they appear. Instead, all syntactic arguments are introduced by functional structure. This means that a lexical item can in principle be combined with many different argument structures; the acceptability of a particular combination depends on extra-linguistic factors such as world knowledge and convention.

Another approach which associates the projection of syntactic arguments directly with functional structure is Ramchand's (2008) "first-phase syntax". In Ramchand's theory, basic event-structural relations are represented by syntactic structures freely formed by Merge. Ramchand proposes three separate functional projections, of which each denotes a core component of event structure: *res* (for result state); *proc* (for process) and *init* (for initiation). Each of these three categories can host a DP in its specifier. These DPs correspond to the

[8] In Borer's analysis, lexical items are inherently unspecified for syntactic category and are categorized by the functional environment in which they occur. This assumption is shared by the Distributed Morphology-framework (e.g. Marantz 1997, 2001), which I discuss in section 6.

syntactic arguments which function as the subjects of the predicational relations expressed by *res*, *proc* and *init* (subject of result = resultee; subject of process = undergoer; subject of initiation = initiator). Different types of event structures are represented by different combinations of the three basic event projections. In contrast to Borer's theory, lexical items do not combine freely with these argument projections in Ramchand's (2008) model. Rather, verbs are lexically specified for exactly which of the three core eventive projections they license.

In sharp contrast to the generative "neo-constructionist" theories of Borer and Ramchand, HPSG (much like traditional MGG) lists all information about argument structure directly in the lexical entries of verbs and other theta role assigners. Argument roles are specified in the semantic feature description of the verb, and linking is achieved via structure-sharing between these roles and the values of the SUBCAT-feature, in combination with the Head-Argument Schema (see section 2).[9] Generalizations about the syntactic realization of arguments can be stated in terms of inheritance hierarchies, linking principles, and constraints which distinguish different types of verbs on the basis of the thematic roles they select (see e.g. Davis & Koenig 2000).

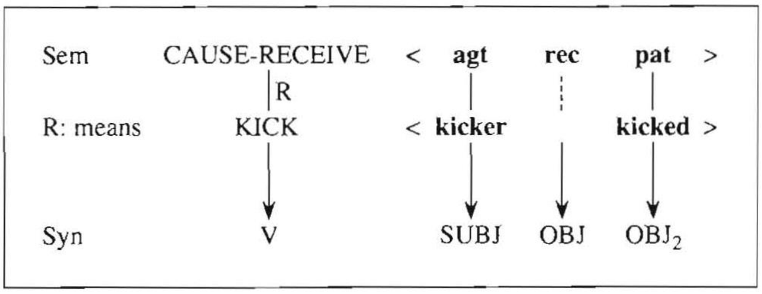
In CxG, argument structure is not projected from the lexical properties of verbs, but is a property of lexically listed phrasal *argument structure constructions* (Goldberg 1995, 2013). A key observation that motivates this view is that one and the same verb can appear in many different argument structure constructions. The following example, from Goldberg (1995: 11), illustrates this with the verb *kick*:

(14) a. Pat kicked the wall.
b. Pat kicked Bob black and blue.
c. Pat kicked the football into the stadium.
d. Pat kicked at the football.
e. Pat kicked his foot against the chair.
f. Pat kicked Bob the football.
g. The horse kicks.
h. Pat kicked his way out of the operating room.

[9] In most contemporary HPSG-analyses, which use the valency features SPR, COMPS and SUBJ, the verb's feature description includes an additional feature description ARG-ST (argument structure), which is required for the analysis of Binding phenomena. See S. Müller (2015).

According to CxG, the different argument structures in (14), and the respective special meanings of these sentences, are attributed to the phrasal constructions in which the verb in (14) appears. For example, (14b) is an example of the resultative construction, and (14h) illustrates the *way*-construction (a constructional idiom already mentioned in section 2). A verb is integrated into (unified with) a construction by fusing its core participant roles (the roles implicitly understood in the conceptualization of the verbal event) with the arguments specified by the construction. Importantly, it is possible that some participant roles remain unexpressed after unification, or that a construction adds arguments not specified by the verb's meaning. For example, the verb *kick* specifies two participants (the kicker and the kicked), but the ditransitive construction introduces three arguments. When the verb and the ditransitive construction are unified, a recipient argument is contributed by the construction (Goldberg 1995: 54):

(15) Ditransitive construction + *kick*



Argument structure-constructions in CxG are polysemous; there is typically a family of closely-related meanings associated with the same syntactic form. For example, Goldberg (1995) distinguishes a number of different senses of the ditransitive construction, which are all related to the "central sense" of this construction, defined as "Agent successfully causes Recipient to receive Patient". The central sense is realized only with a small class of verbs such as *give*; other verbs appear with different senses. For example, in *He baked John a cake*, the cake is only intended to be received by John; in *He allowed John an ice-cream*, the transfer is enabled, not caused; in *He refused John an ice-cream*, the transfer is negated etc.

The view advocated by CxG – that structures contribute meaning, and that the projection of arguments is a property of syntactic constructions – is in some respects comparable to the neo-constructionist theories of Borer (2005a, b, 2013) and Ramchand (2008). However, the fundamental difference between CxG and the latter theories is that constructions in CxG are lexical items, memorized phrasal schemata that are stored in the mental lexicon. In contrast, in Borer's and Ramchand's approaches, argument structure syntax is systematically constructed by Merge from a limited inventory of functional heads. Ramchand (2008: 11) therefore calls her theory "generative-constructivist".

Culicover & Jackendoff (2005) adopt a mixture of these various approaches to argument structure. They assume that the default argument structure of a sentence is licensed by the verb, but they also allow syntax-semantics interface rules to license additional arguments via constructions and constructional idioms in the spirit of CxG.

5 Linearization

In standard MGG, linear order is established derivationally via Merge. If an element α is merged to the left of an element β, it precedes it; otherwise α follows β. These are the only two options in a binary branching syntax. The head directionality parameter is instantiated by specifying on which side a particular head selects its complement.

In constraint-based grammatical frameworks like Simpler Syntax or HPSG, constraints that govern linear order are distinguished from constraints that build structures. For example, in their sketch of English phrase structure in Simpler Syntax, Culicover & Jackendoff (2005: 145) separate constituency rules from word order rules. Similarly, HPSG postulates linear precedence constraints, which determine the order of heads with respect to their complement(s). In addition, HPSG can account for scrambling languages by formulating the *Head-Argument Schema* in such a way that the nonhead daughter can be any arbitrary element from the head's valency list. This licenses free constituent order (see S. Müller 2016: 287).

To the best of my knowledge, only MGG has put forward explicit accounts to *derive* linear order from hierarchical structure. The landmark proposal is Kayne's (1994) *Linear Correspondence Axiom*, which stipulates that there is a direct 1:1 mapping between asymmetric c-command relations and linear precedence:[10]

[10] (16) differs slightly from the version of the LCA given in Kayne (1994), because Kayne's original LCA was formulated using X-bar-theoretical notions and required unary branching. It is therefore incompatible with the Bare Phrase Structure theory of the MP.

(16) *Linear Correspondence Axiom (LCA)*

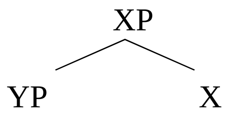
(Biberauer, Holmberg & Roberts 2014: 207)

α precedes β if and only if α asymmetrically c-commands β or if α is contained in γ, where γ asymmetrically c-commands β.

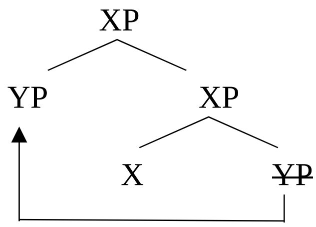
The key idea behind linearization theories is that syntax is only concerned with the hierarchical relations established by Merge. The linear order of the elements is established at the PF-interface, in the mapping to the sensory-motor system, via the algorithm in (16).

According to the LCA, the c-command relations between heads, complements and specifiers imply that specifiers always precede heads, and heads precede (material contained in) their complements. This means that head-final constructions, in which a head X follows its complement YP, cannot have the syntax in (17a), but must be derived via YP-movement to a higher position, as e.g. in (17b):

(17) a.



b.



Since X c-commands YP in (17a), the tree in (17a) would still be mapped onto a linear order where X precedes material dominated by YP – the fact that X is represented to the right of YP in the tree diagram is irrelevant. However, movement of YP to a position where YP asymmetrically c-commands X as in (17b) results in a head-final order YP < X.

A problem with the LCA and other linearization theories is that, in order to derive attested word orders from syntax, proponents of these theories often have to assume otherwise unmotivated movement operations of the sort shown in (17b). Moreover, empirical problems for the LCA are raised by constructions in which an element X that follows Y can nevertheless be shown to c-command Y. On the plus side, apart from its conceptual appeal, the LCA has also offered a principled explanation for a variety of linguistic phenomena and generalizations. For example, Biberauer, Holmberg & Roberts (2014) provide an LCA-based account of the so-called *Final-Over-Final Constraint* (FOFC), which they state informally as in (18) (see also Sheehan et al. 2017):[11]

(18) *The Final-over-Final Constraint (FOFC)*
(Biberauer, Holmberg & Roberts 2014: 171)

A head-final phrase αP cannot dominate a head-initial phrase βP, where α and β are heads in the same extended projection.

The FOFC captures the empirical observation that a head α can only follow its complement βP if βP is also head-final. While one finds "mixed-headedness" in extended projections in which a head precedes a head-final complement, what does not seem to be attested in natural language are constructions in which a head follows a head-initial complement. To cite just one example, Holmberg (2000) shows that Finnish allows for three possible word orders with auxiliaries, verb and object: Aux-V-O (both AuxP and VP are head-initial), O-V-Aux (both AuxP and VP are head-final), and Aux-O-V (AuxP is head-initial, VP is head-final). However, the fourth theoretically possible order, V-O-Aux, is not attested and ruled out by the FOFC. Biberauer, Holmberg & Roberts (2014) explain the FOFC in terms of the LCA. If head final orders are derived by internal Merge (as in (17b)), and if internal Merge is triggered by movement features (EPP-features or EFs; see section 3), then it follows that the head of a head-final phrase must have such a feature. Biberauer, Holmberg & Roberts (2014) argue that the relevant feature always originates on the lowest head in an extended projection from where it can percolate upwards. It follows that, if a higher head in an extended projection has the feature, then every lower head must have it too. This derives the FOFC.

6 Morphology

The morphological framework most closely connected with the assumptions and mechanisms of the MP is *Distributed Morphology* (DM) (see e.g. Bobaljik 2012, 2015; Embick 2010; Halle & Marantz 1993; Harley 2014; Marantz 1997, 2001, 2007, among many others). According to DM, morphology is "distributed" across different places in the architecture of grammar; it comprises both syntactic and post-syntactic operations. While the structural

[11] Note that in Sheehan et al. (2017), the Final-over-Final *Constraint* has become the Final-over-Final *Condition*.

aspects of morphology are determined by the same generative principles and mechanisms that also drive syntactic derivations (morphology is "syntax all the way down"), the association of phonological form and semantic interpretation with the terminal nodes of syntax happens after Transfer in the interpretative components of grammar, i.e. at the PF- and LF-interfaces.

DM distinguishes three "lists" (Halle & Marantz 1993). List 1 comprises the *morphemes*, the terminal nodes of syntax. Morphemes are combined into larger structures (complex heads and phrases) by Merge and (head) movement. List 2 in DM is the Vocabulary. This list includes the phonological exponents (*vocabulary items*) that are associated with the terminal nodes of a syntactic representation in the mapping to PF via the process of Vocabulary Insertion (VI) ("late insertion"). The insertion of a particular vocabulary item is licensed by the grammatical feature specification of the morphemes; like A-morphous Morphology (Anderson 1992), Paradigm Function Morphology (Stump 2001), or Construction Morphology (Booij 2010a, b), DM is a *realizational* morphological theory. Finally, List 3 is the encyclopedia, the list of "special meanings", which are associated with terminal nodes in the mapping to LF. The semantics of a sentence is computed by combining these special meanings with the meaning of the interpretable features associated with functional elements.

List 1 includes two types of morphemes: roots (symbolized by √; Pesetsky 1995), and functional categories. Roots are "pure units of structural computation" (Harley 2014: 226), i.e. category-neutral elements that appear initially without semantic or phonological properties[12]; functional heads are bundles of grammatical features relevant for syntactic operations (see section 3). The inventory of functional morphemes includes the core functional categories D, T and C discussed in section 2, but also heads such as "little" v, a and n, which may be phonologically zero or overtly realized by derivational affixes, and which determine the syntactic category of the a-categorial roots with which they merge. Verbs, adjectives and nouns are therefore syntactically complex; they are created by combining an unspecified root with a category-defining head, e.g.:

(19) a. *work* (verb): [v v √WORK] b. *work* (noun): [n n √WORK]

[12] Following Pfau (2000), Harley (2014) assumes that roots are identified by an abstract index that provides a link between the vocabulary items and encyclopedic entries supplied at PF and LF respectively. De Belder & van Craenenbroek (2015) argue that the syntax generates structurally empty, abstract root positions at the initial stage of the derivation into which specific vocabulary items are inserted post-syntactically.

(19) illustrates that a word such as *work* is not inherently specified as a verb or a noun; it is an a-categorial root which only becomes verbal or nominal in the context of verbalizing or nominalizing functional heads (Marantz 1997; Harley 2014).[13]

In the mapping to PF, the vocabulary items of List 2 are inserted into terminal nodes according to a node's feature specification. A vocabulary item may be underspecified and include fewer features than the target morpheme, but its features must be a subset of the features of the target. Features of a terminal node can in principle be matched by more than one vocabulary item, which means that different matching vocabulary items compete for insertion. This competition is governed by the Subset Principle (Halle 1997), which requires that the most specified vocabulary item that matches the feature specification of the target node is chosen for insertion. An example is the present tense in English, for which there are two vocabulary items: -*s* for third person singular, and the zero morpheme Ø for all other cases. Both exponents are candidates for insertion into a T-node with the feature [present], but if T's φ-features have been valued in the syntax as 3rd Pers Sg by the operation Agree (see section 3), then -*s* must be chosen, because it is more specific. However, if T's φ-features have any other value, then -*s* cannot be inserted, because its feature specification is no longer a subset of T's features, and Ø must be chosen as the elsewhere case (Bobaljik 2015).

Standard DM assumes that the PF-interface includes various morpho-phonological operations that can alter the structural output of syntax before vocabulary items are inserted, or change the order or the phonological form of affixes after VI. For example, the operation *fusion* can combine two terminal nodes into one; *fission* can split a terminal node into two morphemes; *readjustment rules* can change the phonological form of vocabulary items, and *local dislocation* can change the linear order of affixes, after VI (Embick & Noyer 2001; Halle & Marantz 1993). The proliferation of these post-syntactic operations has been criticized in more recent DM-work (see e.g. Haugen & Siddiqi 2016), and some studies within DM, as well as alternative late-insertion theories, assume a model of morphology with Vocabulary Insertion as the only post-syntactic operation in the PF-component. For example, in Nanosyntax theory (Caha 2009; Starke 2009), every grammatical feature corresponds to a "sub-morphemic" terminal node in syntax, and vocabulary items can realize more than one

[13] The idea that roots are a-categorial and "typed" in the context of specific functional structure also underlies Borer's (2005a, b; 2013) exo-skeletal theory (see section 4). However, in contrast to DM, Borer does not assume phonologically empty category-defining heads; instead, roots are typed in the context of inflectional functional categories or derivational affixes.

terminal node at once, a process known as "Spanning" (Merchant 2015; Svenonius 2012; Taraldsen 2010). Phonological (and semantic) interpretations are associated with entire subtrees generated in the syntax, and lexical entries therefore include stored information about the structures they license.

An important insight of DM is that morphological operations at both interfaces are constrained by Locality; i.e. that the association of phonological and semantic material from Lists 2 and 3 with a particular terminal node depends on the node's syntactic environment (Embick & Marantz 2008; Marantz 1997, 2001, 2007, 2013). The classical example of how Locality constrains Vocabulary Insertion, going back to Halle & Marantz (1993) and explored in detail in e.g. Embick (2010) and Bobaljik (2012), is *contextual allomorphy*. For example, the past tense morpheme in English can be spelled out as *-t* (as in *left, spent*), as Ø (as in *hit* or *put*), or as *-ed* in the elsewhere case (Embick & Marantz 2008). The choice between these vocabulary items is determined by the verb that appears in the local context of T[past]. At the same time, T[past] can trigger contextual root allomorphy. For example, the root √LEAVE is realized by the vocabulary item *lef-* in the local domain of a past tense morpheme, at the same time as this root licenses the vocabulary item *-t* to be inserted into T. *Suppletion* is just an "extreme" version of contextual root allomorphy, e.g. *am* is simply the form of √BE in the context of a phonologically null T[present; 1st Pers Sg].

The idea that not only allomorphy, but also allosemy, is constrained by Locality goes back to Marantz's (1984) observation that idioms with a fixed subject and a variable slot for objects are rare in comparison to VP-idioms with fixed objects and open subjects. Since the external argument is introduced by little v (see sections 2 and 4), it was natural to assume that special meanings are restricted to certain syntactically defined domains, such as VP. This idea was then generalized in Marantz (1995, 1997, 2001), who suggests that the local syntactic environment of a morpheme determines its interpretation, with different contexts giving rise to contextual allosemy. To illustrate with an example from Marantz (2013), the root √HOUSE can merge with n to form a noun (*the house*) or with v to form a verb (*to house*), but its meaning in the nominal context is more restricted than its verbal meaning, which does not imply reference to a literal house (you can house someone in a tent). Note that in this example, the different allosemes licensed by the category-defining heads n and v also correspond to different allomorphs of the root; compare *the hou*[*s*]*e*, but: *to hou*[*z*]*e*.

However, the view that late encyclopedic insertion can only target terminal nodes led to a peculiar DM-analysis of phrasal idioms. According to this analysis, the special meaning of a

VP-idiom such as *kick the bucket* is derived compositionally on the basis of the special meanings of its parts, which are established in the syntactic context defined by the idiom. √KICK would mean "to die" in the context of √BUCKET, which in turn receives a null interpretation in the context of √KICK (Bobaljik 2015) (the role of the determiner is not clear in this account). However, in many instances, such an analysis seems unintuitive, and it is doubtful that all idioms can indeed be analyzed in this way (cf. Nunberg, Sag & Wasow 1994; Zeller 2001). In recent work, Marantz (2013: 105) admits that his earlier work was "confusing if not simply wrong in conflating the notion of 'idiom' with the notion of 'special meaning' or 'meaning choice' associated with polysemy", and he accepts that idiomatic meanings can be "built on top of polysemy resolution" (Marantz 2013: 106).

More recent work in DM explores the nature of the local domains that are relevant for insertion of vocabulary items and encyclopedic material. An important hypothesis is that contextual allomorphy and contextual allosemy are restricted to the domain of a *phase*, and that the categorizing functional heads v, n and a can function as phase heads (Arad 2003; Bobaljik 2012; Embick & Marantz 2008; Embick 2010; Marantz 2001, 2007). However, it was noted in section 3 that a phase head triggers transfer of its complement as soon as it is merged (Chomsky 2000, 2001). In order to avoid that merger of the first category-defining head would cause the root to be transferred, Embick (2010) and Marantz (2013) postulate that roots are always spelled out and interpreted in the same domain as the head that "types" them for category.[14] The root and the categorizer are then transferred together as part of the domain of the next cyclic phase head. This domain may also include non-phasal material that projects between the categorized root and the higher phase head. For example, vP, which dominates the verbalized root, will be spelled out as part of the domain of the next phase head C, which also includes the non-phase head T. Since T is in the same spell-out domain as vP, it can determine root allomorphy of verbs, as discussed above.

The categorizing heads n, v and a can also be category-changing and merge with the functional projection corresponding to an already categorized root (that is, another nP, vP or aP). Since n, v and a are phase heads, they will trigger transfer of their complement. In this case, root allomorphy and allosemy can only be determined by the category-*defining* head in the first spell-out domain; the category-*changing* phase head in the next phase has no effect on the form or meaning of a root, as it is not part of this domain. For example, Marantz

[14] One way in which this could be achieved is by assuming obligatory head movement of roots to their categorizing heads (Marantz 2013).

(2013: 103) notes that a verb *to house* can be derived directly from the root √HOUSE (as explained above), but also from the noun *house*, as in "He took a bunch of plastic models and housed the room with them (= filled the room with houses)". In the latter case, the verb *house* is formed by merging v with the nP corresponding to the noun *house*:

(20)
vP
phase head → v ∅ nP → spell-out domain
n √HOUSE

Crucially, the derived verb in (20) preserves the form and the meaning established by the lower category-defining head n; consequently, it is pronounced with the voiceless fricative, and it must include reference to houses.

On the basis of the idea that a categorizing functional head can merge with a bare root or with a categorized nP-, vP- or aP-projection, DM distinguishes two types of derivational morphology. Derivational affixes that directly merge with roots can cause semantic and phonological idiosyncrasies of the root, because they are part of the root's spell-out domain. In contrast, the effects of derivational morphology that attaches outside the functional projection that "types" a root are more regular and predictable (compare (20)). As argued in Marantz (1997, 2001, 2007), this distinction between "inner" and "outer" derivational morphology in DM corresponds to the traditional distinction between "lexical" and "syntactic" word formation. For example, the well-documented differences between adjectival and verbal passives (Wasow 1977), between "lexical" and "syntactic" causatives in Japanese (Kuroda 1965; Miyagawa 1984), or between passives and statives in Bantu (Dubinsky & Simango 1996) follow from the different possible attachment sites of the v-, n- or a-heads whose exponents are the relevant passive, causative etc. affixes (see Harley 2014 for discussion of the "inner" vs. "outer" morphology-hypothesis, which she calls "one of the great insights of the syntacticocentric approach to morphological analysis" (p. 267).

In comparison to derivation and inflection, compounding has received relatively little attention in DM. Harley (2009) suggests that synthetic compounds such as *truck-driver* are derived by merging the root √DRIVE with its argument, the nP *truck* (root √TRUCK plus n). The whole structure is selected by another n-head, the agentive nominalising affix *-er*. The complex terminal node is formed via incorporation of all heads into the highest n-head. "Modificational" compounds such as *quick-acting*, and even root compounds such as *alligator shoes* are derived in the same way (i.e. in the latter case, √SHOE selects the nP *alligator*, and the whole structure is merged with a higher, unpronounced n).

In the remainder of this section I discuss an alternative, constraint-based view of morphology, i.e. Construction Morphology (CM), developed in Booij (2010a, b). CM adopts the general conceptual framework of CxG and regards constructions as the basic units of grammar. In Booij (2010a, b), word formation processes are captured in terms of constructional schemata, which are viewed as abstractions across paradigmatically related words. Complex words are instantiations of these morphological constructions. For example, the lexical entry for the agentive nominalizing suffix *-er* in English is viewed as an abstract generalization over the paradigmatic relationship that characterizes pairs of verbs and nouns, such as *drive* ≈ *driver, read* ≈ *reader* etc. The systematic relationship between these words can be captured by a schema with a variable in the position of the verb, and a fixed element, the affix (Booij 2010a: 8):

(21)
$$\begin{array}{lllll} \omega_i & \leftrightarrow & N_i & \leftrightarrow & [\text{one who PRED}_j]_i \\ [\ \ ]_j[\text{ər}]_k & & V_j \quad Aff_k & & \end{array}$$

The construction in (21) is notated in the format of Jackendoff's (1997, 2002) "tripartite parallel architecture". It expresses the correspondence between phonological, syntactic and semantic information via lexical indices and correspondence arrows. The phonology of the derivational affix is linked to its morpho-syntactic representation as a verbal suffix. The phonological word including the affix is linked to the complex noun in morpho-syntax and to the semantic representation "someone who carries out the action described by the verb". The schema represents a generalization over existing words, but since its verbal part corresponds to an open slot in the phonology and in the semantics, it can be unified with any verb to productively derive new nouns. This means that (21) is a constructional idiom (see section 2).

The representation in (21) shares with DM the idea the lexical entry for an affix must specify its morpho-syntactic context. The crucial difference is that in CM, the affix is only linked to its phonological expression, but not to a semantic representation – the semantics is only specified for the *combination of* verb and affix, which means that the affix has only meaning as part of the schema in (21). This is the constructional aspect of CM. In DM, in contrast, the encyclopedic entry for the affix *-er* is a special meaning, which would be inserted into the terminal node n in the appropriate local context of a vP-complement. The meaning of the agentive noun would then be derived compositionally on the basis of the meanings of the verbalized root and the derivational suffix.

Schemas such as (21) above capture the paradigmatic relation between simple words and corresponding complex words that include the derivational affix. However, paradigmatic relationships can also exist between two morphologically complex words. Booij (2010a) argues that nouns such as *altruist*, *autist* or *pacifist* cannot be derived by attaching the affix -*ist* to a root or stem, since there are no independently existing roots for these derivations. Rather, these complex nouns are systematically related to other morphologically complex nouns, namely *altruism*, *autism* and *pacifism*. Booij (2010a: 33) captures this correlation through a paradigmatic relationship between two word formation schemata, represented by the ≈-symbol:[15]

(22) $< [\text{x-ism}]_{Ni} \leftrightarrow \text{SEM}_i > \; \approx \; < [\text{x-ist}]_{Nj} \leftrightarrow [\text{person with property Y related to SEM}_i]_j >$

In contrast, a DM-analysis would treat words such as *altruism* or *pacifist* as being derived from the roots √ALTRU and √PACIF, which are only licensed in the context of a categorizing little n head if this head is spelled-out as a derivational affix -*ism* or -*ist*.

Affix schemata such as those in (21) and (22) are related to other constructions via the familiar inheritance hierarchies of CxG. For example, the lexical entry in (21) inherits properties from a more abstract schema that specifies general properties of derived nouns. At the same time, word formation schemata dominate the words that instantiate them. E.g. any complex noun formed through the unification of (21) and a verb inherits information from both (21) and the verb (Booij 2010a: 26).

Schemata related through inheritance rules are also used for compounding. For example, Booij (2010b: 4) proposes a schema for nominal compounds in English, which captures the fact that these compounds adhere to the Right-hand Head Rule (Williams 1981):

(23) $< [[a]_{Xk}\ [b]_{Ni}]_{Nj} \quad \leftrightarrow \quad [\text{SEM}_i \text{ with relation R to SEM}_k]_j >$

[15] In (22), each constructional schema is represented inside angular brackets. The phonological and morpho-syntactic information in the schemata has been conflated in (22). SEM represents the arbitrary meaning of the word ending in -*ism*.

The schema specifies that endocentric nominal compounds are based on the category and meaning of their right-hand member, which stands in an arbitrary semantic relation R to the left-hand member. The category of the left-hand member is represented by the variable X, which stands for the major lexical categories. VN, NN, AN and PN compounds are represented by lower-level schemata which inherit the general properties of (23).

Van der Spuy (2017) provides CM-representations for all major types of inflectional morphology. For example, he represents English plural formation with the affix /z/ as in (24) (2017: 64):[16]

(24) $/X_{[N]}\text{-z}/_{[N\ pl]} \leftrightarrow$ SEM

(24) specifies that a noun ending in /z/ is a plural form of that noun and linked to its specific plural semantics. Note that according to (24), a plural noun ending in /z/ is not morpho-syntactically complex; the phonological form ending in /z/ spells out a noun which has the feature [plural], but there is no internal morphological structure. This view, which is characteristic of "Word-and-Paradigm" theories such as Anderson's (1992) *A-morphous Morphology*, is fundamentally different from the DM-analysis, which associates the exponents of inflectional morphology with functional nodes in the syntax.

According to Van der Spuy (2017), inflectional processes involving allomorphy or suppletion are captured in CM in terms of paradigmatic relationships of the sort shown in (22) above. For example, Van der Spuy (2017: 65) provides the following constructional representation for suppletive *am* as the form of [1st Pers Sg] *be*:

(25) $/\text{bi}/_{[V]} \leftrightarrow \text{BE} \approx /\text{æm}/_{[V\ 1sg\ pres]} \leftrightarrow \text{BE}$

Again, in contrast to the DM-analysis of suppletion discussed earlier in this section, (25) does not postulate a separate functional category in the morpho-syntax that would contribute the inflectional features. Instead, (25) simply specifies that the verb *am* is a variant of the verb *be* with the relevant feature specification [1st Pers Sg].

[16] In contrast to Booij (2010a), Van der Spuy (2017) does not use angular brackets to demarcate constructions. He represents phonological forms between slashes, and morpho-syntactic features as subscripts in square brackets. Van der Spuy also omits the corresponding arbitrary semantics, represented by SEM in Booij's (2010a) schemata, in some of his abbreviated representations.

7 Conclusion

In this chapter, I have provided a synopsis of current ideas and developments in the MP and in DM, and I have contrasted the theoretical assumptions and analyses of MGG with syntactic and morphological approaches that are representative of constraint-based theories. The major distinguishing attributes of the grammatical frameworks I have discussed in this chapter are the following: 1. In contrast to constraint-based grammars, MGG assumes a *derivational* syntax; 2. The postulation of *empty structure and elements* is common in MGG, while constraint-based theories are *more surface-oriented*; 3. In contrast to MGG, there is no strict boundary between *the lexicon and grammar* in constraint-based theories. These aspects characterize the different approaches to core syntactic phenomena in the MP, HPSG, CxG and Simpler Syntax, as well as the different treatments of inflectional morphology and word formation processes in DM and CM.

There are other, perhaps more fundamental, differences. For example, while work in HPSG, CxG and Simpler Syntax is usually strongly data-driven, analyses in the MP and DM are guided by the search for "deeper" reasons and principled explanations (the "Galilean style" of scientific inquiry), an approach that has led to the perception among MGG sceptics that conceptual considerations are prioritized in the MP and DM at the expense of empirical accuracy. However, despite the different views about the architecture of grammar, and the fierce debates about what should be the overall goal of linguistic analysis, both derivational and constraint-based approaches have made important contributions to the development of a descriptively adequate and cognitively plausible theory of natural language.

8 References

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Arad, Maya. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language & Linguistic Theory* 21(4): 737–779.

Baker, Mark C. 1988. *Incorporation. A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Biberauer, Teresa, Anders Holmberg and Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2): 169-225.

Bobaljik, Jonathan D. 2012. *Universals in Comparative Morphology: Suppletion, Superlatives, and the Structure of Words*. MIT Press.

Bobaljik, Jonathan D. 2015. Distributed Morphology. Ms., UConn, Storrs.

Boeckx, Cedric. 2008. *Understanding Minimalist Syntax*. Oxford: Blackwell.

Boeckx, Cedric and Sandra Stjepanovic. 2001. Head-ing toward PF. *Linguistic Inquiry* 32: 345-355.

Booij, Geert. 2010a. *Construction Morphology*. Oxford: Oxford University Press.

Booij, Geert E. 2010b. Construction morphology. *Language and Linguistics Compass* 4(7): 543–555.

Borer, Hagit. 2005a. *Structuring sense, Vol. 1: In name only*. Oxford: Oxford University Press.

Borer, Hagit. 2005b. *Structuring sense, Vol. 2: The normal course of events*. Oxford: Oxford University Press.

Borer, Hagit. 2013. *Structuring sense, Vol. 3: Taking form*. Oxford: Oxford University Press.

Bošković, Željko. 2014. Now I'm a Phase, Now I'm Not a Phase: On the Variability of Phases with Extraction and Ellipsis. *Linguistic Inquiry* 45(1): 27-89.

Bresnan, Joan. 2001. *Lexical-functional syntax*. Oxford: Blackwell.

Caha, Pavel. 2009. *The Nanosyntax of Case*. University of Tromsø: Doctoral dissertation.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

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

Chomsky, Noam. 2000. Minimalist Inquiries: the Framework. In Roger Martin, David Michaels and Juan Uriagereka (eds.), Step by Step, pp. 89-155. Cambridge: MIT Press.

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

Chomsky, Noam. 2004. Beyond Explanatory Adequacy. In Adriana Belletti (ed.), *Structures and Beyond – The Cartography of Syntactic Structure*, Vol 3, pp. 104-131. Oxford: Oxford University Press.

Chomsky, Noam. 2007. Approaching UG from Below. In Uli Sauerland and Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language*?, pp. 1-29 Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On Phases. In Freidin, Robert, Carlos Otero and Maria-Luisa Zubzarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, pp. 133-166. Cambridge: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33-49.

Chomsky, Noam. 2015. Problems of Projection: Extensions. In Elisa Di Domenico, Cornelia Hamann and Simona Matteini (eds.), *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, pp. 1-16. Amsterdam: John Benjamins.

Cinque, Guglielmo & Luigi Rizzi. 2010. The cartography of syntactic structures. In Bernd Heine and Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, pp. 51–65. Oxford: Oxford University Press.

Citko, Barbara. 2005. On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge. *Linguistic Inquiry* 36(4): 475-496.

Culicover, Peter W. and Ray S. Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Davis, Anthony R. & Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76(1): 56–91.

De Belder, Marijke & Jeroen van Craenenbroeck. 2015. How to merge a root. *Linguistic Inquiry* 46(4): 625-655.

Dubinsky, Stan and Ron S. Simango. 1996. Passive and stative in Chichewa: Evidence for modular distinctions in grammar. *Language*, 72(4): 749–781.

Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology*. Cambridge: MIT Press.

Embick, David and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4): 555-595.

Embick, David and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39(1): 1-53.

Fillmore, Charles J. 1988. The mechanisms of "Construction Grammar". In Shelley Axmaker, Annie Jaisser & Helen Singmaster (eds.), *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*, pp. 35–55. Berkeley, CA: Berkeley Linguistics Society.

Fillmore, Charles J. and Paul Kay. 1993. *Construction Grammar*. Berkeley: University of California.

Fried, Mirjam. 2015. Construction Grammar. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42 (Handbooks of

Linguistics and Communication Science), pp. 974–1003. Berlin: Mouton de Gruyter (2nd edn.).

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago/London: The University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language* (Oxford Linguistics). Oxford, New York: Oxford University Press.

Goldberg, Adele E. 2013. Argument structure Constructions vs. lexical rules or derivational verb templates. *Mind and Language* 28(4): 435–465.

Hale, Ken and Samuel J. Keyser. 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In Ken Hale and Samuel J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, pp. 53-110. Cambridge: MIT Press.

Halle, Morris, 1997. Distributed Morphology: Impoverishment and Fission. *MIT Working Papers in Linguistics*, 30, pp. 425–449.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale and Samuel J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, pp. 111–176. Cambridge: MIT Press.

Harley, Heidi. 2009. Compounding in Distributed Morphology. In Rochelle Lieber and Pavol Štekauer (eds.), *The Oxford Handbook of Compounding*, pp. 129-144. Oxford: Oxford University Press.

Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40: 225–276.

Haugen, Jason D. and Daniel Siddiqi. 2016. Towards a Restricted Realizational Theory: Multimorphemic monolistemicity, portmanteaux, and post-linearization spanning. In Daniel Siddiqi and Heidi Harley (eds.), *Morphological Metatheory*, pp. 343-385. Amsterdam: John Benjamins.

Hofmeister, Philip & Ivan Sag. 2010. Cognitive constraints and island effects. *Language* 86: 366-415.

Holmberg, Anders. 2000. OV order in Finnish. In Peter Svenonius (ed.) *The derivation of VO and OV*, pp. 123-152. Amsterdam: John Benjamins.

Jackendoff, Ray S. 1990. *Semantic Structures*. Cambridge: MIT Press.

Jackendoff, Ray S. 1997. *The Architecture of the Language Faculty*. Cambridge: MIT Press.

Jackendoff, Ray S. 2002. *Foundations of language*. Oxford: Oxford University Press.

Jackendoff, Ray S. 2008. Construction after Construction and its theoretical challenges. *Language* 84(1): 8–28.

Jackendoff, Ray S. 2011. What is the human language faculty? Two views. *Language* 87(3): 586–624.

Jackendoff, Ray S. 2013. Constructions in the Parallel Architecture. In Thomas Hoffmann and Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, pp. 70-92. Oxford: Oxford University Press.

Kay, Paul & Charles J. Fillmore. 1999. Grammatical Constructions and linguistic generalizations: The What's X Doing Y? Construction. *Language* 75(1): 1–33.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.

Kiss, Tibor and Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42 (Handbooks of Linguistics and Communication Science). Berlin: Mouton de Gruyter (2nd edn.).

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck and Laurie Zaring (eds.), *Phrase structure and the lexicon*, pp. 109–137. Dordrecht: Kluwer Academic Publishers.

Kuroda, Sige-Yuki, 1965. Causative forms in Japanese. *Foundations of Language* 1: 30-50.

Levin, Beth, and Rappaport-Hovav, Malka. 1995. *Unaccusativity at the syntax-lexical semantics interface.* Cambridge: MIT Press.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge: MIT Press.

Marantz, Alec. 1995. 'Cat' as a phrasal idiom. Ms., Cambridge, Mass: MIT.

Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. *UPenn Working Papers in Linguistics* 4: 201-225.

Marantz, Alec. 2001. Words and things. Handout of a talk, MIT, Cambridge, MA.

Marantz, Alec. 2007. Phases and words. In Sook-Hee Choe (ed.), *Phases in the theory of grammar*, pp. 191–222. Seoul: Dong In.

Marantz, Alec. 2013. Locality Domains for Contextual Allomorphy across the Interfaces. In Ora Matushansky and Alec Marantz (eds.), *Distributed Morphology Today. Morphemes for Morris Halle*, pp. 95-115, Cambridge: MIT Press.

Matushansky, Ora. 2006. Head Movement in Linguistic Theory. *Linguistic Inquiry* 37(1): 69-109.

Merchant, Jason. 2015. How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46: 273-303.

Michaelis, Laura A. 2013. Fluid Construction Grammar. In Thomas Hoffmann and Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, pp. 153-167. Oxford: Oxford University Press.

Miyagawa, Shigeru. 1984. Blocking and Japanese causatives. *Lingua* 64(2–3): 177–207.

Müller, Stefan. 2013. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4): 920–950.

Müller, Stefan. 2015. HPSG – A synopsis. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science 42.2), pp. 937–973. Berlin: Walter de Gruyter (2nd edn.).

Müller, Stefan. 2016. *Grammatical theory. From transformational grammar to constraint-based approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press.

Nunberg, Geoffrey, Ivan Sag, and Thomas Wasow. 1994. Idioms. *Language* 70: 491-538.

Pesetsky, David. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge: MIT Press.

Pesetsky, David and Esther Torrego. 2001. T-to-C movement: causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale. A Life in Language*, pp. 355-426. Cambridge Mass.: MIT Press.

Pfau, Roland. 2000. *Features and categories in language production*. Ph.D. dissertation, Johann Wolfgang Goethe-Universität zu Frankfurt am Main.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Pylkkanen, Liina. 2008. *Introducing Arguments*. MIT Press, Cambridge, MA.

Ramchand, Gillian C. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press.

Richards, Marc D. 2015. Minimalism. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science 42.2), pp. 803-839. Berlin: Walter de Gruyter 2nd edn.

Roberts, Ian. 2010. *Agreement and head movement. Clitics, incorporation and defective goals.* Cambridge: MIT Press.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3): 486–545.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-based Construction Grammar* (CSLI Lecture Notes 193), pp. 69–202. Stanford, CA: CSLI Publications.

Sheehan, Michelle, Theresa Biberauer, Ian Roberts and Anders Holmberg. 2017. *The Final-Over-Final Condition*. Cambridge: MIT Press.

Starke, Michal. 2009. Nanosyntax: a short primer to a new approach to language. *Nordlyd* 36: 1-6.

Stump, Gregory T. 2001. *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

Svenonius, Peter. 2012. Spanning. Ms., CASTL: Tromsø.

Taraldsen, Tarald. 2010. The nanosyntax of Nguni noun class prefixes and concords. *Lingua* 120: 1522-1548.

Van der Spuy, Andrew. 2017. Construction morphology and Inflection. *Lingua* 199: 60-71.

Wasow, Thomas. 1977. Transformations and the Lexicon. In Peter W. Culicover, Thomas Wasow and Joan Bresnan (eds.), *Formal Syntax*, pp. 327-360. New York: Academic Press.

Williams, Edwin. 1981. On the notions ‘lexically related’ and ‘head of a word’. *Linguistic Inquiry* 12: 245–74.

Zeller, Jochen. 2001. *Particle verbs and Local Domains*. Amsterdam: John Benjamins.