

How to Merge a Root

Marijke De Belder

Jeroen van Craenenbroeck

The main goal of this article is to show that four properties of roots can be derived in a principled manner from the theory of Merge. The properties in question are the following: (a) roots have no grammatical features, (b) roots have no syntactic category, (c) roots are defined structurally rather than lexically, and (d) roots are dominated by functional material (rather than the other way around). We argue that the first Merge operation in each cyclic domain creates a radically empty structural position at the foot of the structure in which a root can be inserted at the level of Vocabulary Insertion. The four abovementioned properties of roots can then be shown to follow straightforwardly from this theory.

Keywords: roots, Merge, Vocabulary Insertion, Distributed Morphology

1 Introduction

The main goal of this article is to show that four properties of roots can be derived in a principled manner from the theory of Merge. The properties in question are the following: (a) roots have no grammatical features, (b) roots have no syntactic category, (c) roots are defined structurally rather than lexically, and (d) roots are dominated by functional material (rather than the other way around). The theory of Merge we put forward builds on work by Jaspers (1998), Fortuny (2008), Zwart (2009a,b, 2011), and others in assuming that this operation is inherently asymmetric. We show that as a by-product of this asymmetry, the first Merge operation in each cyclic domain¹—called Primary Merge here²—creates a radically empty structural position at the foot of the structure in which a root can be inserted at the level of Vocabulary Insertion. The four abovementioned properties of roots can then be shown to follow straightforwardly from this theory.

This article is organized as follows. In section 2, we introduce—and where necessary, provide new supporting evidence for—the four properties of roots under discussion. We show that while all four of them are empirically well-motivated, they do not follow from anything in the theory

We would like to thank Hagit Borer, Guglielmo Cinque, Norbert Corver, Marcel den Dikken, Heidi Harley, Dany Jaspers, Marjo van Koppen, Richard Larson, Omer Preminger, Eric Reuland, Guido Vanden Wyngaerd, Masaya Yoshida, the members of CRISP (Brussels) and UiL-OTS (Utrecht), and two *LI* reviewers for their very insightful questions, comments, and suggestions. Marijke De Belder would also like to gratefully acknowledge that part of this research was made possible by a postdoctoral fellowship (application number 1285513N) from the Flemish Science Foundation (Fonds voor Wetenschappelijk Onderzoek, FWO). The usual disclaimers apply.

¹ As will be made explicit in section 5.2, we define *cyclic domain* as the extended projection of a root.

² We use the term *Primary Merge* rather than *First Merge*, as the latter is often used more generally to refer to the merger between a head and its complement. It should be clear, though, that this is merely a descriptive label and that no theoretical significance should be attached to it.

and hence essentially have to be stipulated. In section 3, we introduce the theory of asymmetric Merge and discuss the special position occupied by Primary Merge in such a system, and in section 4, we show how the four central properties of roots follow from the theory outlined in section 3. In section 5, we explore a number of theoretical consequences of our proposal. We show that the theory of roots put forward here leads to a model of the grammar in which all noncomplements are spelled out first, in a separate derivation, until finally the main derivation is constructed. We further point out that although our analysis is compatible with core assumptions about Vocabulary Insertion in Distributed Morphology, it leads to an empirically more adequate insertion mechanism. In section 6, we sum up and conclude.

2 Four Properties of Roots

2.1 Introduction

In this section, we introduce the four properties of roots under discussion here (see in particular Halle and Marantz 1993, Harley and Noyer 1999, Borer 2005a,b, 2009a,b, 2013). About three of them we are fairly brief, as the arguments are well-known (see the references just mentioned) and hence need not be repeated here. However, the fourth one—the fact that roots must be defined structurally rather than lexically (section 2.4)—we discuss in more detail, as this is one of the areas where existing theories of morphology disagree. We present new evidence from word formation in Dutch supporting the hypothesis that roots are defined structurally.

2.2 *Roots Have No Grammatical Features*

As discussed in detail by Borer (2005a), open- and closed-class lexical items differ considerably in the degree of flexibility they allow in their interpretation. For example, a lexical item such as the English word *stone* can be used in a wide variety of ways, some nominal and some verbal. Examples (1a–d) illustrate a handful of them.

- (1) a. I've got a stone in my hand.
- b. There's too much stone and metal in this room.
- c. They want to stone this man.
- d. Billy-Bob should lay off the weed; he's always stoned.

The first two examples show that *stone* can be used both as a count noun (1a) and as a mass noun (1b), while in (1c) it is a transitive verb and in (1d) a verb that is obligatorily passive. As soon as functional material is added to this root, however, the interpretation becomes completely rigid. A structure such as *three stones*, which contains plural marking and a cardinal,³ is always and only interpreted as a count nominal expression; it can function neither as a mass DP, nor as a verbal element. More generally, while substantive formatives are extremely malleable and can

³ See Borer 2005a:119 on the functional nature of cardinals.

be coerced any which way, subject only to the extent of our imagination,⁴ functional ones are very rigid in their denotation, and any attempt to tweak it leads to ungrammaticality rather than uninterpretability. For example, **three much stones* is illicit, as it is marked both as a mass nominal expression by the functional vocabulary item *much* and as a count nominal expression by the functional vocabulary item *three* and via plural marking.

The absence of grammatical features thus constitutes the first property of roots under discussion here. It is worth pointing out that while this claim provides a straightforward account for the difference in interpretive flexibility between substantive and grammatical formatives, it does not itself follow from anything in the theory on roots (see also section 2.4 for more discussion).

2.3 *Roots Have No Syntactic Category*

The second property of roots follows from the first one. If a root has no grammatical features, the logical consequence is that it has no syntactic category either (see the examples in (1)).⁵ This point of view is further corroborated by two lines of reasoning. First of all, it is becoming increasingly clear that a large number of tasks—if not all of the tasks—traditionally ascribed to (the categorial characterization of) roots, in particular in terms of argument licensing (agreement, case, thematic roles), is in fact performed by (temporal, aspectual, event-introducing) functional projections (see Marantz 1997, Borer 2005a,b, 2009b, Adger 2011). The sole function of the root is to add conceptual meaning to the structures built by syntax, but for this it does not need to have a syntactic category.

The second reason for not assigning category labels to roots is that doing so introduces a large degree of redundancy into the system (Borer 2005a:20). Suppose the word *stone* was categorized either as N or as V. One would then still have to stipulate that the noun *stone* can only merge with D, while the verb *stone* only merges with T. In more technical terminology, D selects for (or checks the categorial features of) N(P) and similarly for T and V(P), thus effectively leading to a reduplication of the nominal/verbal characterization of this extended projection (see Doetjes 1997 and Grimshaw 2005 for proposals along these lines). In the approach adopted here, however, an acategorial root that merges with D is interpreted as nominal and one that merges with T is interpreted as verbal.

Summing up, there are both empirical and theoretical advantages to adopting the idea that roots are acategorial. Ideally, however, one would like to derive this property from more fundamental primitives of the theory, a goal that neither Distributed Morphology nor Borer's (2005a,b, 2009a,b) Exo-Skeletal Model achieves at this point.

⁴ We do not mean to imply that words such as *stone* are completely devoid of content. For one thing, *stone* is listed in the Vocabulary and as such has a phonological matrix (and hence phonological features). For another, the Encyclopedia constrains its meaning by linking the Vocabulary item and the syntactic structure in which it appears to particular conceptual knowledge (Marantz 1995). However, phonological features and information from the Encyclopedia do not affect the syntactic derivation; they only become relevant postsyntactically (Marantz 1995, Borer 2005a).

⁵ Thanks to Eric Reuland (pers. comm.) for reminding us: to the extent that having a grammatical category entails bearing certain grammatical features (see, e.g., Muysken and Van Riemsdijk 1985), the first two properties amount to one and the same.

2.4 *Roots Are Defined Structurally, Not Lexically*

As pointed out in section 2.1, that roots are defined structurally rather than lexically is an issue with respect to which current theories of morphology disagree. We show that a crucial factor in this debate is the question of whether vocabulary items—including roots—are inserted early or late. In the former case, roots must be defined lexically, while in the latter, a structural definition is called for. In this section, we point out a correlation between the way roots are defined and their point of insertion, and we provide new evidence from Dutch in support of the late insertion approach. The conclusion will be that roots are not marked as such in the lexicon; instead, they correspond to a particular position in the syntactic structure.

2.4.1 Roots That Are Inserted Early Are Defined Lexically Suppose the lexical item *stone* was inserted at the very start of the syntactic derivation. In that case, there would be only one way to account for its extreme malleability (see section 2.2)—namely, by stipulating that this lexical item itself has no grammatical features.⁶ In this respect, it would differ from functional items such as *many* or *the*, which are inherently marked for [+count] and [+definite], respectively. More generally, the distinction between roots and nonroots would be made in the lexicon: roots are lexical items without grammatical features, whereas nonroots are lexical items with grammatical features. Whenever a root is merged into the derivation, its (conceptual) meaning can be coerced any which way by the syntactic context, but when a nonroot is merged, its grammatical features determine whether it can be legitimately inserted or not.

The most explicit advocate of the lexical definition of roots is Borer's (2005a,b, 2009a,b) Exo-Skeletal Model (henceforth XSM). In this theory, vocabulary items come in two varieties: lexical ones (henceforth LVIs)⁷ and functional ones (FVIs). The former correspond to open-class items, the latter to closed-class ones. LVIs do not bear any grammatical features, and their semantics is malleable and dependent on the syntactic structure in which they are inserted. FVIs, on the other hand, are marked for grammatical features and have a rigid and fixed denotation.

Borer proposes that phonological indices of vocabulary items (henceforth VIs) are inserted early, that is, in narrow syntax.⁸ As a result, an FVI interacts with the syntactic derivation via its grammatical features, whereas an LVI does not bring any such features into the structure. Each time such a featureless VI is merged, a root is created. This means that roots are defined lexically in the XSM; it is the absence of grammatical features on LVIs that makes them into roots.

Summing up, under a lexical definition the relevant dividing line between roots and nonroots is situated in the lexicon. What distinguishes a root from its functional structure is a lexical

⁶ One could of course also assume a productive—and pervasive—mechanism of category conversion allowing roots of one category (e.g., the noun *stone*) to be converted into another (e.g., the verb *stone*). See De Belder 2011b for arguments against this alternative.

⁷ LVIs are called *listemes* in Borer 2005a,b and *roots* in Borer 2009a,b. They correspond to what we have been calling—and will continue to call—*roots*.

⁸ A phonological index is an abstract index that refers to all possible surface instantiations of one root. For example, *think* and *thought* are both referred to by the same phonological index. Borer assumes that the index bears a phonological resemblance to the root, hence the term *phonological index*.

property; the former corresponds to a featureless VI, the latter to a VI with grammatical features. As a result, there is a strict division between the two classes of VIs.

2.4.2 Roots That Are Inserted Late Are Defined Structurally A key innovation of Distributed Morphology (henceforth DM) compared with preceding morphological theories is the so-called Late Insertion Hypothesis (see Harley and Noyer 1999).⁹ According to this hypothesis, the phonological expression of syntactic terminals is provided only postsyntactically, in the mapping to PF. In other words, syntax manipulates only abstract bundles of grammatical features, not actual VIs. Given that roots do not involve any such grammatical feature bundles, late insertion–based models make a distinction between root and nonroot (or functional) *terminal nodes*. That is, there are specific positions in the syntactic structure that will serve as the insertion site for roots at the level of Vocabulary Insertion. These positions are characterized by the absence of grammatical features and therefore do not play any active role in the syntactic derivation. The way this intuition is usually implemented is by merging the feature [Root] into the structure (see Halle and Marantz 1993, Harley and Noyer 1998, 1999). This feature can be seen as a (syntactically inactive) placeholder or diacritic that signals to the relevant postsyntactic mechanism that a root should be inserted in this position. Technical details aside, it is clear that late insertion of VIs forces one to define roots structurally rather than lexically.

2.4.3 Interim Summary In sections 2.4.1 and 2.4.2, we have made clear that early insertion of Vocabulary items leads to a lexical definition of roots, while late insertion is compatible only with a structural one. A root is defined lexically when it results from merging a *Vocabulary item* that has no syntactic features; it is defined structurally when a particular (type of) *terminal node* is featureless. We now explore the empirical predictions made by these two definitions.

2.4.4 Supporting Evidence for the Structural Account: Functional Vocabulary Items in Root Position We have just outlined the differences between the lexical definition of roots in early insertion theories and the structural one in theories adopting late insertion. The question is now to what extent these two accounts can be empirically distinguished. It is clear that in the standard scenario—FVIs (e.g., *the*) being merged in functional terminal nodes and LVIs (e.g., *book*) in root terminal nodes—the two theories make the same predictions (albeit with different theoretical machinery). As soon as we diverge from this simple picture, however, differences emerge. Suppose we want to use an FVI as a root. In an early insertion model, this state of affairs is simply unformulable. The mere presence of grammatical features on a VI will cause the projection headed by this VI to be recognized as functional rather than lexical. As a result, FVIs can never head lexical projections. In the late insertion model, however, there is no a priori ban on merging a

⁹ Note that the concept of late insertion long predates the advent of DM. The idea that lexical items are not inserted at the beginning of the derivation was prominent in the Generative Semantics era (see in particular McCawley 1968). A noticeable difference between the two, though, is the fact that for McCawley, (certain) lexical items are inserted *during* the derivation (specifically, after certain transformations; see McCawley 1968:72), rather than at the very end of it, as in standard DM. Thanks to a reviewer for pointing out the parallelism, and to Omer Preminger (pers. comm.) for providing the necessary background.

particular type of VI in a root terminal node. Given that this position is nothing but an empty placeholder, it should in principle be possible to merge either LVIs or FVIs there.¹⁰ The use of FVIs as roots thus constitutes a potential testing ground for distinguishing between the lexical and structural definitions of roots.¹¹ Consider in this respect the following examples from Dutch:

- (2) Ik heb het **waarom** van de zaak nooit begrepen.
 I have the why of the case never understood
 ‘I have never understood the motivation behind the case.’
- (3) In een krantenartikel komt het **wat/hoe/wie/waar** altijd voor
 in a newspaper.article comes the what/how/who/where always before
 het **waarom**.
 the why
 ‘In a newspaper article the what/how/who/where always precedes the why.’
- (4) De studenten **jij-en** onderling.
 the students you-INFINITIVE among.one.another
 ‘The students are on a first-name basis with each other.’
- (5) Martha is mijn tweede **ik**.
 Martha is my second I
 ‘Martha is my best friend.’
- (6) Niets te **maar-en**!
 nothing to but-INFINITIVE
 ‘Don’t object!’
- (7) Paard is een **het**-woord.
 horse is a the.NEUTER.DEFINITE-WORD
 ‘*Paard* takes a neuter article.’

Each of these sentences exemplifies the use of an FVI in a root position. In (2) and (3), a *wh*-pronoun is merged under a nominal structure, and in (5), the personal pronoun *ik* ‘I’ is. Examples (4) and (6) illustrate that a personal pronoun and a conjunction can be inserted under a verbal structure, and (7) illustrates the use of a definite article as the nonhead part of a compound. One could of course argue that these are exceptions, and that what is inserted in root position in (2)–(7) is not an FVI, but a root that is homophonous with an FVI. As it turns out, however, the use of FVIs in root position is productive. Consider first the data in (8).

¹⁰ Needless to say, much will depend on the precise insertion mechanism for VIs. We take this issue up in detail in section 5.3.

¹¹ For the opposite state of affairs (i.e., LVIs that are merged in functional terminal nodes), see Borer 2005a:10n4, De Belder 2011b, and De Belder and Van Craenenbroeck 2014.

- (8) a. het getik van de klok
 the GE-tick of the clock
 ‘the ticking of the clock’
 b. het gefluit van de vogels
 the GE-whistle of the birds
 ‘the whistling of the birds’

Dutch has a derivational word-formation process of *ge*-prefixation to form nouns referring to a pluractional event. As is illustrated in (9), this type of word formation productively allows FVIs to occur in root position.¹²

- (9) a. Ik hoef al dat ge-**maar** niet.
 I need all that GE-but not
 ‘I don’t like those constant objections.’
 b. Ik hoef al dat ge-**alhoewel** niet.
 I need all that GE-although not
 ‘I don’t like those constant considerations.’
 c. Ik hoef al dat ge-**of** niet.
 I need all that GE-or not
 ‘I don’t like those constant alternatives.’
 d. Ik hoef al dat ge-**hé** niet.¹³
 I need all that GE-PARTICLE not
 ‘I don’t like this constant need for confirmation.’
 e. Ik hoef al dat ge-**waarom** niet.
 I need all that GE-why not
 ‘I don’t like this constant need for justification.’
 f. Ik hoef al dat ge-**nooit** niet.
 I need all that GE-never not
 ‘I don’t like this constant unwillingness.’
 g. Ik hoef al dat ge-**ik** niet.
 I need all that GE-I not
 ‘I don’t like all this egocentricity.’

Rather than assuming that FVIs are systematically ambiguous between a functional and a root reading—clearly an undesirable move—we take the data in (2)–(9) to show that FVIs can be used as roots. As pointed out above, this finding argues against an early insertion–based lexical

¹² The translations given here are only indicative. The precise interpretation of these examples may vary depending on the context.

¹³ We follow Munaro and Poletto (2003) (among many others) in assuming that sentential particles are FVIs that typically realize a functional head in the clausal left periphery.

definition of roots and in favor of the structural late insertion approach. Simply put, whether or not a lexical item is a root is not due to certain inherent characteristics or properties of that lexical item, but depends solely on the structural position in which this element is merged. Certain slots in the syntactic structure—typically, the low(est) ones (see section 2.5 for discussion)—turn whatever is merged in that slot into a root.

This conclusion is further corroborated by another aspect of the data in (2)–(9). Recall that a structural definition of roots goes hand in hand with late insertion. This predicts that the grammatical features of FVIs used as roots will not have any syntactic effect: at the point of Vocabulary Insertion, the syntactic derivation is already over. This is indeed what we see in (2)–(9). For example, the FVI *waarom* ‘why’ in (2) does not type the sentence as a *wh*-question (see Cheng 1997), nor is it subject to (otherwise obligatory) *wh*-movement. In other words, the syntactic derivation does not take the grammatical features of an FVI into account when this FVI is used as a root. Similarly, the personal pronoun *ik* ‘I’ in (5) does not trigger first person singular agreement when used as a subject; rather, it triggers (default) third person singular (see (10)), suggesting that its inherent ϕ -features are not visible to the syntactic derivation.

- (10) Mijn tweede ik {*ben/is} ongelukkig.
 my second I am/is unhappy
 ‘My best friend is unhappy.’

In the same vein, the conjunction *maar* ‘but’ in (6) cannot take a sentential complement when used as a verb, as shown in (11). This means that its selectional features are inactive in (6).

- (11) *Niets te maar-en je hebt veel werk!
 nothing to but-INFINITIVE you have much work

In short, a root position is like a Bermuda Triangle for grammatical features: regardless of which element is inserted there or what its feature specification is, it will not affect the course of the syntactic derivation in any way. Needless to say, this observation is straightforwardly compatible with a late insertion approach.

2.4.5 Summary In this section, we have introduced the third property of roots under discussion in this article. We have shown that roots should be defined structurally rather than lexically; that is, there are designated positions in the syntactic structure where whatever is merged in that position starts functioning as a root. Just as was the case in sections 2.2 and 2.3, this analysis—successful though it may be in accounting for data such as those in (2)–(11)—remains stipulative and does not follow from independent principles in late insertion models of the grammar.¹⁴

¹⁴ Note in particular that in order to account for the data in (2)–(11) a DM analysis based on the [Root] feature would have to assume that all VIs—both LVIs and FVIs—are endowed with such a feature, thus rendering it technically trivial. See section 5.3 for detailed discussion.

2.5 *Roots Are Merged Lower Than Functional Material*

The fourth and final property is perhaps the most basic of the four, partly also because it long predates the XSM or DM perspective on lexical categories. It concerns the fact that lexical categories are dominated by functional material, rather than the other way around: DP dominates NP, but NP doesn't dominate DP; TP dominates VP, but VP doesn't dominate TP; and so on.¹⁵ In traditional, pre-DM/XSM days, this hierarchical asymmetry could be made sense of by arguing that in a bottom-up derivation, the lexical projection first introduces a concept that the functional material subsequently ties to a particular speech act or situation (by situating the event in space and time, identifying its referents, adding information-structural distinctions, etc.). In the perspective on roots outlined above, however, this simple account breaks down. We have shown that roots are nothing but structural positions that play no role in the syntactic derivation. With this in mind, it is unclear why they should necessarily be merged as the first/lowest element in a cyclic domain. Worse still, given that roots are acategorical and featureless (see sections 2.2 and 2.3), how can any functional head select for (and hence be merged with) a root?¹⁶ In the next section, we show how these problems dissolve under a Merge-based definition of roots.

2.6 *Summary*

In sections 2.2–2.5, we have introduced and discussed four central properties of roots: (a) they have no grammatical features, (b) they have no syntactic category, (c) they are defined structurally rather than lexically, and (d) they are dominated by functional material rather than the other way around. One thing all four properties have in common is that they must be stated as primitives of the theory; they do not follow from any independent properties of the current model of the grammar. What we would like the grammar to generate, then, is a structural position at the beginning of the derivation that is acategorical and radically featureless, where root material can be inserted during (late) Vocabulary Insertion. In section 3, we show that the first instance of asymmetric Merge creates precisely such a position.

3 **Asymmetric Primary Merge and the Null Derivation**

3.1 *Introduction*

In this section, we introduce a theory of Merge that we will use in section 4 to derive the four basic properties of roots outlined above. We proceed in two steps. In section 3.2, we argue, following Jaspers (1998), Langendoen (2003), Zwart (2009a,b, 2011), and others, that Merge is inherently asymmetric and that pair Merge rather than set Merge is the default—and arguably

¹⁵ This statement should of course be relativized to cyclic domains: within one cyclic domain, functional projections dominate lexical ones.

¹⁶ Note in this respect that Borer (2009b:1) must state as an axiom that “Extended projections must have an L [lexical] core.”

the only—structure-building mechanism in natural language. In section 3.3, we focus on the very first Merge operation in a derivation—termed Primary Merge here—and show that it involves merger with the null derivation. In section 3.3, we sum up.

3.2 *Asymmetric Merge*

The standard technical implementation of the structure-building operation Merge in present-day Minimalist theorizing is so-called set Merge. That is, Merge combines two (possibly complex) syntactic objects α and β into the set containing (precisely) these two elements: that is, $\{\alpha, \beta\}$ (see, e.g., Chomsky 1995:243, 2008:138, 2013:40). From the perspective of this bare minimum, Merge seems to be a completely symmetric operation, which takes two elements of equal stature and yields a new object that is neither linearized nor hierarchically organized.¹⁷ More generally, Merge (α, β) = Merge (β, α). In the remainder of this section, however, we present three arguments—taken from Chomsky 1995, Jaspers 1998, Langendoen 2003, and Zwart 2009a,b, 2011—suggesting that this picture is too simple, and that there is an inherent asymmetry to Merge (see also Cormack and Smith 2005, Di Sciullo and Isac 2008, Franco 2011, Osborne, Putnam, and Gross 2011). Accordingly, we adopt Zwart's (2011:101) conclusion that pair Merge rather than set Merge is the basic structure-building principle of natural language.

The first complication is highlighted by Chomsky (1995:243). After introducing the definition given above, he points out that output conditions dictate that mere set formation does not suffice. Given that different types of constituents (e.g., verbal and nominal ones) are interpreted differently at the interfaces, the distinction between them should somehow be encoded in syntax. Put differently, Merge (α, β) is not symmetric because either α or β is the head of the newly formed constituent and as a result projects its category label onto that constituent. Chomsky implements this by assuming that either α or β functions as the label of the newly formed constituent; that is, Merge (α, β) = $\{\alpha, \{\alpha, \beta\}\}$ (with α the label of the complex constituent). As Langendoen (2003:3) points out, however, $\{\alpha, \{\alpha, \beta\}\}$ is set-theoretically equivalent to the ordered pair $\langle \alpha, \beta \rangle$.¹⁸ In other words, by admitting that one of the two elements combined by Merge projects, we are led to conclude that the proper characterization of this operation involves pair Merge rather than set Merge.

A second form of asymmetry was first observed by Jaspers (1998). He draws attention to what he calls derivational asymmetry—the fact that for every Merge operation, one element is

¹⁷ Recursive operations of set Merge obviously do yield hierarchical structure, but we are focusing on the output of a single Merge operation here.

¹⁸ The most standard definition of ordered pairs is that of Kuratowski (1921), given in (i) below (see also Partee, ter Meulen, and Wall 1987:27). The definition that Langendoen (2003) refers to is the so-called short definition (see, e.g., Enderton 1972). For a formal proof that the short definition satisfies the characteristic property of ordered pairs—that is, $\langle a, b \rangle = \langle c, d \rangle$ if and only if $a=c$ and $b=d$ —see <http://us.metamath.org/mpegif/opthreg.html>.

(i) $\langle a, b \rangle =_{\text{def}} \{\{a\}, \{a, b\}\}$

derivationally prior to the other. Put differently, one element was already part of the derivational workspace before Merge took place, while the other is newly added to that workspace as a result of the operation—more specifically, as a result of the application of Select preceding Merge (see section 3.3 for discussion). Derivational asymmetry is also appealed to—though in a slightly different form—by Epstein (1999:337), who rules out c-command from a head to its specifier on the grounds that at the point in the derivation when the c-command domain of the head is determined, the specifier is not yet a member of the derivational workspace. An implementation of Merge in terms of pure set formation does not succeed in capturing derivational asymmetry, as set Merge is a strictly symmetric operation.

Third, Zwart (2009a:163) raises the following conceptual argument against set Merge. The fact that this operation takes precisely two elements as its input remains a stipulation—necessary though it may be to derive binary branching. Chomsky (2007:5) calls this “the simplest case,” but as one is the absolute minimum of elements an operation can manipulate, an implementation of Merge that can reach this minimum is to be preferred over one that uses two syntactic objects as its input. As we discuss below, asymmetric Merge achieves this goal.

Summing up, there are good reasons to assume that Merge is asymmetric and, as a result, that this asymmetry should be built into the technical implementation of this operation. In this article, we adopt and adapt Zwart’s (2009b, 2011) definition, given in (12).

(12) *Unary Merge* (prefinal version)

Merge selects a single element from a resource¹⁹ and includes it in the object under construction.

Zwart (2009b:62, 2011:101) argues that the output of this operation yields an ordered pair, so that when an element α is taken from the resource and added to δ , the derivation currently under construction, the result is $\langle \alpha, \delta \rangle$.

Before we can proceed, there is one aspect of the definition in (12) that needs further clarification. In particular, we want to make explicit what the term *element* refers to in this definition. As Chomsky (1995:383n27), Drury (1998:76n20), and Fortuny (2008:18) point out, the objects combined by Merge are *sets* of features.²⁰ For example, under the (common) assumption that T⁰ (or the C-T complex, under Chomsky’s (2008) inheritance approach) is a combination of person, number, gender, and tense features, it follows that the set consisting of these four features acts as a single atomic element for the operation Merge (Marantz 1997:2). Accordingly, we refine the definition of Unary Merge as follows:

¹⁹ We follow Zwart in using the neutral term *resource* here, rather than *lexicon*, *Numeration*, *lexical (sub)array*, or any of the other alternatives available in the literature. As far as we can tell, our proposal is compatible with all of these implementations.

²⁰ In DM, the operation responsible for creating these sets is called *fusion* or *bundling* (see Bobaljik and Thráinsson 1998, Pykkänen 2002, Marantz 2006, De Belder 2011a for discussion and examples).

(13) *Unary Merge* (final version)

Merge selects a single subset from a resource (e.g., $\{\alpha\}$), includes it in the object under construction (δ), and yields an ordered pair (e.g., $\langle\{\alpha\}, \delta\rangle$, assuming that $\{\alpha\}$ projects).

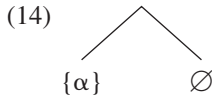
With this discussion as background, we now turn to the first application of Merge.

3.3 *Primary Merge*

An application of Merge that rarely gets a lot of attention in the literature (though see Chomsky 2013:47, to which we return below) is the very first operation in a derivation, which we call Primary Merge here. In a system based on symmetric set Merge, Primary Merge introduces a complication. Normally, Select takes an element from the resource and combines it with (the root of) the current derivation via Merge. In the case of Primary Merge, however, there is no such derivation. This leaves one of two options for Select. The first is to assume that in the case of Primary Merge, Select can exceptionally take two objects from the resource rather than one (as Chomsky (2008:139) seems to imply; see also Zwart 2011:102). It is clear that this is not a very desirable move. By allowing Select to target either one or two objects, we introduce a degree of arbitrariness into the system that can only be maintained by pure stipulation—if 1 or 2, why not 3 or 4 or 42? Moreover, it is not clear how the selection of two objects can be limited to only the first step of the derivation. This option seems to predict the possibility of ternary Merge at a later point in the derivation. The second option is to assume that Select always targets precisely one element from the resource, but that in the case of Primary Merge, the first application of Select does not immediately feed into Merge. Put differently, Select takes the first element from the resource in anticipation of the merger operation that will take place after it has applied a second time. Like the first scenario, this one has little appeal, as we are now introducing into the system a substantial degree of lookahead and a concomitant increase in computational workload.

Part of this problem is alleviated under the perspective sketched in section 3.2. As is clear from the definition of Unary Merge in (13), this operation always and without exception targets a single element from the resource. What remains to be determined, then, is what it means to be “included in the object under construction” when in fact there is no such object yet. We propose to take the definition in (13) as literally as possible. When an element $\{\alpha\}$ is the first one to be taken from the resource by Unary Merge, it is included into an empty derivational workspace; that is, the object under construction is the empty set \emptyset (see also Zwart 2011:102). The output of this instance of Merge is no different from any other: it yields an ordered pair, in this case $\langle\{\alpha\}, \emptyset\rangle$. In other words, Primary Merge is identical to all other Merge operations. All of them yield an ordered pair, the only difference being that in the case of Primary Merge, the right-hand member of this ordered pair is the empty set.

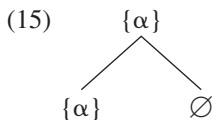
Let us consider how a(n abstract) derivation would proceed under the assumptions outlined above. Suppose we use the resource $R = \{\alpha, \beta\}$ as the input for a derivation. The first instance of Unary Merge takes a subset from R —say, $\{\alpha\}$ —and includes it in the object under construction (i.e., \emptyset), yielding the structure in (14).



The question at this point is what the label of this complex constituent is. So far, we have been proceeding under the assumption that merging an element from the resource to the derivation under construction yields an ordered pair of which that newly merged element is the left-hand member. However, recall that in section 3.2 we converted Chomsky’s labeled structure $\{\alpha, \{\alpha, \beta\}\}$ into the ordered pair $\langle \alpha, \beta \rangle$. This means that being the left-hand member of an ordered pair correlates with projection, not with being the most recently merged element. When discussing projection/labeling, Chomsky (1995, 2008, 2013) makes a distinction among three possible Merge outputs: $\{H, XP\}$, $\{XP, YP\}$ (further differentiated on the basis of whether this object is the result of Internal or External Merge), and $\{H, H\}$. It is the third case that is of interest here. Chomsky’s discussion of it is limited to the following short paragraph:

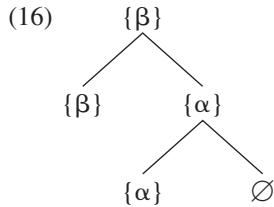
Another long-standing problem has to do with head-head constructions, the first step in a derivation. If the Marantz-Borer conception is adopted, these will be of the form f -root, where f is one of the functional elements determining category. Suppose that root, like conjunction, does not qualify as a label. In that case these constructions will be labeled f , as intended, because no other element is visible to LA [the labeling algorithm]. (Chomsky 2013:47)

We agree with Chomsky’s conclusion (i.e., that roots do not project), but want to improve upon the stipulation/assumption that is inherent in the above quotation (i.e., “*Suppose* that root . . . does not qualify as a label”). The partial derivation just outlined allows for exactly that: in the structure in (14), $\{\alpha\}$ necessarily projects, without the need for an additional assumption. The reason for this is twofold. First, \emptyset is completely and radically empty: it has no category, no grammatical features, no specification of any kind. Under the uncontroversial assumption that projection involves passing information from a daughter node on to its mother, or copying information from a daughter node onto its mother, this radical emptiness makes \emptyset inherently incapable of projecting. Second, if \emptyset were to project, the constituent in (14) would incorrectly be identified as an empty derivation, thus obscuring the fact that (the nonempty element) $\{\alpha\}$ has already been merged into that derivation.²¹ Summing up, then, the complete representation of the Primary Merge operation described above is as in (15).



²¹ As a reviewer points out, projection of the empty set might be employed in “situations of radical ellipsis—if these exist.” While we agree that this is a possible perspective, one that might tie in with recent discussions of opacity effects in ellipsis sites (see Aelbrecht 2010, Baltin 2013), pursuing it here would lead us too far afield.

At this point, Unary Merge takes $\{\beta\}$ from R and merges it with the structure in (15). Supposing that $\{\beta\}$ is the projecting element, this Merge operation yields the ordered pair $\langle\{\beta\}, \langle\{\alpha\}, \emptyset\rangle\rangle$, which can be represented as in (16).



In short, the theory of Merge developed in this section and the preceding one entails that every derivation begins with a radically empty element, which accordingly sits at the very foot of the syntactic structure. While representations such as the one in (16) might seem unorthodox at first, they are certainly not unprecedented. For instance, Zwart (2009a, 2011) also assumes that Primary Merge involves “merger with nothing/the empty set” (Zwart 2011:105). His implementation, however, differs from ours in two ways. First, his system is strictly top down—or rather, left to right: Merge splits the resource into an ordered pair consisting of one item from the resource as left-hand member and the remainder of the resource on the right. When the last member of the resource is thus split off, what remains as the right-hand member is the empty set. A second, more important difference is that Zwart does not assume this empty set to occupy a structural position in the phrase structure representation (Zwart 2009a:164). As the above discussion makes clear (and see also section 4), in our proposal the position created by merger with the empty set is real and plays a central role in natural language.²²

The idea that Primary Merge involves merger with the empty set is also proposed by Fortuny (2008). He starts out from a particular implementation of set Merge, whereby this operation takes two subsets from the resource and yields the union of those subsets (Fortuny 2008:18). Moreover, at least one of the two subsets must be the output of an immediately preceding application of Merge. For Primary Merge, this entails that one of the two elements targeted by Merge must be the empty set \emptyset . When \emptyset is merged with a first subset from the resource (e.g., $\{a\}$), the output is the union of those two sets (i.e., the singleton $\{a\}$). This set can then be used as input for the second application of Merge (e.g., $\{b\}$), to yield the union of those sets (i.e., $\{a, b\}$), and so on. Note that, just as with Zwart’s system, the effect of the initial empty set on the remaining derivation is nonexistent. This is exactly where Fortuny’s analysis differs from ours: while we agree that

²² As Zwart (2011:98) points out, the Linear Correspondence Axiom might necessitate positing an empty position independently of any considerations related to Merge. Given that this principle depends on asymmetric c-command in order to convert hierarchical structure into linear order, the merger of two nonbranching nodes—precisely the kind of structure that is created by Primary Merge—yields a nonlinearizable structure. Moro (2000) proposes that such configurations must be rendered asymmetric via movement before they reach the PF interface, but more generally, one might require that every right-branching structure end in an empty position (see Johnson 2007:5n6 for a similar solution). The fact that the perspective on Merge developed here yields precisely this result on independent grounds thus serves as additional support for our proposal.

Primary Merge involves the empty set as one of its members, we claim that this empty position is a syntactic terminal that receives phonological exponence in the postsyntactic morphological module (see section 4).

A third parallel to the existing literature concerns Guimarães's (2004:221ff.) Starting Axiom, which states that the first branching node in every derivation involves the empty set as one of its sisters. Just like Zwart's system, though, Guimarães's works top to bottom, and as in both Zwart's and Fortuny's systems, the empty set postulated by the Starting Axiom plays no further role in the syntactic derivation. It is there 'just to guarantee the appropriate syntactic configuration' (Guimarães 2004:222). As we will now demonstrate, this empty position has a far more central role to play in the system developed here.

3.4 Summary

In this section, we introduced and discussed our theory of Merge. Following Jaspers (1998), Langendoen (2003), Zwart (2009a,b, 2011), and others, we argued that this operation is asymmetric and that it yields ordered pairs rather than unordered two-membered sets. We then focused on the very first instance of Merge in a derivation (Primary Merge) and concluded that it involves the empty set as one of its members. This implies that every derivation begins with a radically empty and featureless slot at the most deeply embedded position in the structure.

Before we proceed to the next section, we want to explore very briefly a residual issue brought up by a reviewer. The reviewer agrees that there is an asymmetry in Merge, but wonders whether this asymmetry is inherent (i.e., part of the Merge operation itself, as we have been suggesting) or derivative (e.g., due to the elements undergoing Merge or to other, more general properties of the derivational procedure). For example, derivational asymmetry might be due/related not to Merge itself, but to the general characterization of a derivation as an ordered sequence of consecutive operations. Similarly, a crucial ingredient of Chomsky's (2008, 2013) labeling algorithm is that the label of a syntactic object can be wholly derived from the (types of) elements undergoing Merge (possibly in combination with the type of Merge: Internal vs. External), and thus it need not be coded in the Merge operation itself. Settling this issue is a far from trivial matter and would lead us further afield than this article—whose main focus is root syntax—allows. Moreover, nothing in what follows hinges on the inherent vs. derived nature of the asymmetry in Merge. Accordingly, we will continue to follow Jaspers (1998), Langendoen (2003), Cormack and Smith (2005), Di Sciullo and Isac (2008), Zwart (2009a,b, 2011), Franco (2011), and Osborne, Putnam, and Gross (2011) in adopting an inherently asymmetric view on the Merge operation itself. However, the reader should bear in mind that this is not the only option.

4 Deriving the Properties of Roots

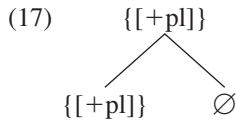
4.1 Introduction

This section combines the insights of the previous two. In section 2, we introduced four central properties of roots, and from the evidence in section 3 we concluded that Merge is asymmetric

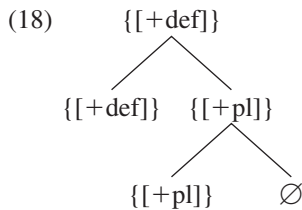
and that the first application of Merge involves the empty set as one of its members. Here, we argue that the empty structural position thus created serves as the vocabulary insertion site for roots. In section 4.2, we go through a sample derivation of a nominal constituent to illustrate how exactly the theory of Merge interacts with that of roots, and in section 4.3, we return to the basic properties of roots outlined in section 2.

4.2 *Asymmetric Primary Merge of Roots: A Sample Derivation*

In this section, we present a sample derivation of the nominal constituent *the books* to illustrate how our view on Merge meshes with the syntax of roots. Recall that we adopt late insertion of VIs. This implies that the resource from which Merge draws contains only grammatical features. Moreover, roots play no role in the syntactic derivation, and they are defined structurally. As a result, there are no features in the resource that refer to or anticipate the merger of a root. For the example at hand, this means that the resource is a set containing a definiteness feature and a plural feature—that is, $R = \{[+def], [+pl]\}$.²³ Given this resource, the derivation proceeds as follows. Unary Merge first selects the singleton set containing the plural feature from the resource and merges it with the empty set. Given that the latter is featureless, it is the plural feature that projects (see section 3.3 for discussion). This is shown in (17).



Next, the definiteness feature is targeted by Merge. It too projects its own structure, yielding the representation in (18).



Given that $[+def]$ is the highest feature in the nominal domain, it is a phase head (Svenonius 2004, Bošković 2014), and merging it triggers Transfer of the syntactic derivation to PF.²⁴ One of the operations on the way to the interface with the articulatory-perceptual system is Vocabulary

²³ For the sake of exposition, we abstract away from any other grammatical features that might underlie the DP *the books*. For example, most DM analyses would include a so-called little *n* in the resource.

²⁴ Note that in this particular instance, the resource is also depleted after the merger of $\{[+def]\}$ and so by definition the syntactic derivation is finished (at least as far as External Merge is concerned). See section 5.2, though, for derivations where phase-head-triggered Spell-Out does not coincide with resource depletion.

Insertion (see Harley and Noyer 1999 for discussion). When confronted with the structure in (18), the grammar searches its Vocabulary for matching VIs, and it encounters the following ones:²⁵

- (19) a. /ðə/ ↔ [+def]
 b. /s/ ↔ [+pl]
 c. /bʊk/ ↔ ∅

The phonological exponents on the left-hand side of the VIs in (19) are inserted into the terminal nodes of the structure in (18), and the derivation converges as the nominal constituent *the books*.²⁶

4.3 Returning to the Four Root Properties

The derivation in (17)–(19) demonstrates how the empty position created by asymmetric Primary Merge can serve as an insertion site for roots. Here, we return to the four properties of roots introduced in section 2 and show how they follow from the theory developed so far. Recall that the properties of roots under discussion here can be summarized as in (20).

- (20) a. Roots have no grammatical features.
 b. Roots have no grammatical category.
 c. Roots are defined structurally, not lexically.
 d. Roots are merged lower than functional material.

The property in (20c) is the one that has featured most prominently in the preceding discussion. We pointed out that late insertion–based theories require that roots be defined as designated positions in the structure rather than as a special featural marking on a specific subset of the lexicon. This conclusion follows directly from our theory of asymmetric Primary Merge. By merging an element from the resource to the empty set, the basic structure-building mechanism creates just the required syntactic terminal. Moreover, given that this terminal is the empty set, it also follows that this position is completely featureless (see (20a)). Regardless of which VI is inserted in this position in the morphological component (and see section 5.3 for discussion of the insertion mechanism), during the syntactic derivation this position will be completely inert (see also Bresnan 1978, 1995 and Hale and Keyser 1993 for early discussion of the nondeterministic relation between lexical semantics and syntactic structure).

The fact that roots are acategorical also follows from our theory. Recall that ∅ never projects. Given that it is radically featureless, it cannot pass its features on to a higher node, or copy its features onto it. It is always the set that merges with ∅ that projects and thus determines the category of the whole. If that set contains nominal features (as in (17)), then the VI inserted in ∅ is interpreted as a noun; if it is verbal, then the root is interpreted as a verb; and so on.

²⁵ We return to the properties of the vocabulary insertion mechanism in section 5.3.

²⁶ Note that the plural morpheme is spelled out to the right of the root whereas the tree in (18) suggests it is linearized to its left. See Embick and Noyer 2001 for discussion of the various ways in which this inversion can come about. Given that this issue is orthogonal to the reasoning developed in the main text, we gloss over it for expository purposes.

Finally, the fact that roots are merged lower than functional material is also an integral part of the analysis. The only stage at which the derivation(al workspace) is empty is the very beginning, and so the only Merge operation that can involve \emptyset as one of its members is Primary Merge. Given that the occurrence of roots directly depends on merger with \emptyset , this implies that roots are never inserted midcycle, that is, dominating previously merged functional material.

Summing up, then, the theory of asymmetric Primary Merge outlined in section 3 derives the four properties of roots discussed in section 2, thus reducing them to theorems of this theory. Given that these properties are empirically well-motivated (see section 2 for discussion and references), we take this to be additional support for our theory of Merge.

5 Theoretical Consequences of the Analysis

5.1 Introduction

In section 4, we have shown how the basic properties of roots follow from the theory of (asymmetric) Merge. At the same time, however, our analysis has left two fairly central questions unanswered so far. We address them in this section. First (in section 5.2), we focus on the derivation of linguistic expressions containing more than one root. Given that root positions can only be associated with Primary Merge, we are led to adopt layered derivations for such expressions (see also Zwart 2009a, 2011). We explore the consequences of this analysis and conclude that all noncomplements are spelled out first, in a separate derivation, until finally the main derivation is constructed. In section 5.3, we address the consequences of our proposal for Vocabulary Insertion. We argue that the absence of any grammatical features in root positions allows all VIs to be inserted in that position. In section 5.4, we summarize.

5.2 Derivations with Multiple Roots: Layered Derivations

5.2.1 Introduction Recall from section 3 that Primary Merge merges a subset of the resource to the empty derivational workspace, that is, to the empty set. This empty syntactic terminal then serves as the insertion site for a root at the postsyntactic level of Vocabulary Insertion. Given that a derivational workspace by definition contains only one instance of Primary Merge, it follows that there is a one-to-one correspondence between the number of roots a structure hosts and the number of derivational workspaces it is the output of. We can formulate this as in (21).

(21) *One Derivational Workspace, One Root (ODWOR)*

In every derivational workspace, there is exactly one root, and for every root there is exactly one derivational workspace.

Needless to say, this principle raises analytical questions with respect to strings containing multiple roots such as the VP in (22a) or the clause in (22b). The ODWOR principle predicts that (22a) should be the result of two independent derivational workspaces (one for each root), while (22b) should contain three. More generally, the approach developed so far leads us to adopt the concept

of layered derivations (in the sense of Zwart 2009a, 2011), whereby the output of one derivation(al workspace) can appear as an atom in the next one.²⁷

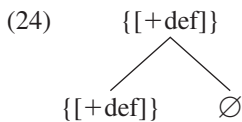
- (22) a. eat the cookie
b. The child eats the cookie.

In the remainder of this section, we combine our analysis of roots with the layered-derivations approach and show how data such as those in (22) can be accounted for in such a system. We consider two possible implementations. The first (section 5.2.2) is a conservative one, whereby layering of derivations simply means readmittance to the resource without any further syntactic—in particular, opacity—effects. The second implementation (section 5.2.3) is more ambitious and assumes that layered derivations involve multiple Spell-Out, with concomitant opacity effects. The model we arrive at is very much akin to the radical version of Uriagereka’s (1997) multiple Spell-Out model whereby all noncomplements—including, in our view, the v-V complex—are spelled out first, in separate derivations, until finally the main derivation is constructed. In section 5.2.4, we sum up.

5.2.2 Layered Derivations as Readmittance to the Resource Zwart (2009a:161) defines derivations as layered when “the output of a previous derivation [appears] as an atom in the numeration for the next derivation.” Clearly, the most straightforward way of implementing this is by allowing the end result of a derivation or derivational workspace to be readmitted to the resource from which its members were originally drawn. To show how this would work, we now go through a sample derivation of the VP in (22a). For expository purposes, let’s assume that this VP is the spell-out of a syntactic derivation built from the resource R in (23). (Recall that in our analysis, the resource contains no information about roots.)

- (23) $R = \{v, [+def]\}$

The first step of the derivation involves Primary Merge of $\{[+def]\}$ with the empty derivational workspace (i.e., with the empty set). This is represented in (24). R is revised accordingly as in (25).



- (25) $R = \{v\}$

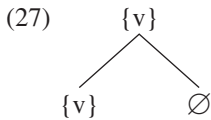
At this point, the initial derivation is stuck. In particular, Merge cannot append $\{v\}$ to the structure in (24), as this would violate the nominal functional sequence that is being built in

²⁷ Zwart (2009a, 2011) is not the only researcher proposing (something like) this. See, for example, Uriagereka 1997, Johnson 2003, Starke 2009, 2011, and Hunter 2010 for similar proposals.

(24).²⁸ Note that $\{v\}$ cannot be (Primary-)Merged to the empty set either, as the derivation is no longer in its null state: it contains the structure in (24).²⁹ The solution lies in layering the derivation—that is, in spelling out the structure in (24) (recall from section 4.2 that $[+def]$ is a phase head triggering Spell-Out) and readmitting the object thus constructed to the resource. This yields the following representation for R:

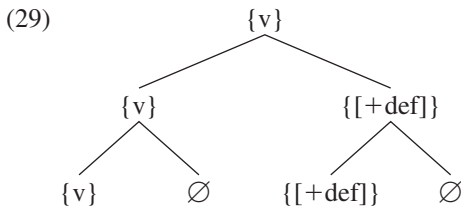
$$(26) R = \{v, \langle [+def], \emptyset \rangle\}$$

This readmittance operation has cleared the workspace, and hence makes possible a new application of Primary Merge. This time, the operation selects $\{v\}$ from R and yields the structure in (27). R is revised as in (28).



$$(28) R = \{\langle [+def], \emptyset \rangle\}$$

Merge now targets the one remaining (complex) element in R and appends it to the structure in (27), yielding the representation in (29).³⁰



At the level of Vocabulary Insertion, the VIs *eat*, *the*, and *cookie* are inserted into the relevant syntactic terminals, and the structure in (29) is spelled out as the VP *eat the cookie*.

The derivation in (23)–(29) has shown how the Primary Merge–based analysis of roots can yield surface strings containing more than one root. A key ingredient of this analysis is the

²⁸ We are assuming that functional projections are merged according to a universal functional sequence, as is common in cartographic work (see Cinque and Rizzi 2010) and that failure to adhere to this sequence leads to ill-formed structures (see also Abels 2009 and Williams 2009 for related proposals). The question of whether this functional sequence is part of Universal Grammar or whether it can be further derived (e.g., from the selectional requirements of the heads involved) is orthogonal to our main concern. See also Starke 2001 for discussion.

²⁹ More generally, this illustrates that our conception of Primary (and Unary) Merge prohibits multiple configurations from being built in tandem. For argumentation that a single-tree workspace is to be preferred over a multiple-tree workspace, see Postma and Rooryck 2009.

³⁰ The representation in (29) glosses over one complication: the fact that Merge should not select the member of R in (28) as is; rather, it should select the singleton set containing that member (see the definition in (13)). We return to this issue below.

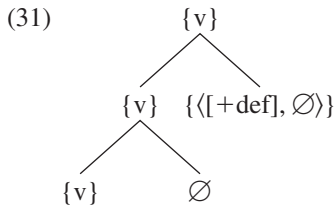
Note that (29) illustrates another consequence of our analysis of roots: a root can never directly take arguments. For example, merging the direct object in (24) directly with the verbal root—that is, readmitting it to the resource and then appending it to the null derivation via Primary Merge—would only cause the object to project further, and would not yield a verbal structure. For argumentation that roots indeed have no argument structure, see Borer 2013.

assumption that the resource can contain elements that are the output of preceding derivations and hence syntactically complex. This assumption is corroborated by Ackema and Neeleman (2004:122–129) (see also Zwart 2009a:172), who present a host of data that involve syntactically complex phrases as parts of words. Some relevant examples are given in (30).

- (30) a. a sit-on-the-sidelines Euro policy
 b. animal-to-human transplant experiments
 c. go-anywhere-at-any-time access
 d. I feel particularly sit-around-and-do-nothing-ish today.

Summing up, the ODWOR principle in (21), which is a consequence of our theory of roots, leads to the conclusion that a VP such as the one in (22a) contains two subderivations. Here, we have provided a first, simple implementation of this idea, based on Zwart's (2009a, 2011) notion of layered derivations. Next, we will add a complication to the picture: namely, the idea that layered derivations create opacity effects.

5.2.3 *Layered Derivations and Opacity* As pointed out in footnote 30, the structure in (29) glosses over the fact that the operation *Select* does not target elements from the resource directly; rather, it targets subsets of that resource. Accordingly, a more accurate representation of that structure would be the following:



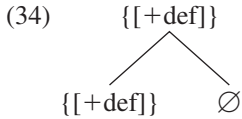
The transition from (29) to (31) is not innocuous. Given that $\{[+def], \emptyset\}$ is not of the form $\{\alpha, \{\alpha, \beta\}\}$ —that is, the output of a(n asymmetric) Merge operation—it is not recognized by C_{HL} (the computational system of human language) as a syntactically complex element, and instead is treated as an atom. Put differently, the internal structure of this constituent has become opaque to any further syntactic operations (see also Zwart 2009a:173 on layered derivations creating opacity effects). Given that direct objects such as *the cookie* in the VP in (22a) are generally not islands for extraction, this implies that the derivation presented in section 5.2.2 cannot be correct. What we propose to change is the order in which the various subderivations are readmitted to the resource. To illustrate this, we turn to the more complex example in (22b), repeated here.

- (32) The child eats the cookie.

Given that this sentence contains three roots, it will involve a three-layered derivation (or three subderivations). For the sake of concreteness (and abstracting away from tense), let us assume the resource *R* starts out as in (33).

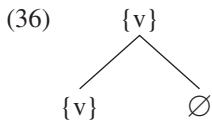
- (33) $R = \{v, [+def], [+def]\}$

As pointed out above, derivations that are readmitted to the resource become opaque for further syntactic operations. The fact that subjects are islands for extraction suggests that the subject DP *the child* is merged first. This yields the structure in (34) and the modified resource in (35).



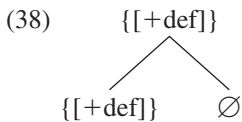
(35) $R = \{v, [+def]\}$

The structure in (34) is then readmitted to R , thus emptying the derivational workspace. The next subderivation involves the verbal structure. Zwart (2009a:175–178) argues that the v - V complex constitutes a separate derivation, with concomitant opacity effects (and see Gallego 2010 for a very similar view). Translated into the machinery developed here, this means that v is merged with the null state of the derivation, yielding the structure in (36) and the modified resource in (37).



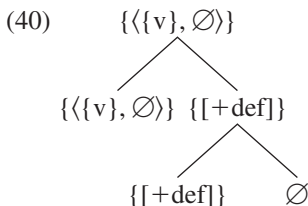
(37) $R = \{\{[+def]\}, \emptyset, [+def]\}$

At this point, the verbal structure is readmitted to R and merger of the direct object can proceed, as illustrated in (38). R now contains only the output of the previous two subderivations; see (39).

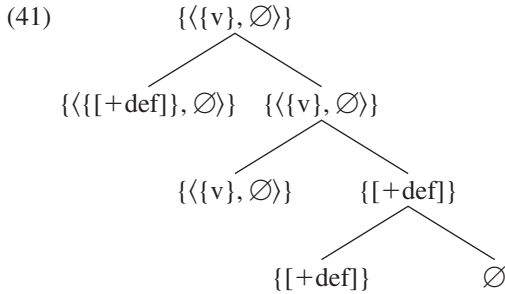


(39) $R = \{\{[+def]\}, \emptyset, \langle\{v\}, \emptyset\rangle\}$

The DP in (38) is then merged with the v - V complex, as in (40).



Finally, the subject is merged, yielding the representation in (41).

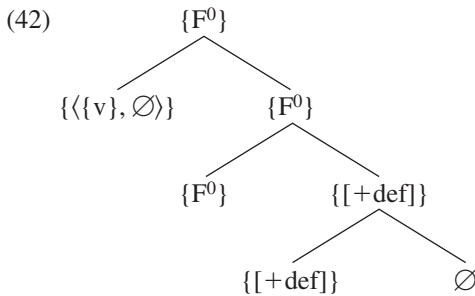


At this point, the vP phase is completed and the structure can be spelled out. At the level of Vocabulary Insertion, the syntactic terminals of (41) are matched up with their respective phonological exponents and the structure is spelled out as *The child eats the cookie*. Moreover, given that the direct object is not part of a structure that has been readmitted to R, it is correctly predicted not to be opaque for extraction. The subject and the v-V complex, on the other hand, do resist such extraction. More generally, the approach just developed suggests that everything but the most deeply embedded internal argument is spelled out first, in a separate derivation. It is only when the most deeply embedded internal argument is reached that all subderivations are brought together into one syntactic representation.³¹

While the order of derivation just described correctly predicts the distribution of opacity effects across subject, verbal complex, and direct object (opaque in the case of the former two, transparent in the case of the latter), it is clear that nothing we have presented so far *forces* the direct object to be merged last. While space limitations prevent us from exploring this issue in full here, we do want to sketch the general direction an answer to this question might take, which will at the same time lead to a further refinement of our analysis. Our approach consists in bringing together two independent strands of research. The first is Johnson's (2003) analysis of adjunct islands. The key difference between the direct object, on the one hand, and the subject and the verbal complex, on the other, is the fact that the former is a complement whereas the latter two are not. Let us start from the (uncontroversial) assumption that being a complement implies having been merged with a head. Johnson (2003) primarily focuses on one-half of that equation: if something merges with a head, then it must be a complement. He uses this to derive the fact that in a VP such as *flew after this talk* the adjunct *after this talk* is merged in a separate derivational workspace. Had it not been, then it would have been a sister of V^0 and hence would (incorrectly) have been interpreted as an argument. We want to focus on the mirror image of Johnson's principle: if something is a complement, it should be merged with an X^0 . When applied to the direct object *the cookie* in the derivation in (32)–(41), this principle runs into problems: given that the verb is merged in its own separate derivational workspace—the only way to create the verbal root (see the previous sections for discussion)—the merger between the verb and the direct object is not one between a head and a maximal projection, but one between two maximal projections, thus incorrectly identifying the direct object as a noncomplement.

³¹ We will not go into the technical details of how exactly—or even if—the various subderivations are brought together into one representation in narrow syntax, as this exceeds the main topic of this article. See Hoffman 1996, Uriagereka 1997, and Van Gelderen 2003, among others, for relevant discussion.

Enter the second ingredient of our analysis. As pointed out in section 2.3, there is a growing consensus in the literature that the introduction and licensing of arguments proceeds not via lexical heads directly, but through the mediation of functional (temporal, aspectual, event-introducing) material (see Marantz 1997, Pylkkänen 2002, Borer 2005a,b, 2009b, Adger 2011). This means that it should not be the verbal complex itself that introduces the direct object in (32)–(41); rather, it should be a functional head. This in turn solves the problem raised above in that the direct object can once again be properly identified as a complement, that is, as a phrase that has merged with an X^0 . Moreover, this X^0 is arguably part of the functional spine of the clause and hence needs to project. If we adopt Uriagereka’s (1997:117) assumption that material that has been readmitted to the resource can no longer project,³² it now follows that the direct object must be merged last. The tree in (42) represents the modified syntactic structure for the VP *eat the cookie*.



Needless to say, the sketch we have just provided leaves several questions unanswered, but at the same time it forms a promising basis for further research. Moreover, it is worth pointing out that the order of Merge operations we have proposed bears a very close resemblance to the one Uriagereka (1997) and Drury (1998) arrive at on independent grounds, thus lending further plausibility to the proposal.

Summing up, under the assumption that layered derivations create opacity effects, the analysis proposed in section 5.2.2 had to be revised. We have shown that the assumptions adopted in this article lead to a model of the grammar in which all noncomplements—including, in our view, the v-V complex—are spelled out first, in separate derivations, until finally the root derivation is constructed.

5.2.4 Conclusion If the insertion of roots is a by-product of asymmetric Primary Merge, then each derivational workspace contains precisely one root. Accordingly, strings containing more than one root must be the result of multiple—or more accurately, layered—derivations. In this section, we have provided a concrete implementation of how such a multiroot derivation proceeds. We now turn to the insertion mechanism responsible for assigning a phonological exponence to syntactic terminals.

³² We are translating Uriagereka’s proposal into our own terminology here: phrases that we consider to be readmitted to the resource constitute separate “command units” in his proposal.

5.3 *The Vocabulary Insertion Mechanism*

5.3.1 Introduction An aspect of our theory we have said very little about so far concerns the precise mechanism that matches up syntactic terminals with the appropriate VIs. This issue is complicated by the fact that unlike in previous late insertion models, in our model FVIs can be inserted in root positions (see the data in (2)–(11)). Most DM accounts of Vocabulary Insertion assume different insertion mechanisms for functional terminal nodes and root terminal nodes (Halle and Marantz 1993, Harley and Noyer 1999). While the former are subject to competition (the VI that matches the syntactic terminal most closely in features being the preferred insertion candidate), insertion of the latter is determined by free choice. In DM, this dual insertion mechanism leads to a strict division of labor between the functional and lexical domains: FVIs always and only spell out functional terminal nodes, whereas LVIs always and only spell out root terminal nodes. At first sight, then, the DM mechanism of Vocabulary Insertion is ill-suited for the data discussed earlier in this article. Recall from section 2.4.2, however, that there is a crucial difference between the structurally defined root positions in DM and those in the present system. While in our model root terminal nodes are radically devoid of features, in DM these positions are explicitly marked as such. For example, classical DM postulates a special [Root] feature in these positions (Halle and Marantz 1993). Only LVIs can be inserted in these root positions because they are the only VIs that bear this feature. In what follows, we show that our proposal to do away with [Root] features—or any comparable diacritic—has as a welcome side effect that the DM mechanism of Vocabulary Insertion now allows FVIs to be inserted in root terminal nodes as well.

We proceed in three steps. In section 5.3.2, we lay out the mechanism of Vocabulary Insertion used in DM and show how it leads to a strict division of labor between the lexical and functional domains. In section 5.3.3, we argue that postulating a [Root] feature is problematic in light of the data discussed in section 2. Finally, in section 5.3.4, we show that giving up [Root] features altogether leads to a situation in which both FVIs and LVIs can be inserted postsyntactically in root terminal nodes.

5.3.2 Vocabulary Insertion in Distributed Morphology As pointed out above, DM assumes different modes of insertion for functional terminal nodes and root terminal nodes. Insertion into functional terminal nodes is regulated by competition. More specifically, FVIs are inserted into functional terminal nodes according to the Subset Principle in (43) (Kiparsky 1973, Anderson 1986, Halle 1997).

(43) *The Subset Principle*

The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen.³³ (Halle 1997:428)

³³ Halle uses the term (*terminal*) *morpheme* for what we have been calling—and will continue to call—(*syntactic*) *terminal node*.

This procedure ensures that the VI whose feature specification matches that of the terminal node most closely will be the winner, essentially via an implementation of the Elsewhere Principle (see Caha 2009 for detailed discussion as well as an alternative in terms of supersets).

Root terminal nodes, on the other hand, do not trigger competition. Their insertion is based on free choice (Harley and Noyer 1998), although the precise way in which this free choice is implemented tends to vary. For example, Harley and Noyer (1998) propose that LVIs carry selectional features or at least a specification of the context in which they can be inserted (see also Embick 2010), while Harley and Noyer (1999) suggest that LVIs are endowed with the feature [Root] or with an index. What these options have in common, however, is that all LVIs have some sort of marking that they share with root terminal nodes, and that it is this marking that allows all of them to be inserted freely in such positions.

In sum, DM proposes that while functional terminal nodes trigger competition, root terminal nodes trigger free choice. The technical implementation of this dichotomy is feature-based. On the one hand, functional terminal nodes have specific grammatical features, and only those VIs that carry a subset of those features can enter the competition for insertion. On the other hand, both root terminal nodes and LVIs are endowed with a [Root] feature, making it possible for any LVI to be inserted in any root terminal node.

5.3.3 [Root] Problems As is clear from the previous section, root positions in our analysis are radically devoid of features. This means that there is no room for a [Root] feature or any other such diacritic or placeholder signaling that a particular structural position is reserved for roots. We believe eliminating the [Root] feature in this way is a step up from existing DM accounts in that it leads to a theory that is both theoretically and empirically more adequate. First of all, the [Root] feature is the odd one out in the presyntactic inventory of features in that it is not a morphosyntactic feature in any clear sense of the word. It is not syntactic in that it plays no role in the syntactic derivation (see section 2 for discussion), and it is not morphological in that there is no morphological trait or property distinguishing all roots from all nonroots. This idiosyncratic nature of the [Root] feature makes it theoretically suspect, and a theory that can do away with it arguably has the upper hand.

Second, recall from the data in (2)–(11) that FVIs can be inserted in root terminal nodes. This implies that the diacritic that allows LVIs to be inserted in such positions—the [Root] feature—should be present on FVIs as well. Such a move, however, would render this diacritic meaningless, as it would now be present on *all* VIs and therefore would no longer distinguish roots from nonroots.

Third, if FVIs were endowed with a [Root] feature, then by the Subset Principle this feature should also be present on functional terminal nodes, thus further hollowing out the concept of a [Root] feature: it would now be present on all types of VIs and on all types of terminal nodes.

Summing up, regulating Vocabulary Insertion in root terminal nodes by means of a [Root] feature (or any other comparable diacritic) faces considerable problems, both theoretical and empirical. Accordingly, a theory that can do away with this feature—such as the one proposed

in the preceding sections—is to be preferred. Next, we revisit the DM mechanism of Vocabulary Insertion from the perspective of a [Root]-less grammar.

5.3.4 Vocabulary Insertion without Root Markers Recall that in DM, root terminal nodes trigger free choice. This means that anything can be inserted, as long as it meets one basic requirement: it must be marked with the same diacritic as the root terminal node (e.g., a [Root] feature). As argued above, we have given up this diacritic. As a result, the one requirement on insertion in root terminal nodes also disappears, and free choice is truly unlimited. In particular, *any* VI, whether lexical or functional, can now be inserted in root terminal nodes. In other words, by giving up the [Root] feature, the DM mechanism of Vocabulary Insertion allows both functional and root VIs to realize root terminal nodes. Given that this is exactly what we find (see the data in (2)–(11)), this is a desirable outcome.

5.4 Conclusion

In this section, we have explored two consequences of the theory outlined in sections 1–4. On the one hand, we have shown how the Primary Merge analysis of roots leads to layered derivations in the case of strings containing more than one root, while on the other we have argued that the elimination of specific root diacritics allows both FVIs and LVIs to be inserted in root terminal nodes.

6 Conclusions and Prospects

In this article, we have shown that four widely discussed properties of roots follow from an articulated theory of Merge. If we agree with Chomsky (1995:378, 2000:101) that the operation Merge is “inescapable” or “indispensable” for any language-like system, then the reduction proposed here is clearly a step forward for the theory: the properties of roots no longer need to be stated separately, but can be derived as theorems from a more fundamental part of the theory.

Moreover, in the preceding discussion we have focused on Primary Merge, an aspect of the theory of Merge that (a) is often left implicit, and (b) when made explicit, is typically given a special status and thus diverges from the simplest case. We have provided an explicit discussion of this Merge operation; moreover, we have proposed treating it as identical to all other applications of this operation, thus simplifying existing accounts. As a side effect of this uniformity, every syntactic derivation starts out with a structurally empty syntactic terminal in its most deeply embedded position, and it is this position we have identified as the insertion site for roots. This derives their featurelessness and acategorical nature, while allowing for the insertion of FVIs in root terminal nodes. This last point is an important—and, as we have shown, empirically well-motivated—aspect setting our analysis apart from previous DM-based accounts of Vocabulary Insertion.

At the same time, we have shown that our proposal has far-reaching consequences in quite disparate portions of the grammar. It provides new evidence for the layered-derivations approach to expressions containing more than one root and for the hypothesis that roots never directly take

arguments. Exploring some of these consequences in more detail is a task we look forward to taking up in future research.

References

- Abels, Klaus. 2009. Some implications of improper movement for cartography. In *Alternatives to cartography*, ed. by Jeroen van Craenenbroeck, 325–359. Berlin: Mouton de Gruyter.
- Ackema, Peter, and Ad Neeleman. 2004. *Beyond morphology: Interface conditions on word formation*. Oxford: Oxford University Press.
- Adger, David. 2011. Labels and structure. Ms., Queen Mary University of London. Available at ling.auf.net/lingbuzz/001252/current.pdf.
- Aelbrecht, Lobke. 2010. *The syntactic licensing of ellipsis*. Amsterdam: John Benjamins.
- Anderson, Stephen R. 1986. Disjunctive ordering in inflectional morphology. *Natural Language and Linguistic Theory* 4:1–31.
- Baltin, Mark. 2013. Deletion versus pro-forms: An overly simple dichotomy? *Natural Language and Linguistic Theory* 30:381–423.
- Belder, Marijke De. 2011a. A morphosyntactic decomposition of countability in Germanic. *Journal of Comparative Germanic Linguistics* 14:173–202.
- Belder, Marijke De. 2011b. Roots and affixes. Doctoral dissertation, Utrecht University.
- Belder, Marijke De, and Jeroen van Craenenbroeck. 2014. On Vocabulary Insertion. *Linguistic Analysis* 39: 173–210.
- Bobaljik, Jonathan David, and Höskuldur Thráinsson. 1998. Two heads aren't always better than one. *Syntax* 1:37–71.
- 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. 2009a. Categorizing roots. Paper presented at Roots. Universität Stuttgart, 10–12 June.
- Borer, Hagit. 2009b. Roots and categories. Paper presented at the 19th Colloquium on Generative Grammar. Faculty of Letters (Gradu Aretoa), University of the Basque Country, 1–3 April.
- 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* 45:27–89.
- Bresnan, Joan. 1978. A realistic transformational grammar. In *Linguistic theory and psychological reality*, ed. by Morris Halle, Joan Bresnan, and George A. Miller, 1–59. Cambridge, MA: MIT Press.
- Bresnan, Joan. 1995. Lexicality and argument structure. Paper presented at the Paris Syntax and Semantics Conference, Paris, 12–15 October.
- Caha, Pavel. 2009. The nanosyntax of case. Doctoral dissertation, University of Tromsø.
- Cheng, Lisa Lai-Shen. 1997. *On the typology of wh-questions*. New York: Garland.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + recursion = language?*, ed. by Uli Sauerland and Hans-Martin Gärtner, 1–29. New York: Mouton de Gruyter.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.

- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.
- Cinque, Guglielmo, and Luigi Rizzi. 2010. The cartography of syntactic structures. In *Mapping spatial PPs: The cartography of syntactic structures, vol. 6*, ed. by Guglielmo Cinque and Luigi Rizzi, 3–25. New York: Oxford University Press.
- Cormack, Annabel, and Neil Smith. 2005. Linearisation: Adjuncts and arguments. *UCL Working Papers in Linguistics* 17:111–129.
- Di Sciullo, Anna Maria, and Daniela Isac. 2008. The asymmetry of Merge. *Biolinguistics* 2:260–290.
- Doetjes, Jenny. 1997. Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English. Doctoral dissertation, Leiden University.
- Drury, John E. 1998. The promise of derivations: Atomic Merge and Multiple Spell-Out. *Groninger Arbeiten zur germanistischen Linguistik* 42:61–108.
- Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.
- Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.
- Enderston, Herbert B. 1972. *A mathematical introduction to logic*. New York: Academic Press.
- Epstein, Samuel David. 1999. Un-principled syntax and the derivation of syntactic relations. In *Working Minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 317–345. Cambridge, MA: MIT Press.
- Fortuny, Jordy. 2008. *The emergence of order in syntax*. Amsterdam: John Benjamins.
- Franco, Ludovico. 2011. The strict asymmetry of Merge. Ms., Università Ca'Foscari, Venice.
- Gallego, Ángel J. 2010. Locality and anti-locality in a Merge-based system. Paper presented at the workshop Antilocality and Snowballing Movement, GIST, Ghent, 24–25 June.
- Gelderen, Véronique van. 2003. Scrambling unscrambled. Doctoral dissertation, Leiden University.
- Grimshaw, Jane. 2005. Extended projection. In *Words and structure*, 1–73. Stanford, CA: CSLI Publications.
- Guimarães, Max. 2004. Derivation and representation of syntactic amalgams. Doctoral dissertation, University of Maryland, College Park.
- Hale, Kenneth, and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 53–109. Cambridge, MA: MIT Press.
- Halle, Morris. 1997. Distributed Morphology: Impoverishment and fission. In *PF: Papers at the interface*, ed. by Benjamin Bruening, Yoonjung Kang, and Martha McGinnis, 425–449. MIT Working Papers in Linguistics 30. Cambridge, MA: MIT, MIT Working Papers in Linguistics.
- Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.
- Harley, Heidi, and Rolf Noyer. 1998. Licensing in the non-lexicalist lexicon: Nominalizations, Vocabulary items and the Encyclopedia. In *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, ed. by Heidi Harley, 119–137. MIT Working Papers in Linguistics 32. Cambridge, MA: MIT, MIT Working Papers in Linguistics.
- Harley, Heidi, and Rolf Noyer. 1999. State-of-the-article: Distributed Morphology. *Glott International* 4: 3–9.
- Hoffman, Joel M. 1996. Syntactic and paratactic word order effects. Doctoral dissertation, University of Maryland, College Park.
- Hunter, Timothy Andrew. 2010. Relating movement and adjunction in syntax and semantics. Doctoral dissertation, University of Maryland, College Park.
- Jaspers, Dany. 1998. Categories and recursion. *Journal of Applied Linguistics* 12:81–112.

- Johnson, Kyle. 2003. Towards an etiology of adjunct islands. *Nordlyd* 31:187–215.
- Johnson, Kyle. 2007. LCA + alignment = RNR. Paper presented at the Workshop on Coordination, Subordination and Ellipsis, University of Tübingen, 6–7 June.
- Kiparsky, Paul. 1973. Elsewhere in phonology. In *A festschrift for Morris Halle*, ed. by Stephen R. Anderson and Paul Kiparsky, 93–106. New York: Holt, Rinehart and Winston.
- Kuratowski, Kazimierz. 1921. Sur la notion de l'ordre dans la théorie des ensembles. *Fundamenta mathematicae* 2:161–171.
- Langendoen, D. Terence. 2003. Merge. In *Formal approaches to function in grammar: In honor of Eloise Jelinek*, ed. by Andrew Carnie, Heidi Harley, and MaryAnn Willie, 307–318. Amsterdam: John Benjamins.
- Marantz, Alec. 1995. *Cat* as a phrasal idiom: Consequences of late insertion in Distributed Morphology. Ms., MIT, Cambridge, MA.
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, ed. by Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams, 201–225. Pennsylvania Working Papers in Linguistics 4.2. Philadelphia: University of Pennsylvania, Penn Linguistics Club.
- Marantz, Alec. 2006. Morphology lectures: Streaming video. Lectures at the DEC of the ENS, Paris, October. Available at <http://www.diffusion.ens.fr/index.php?res=themesandidtheme=18>.
- McCawley, James D. 1968. Lexical insertion in a transformational grammar without deep structure. In *Papers from the 4th regional meeting of the Chicago Linguistic Society*, ed. by Bill J. Darden, Charles-James N. Bailey, and Alice Davison, 71–80. Chicago: University of Chicago, Chicago Linguistic Society.
- Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge, MA: MIT Press.
- Munaro, Nicola, and Cecilia Poletto. 2003. Sentential particles and clausal typing in the Veneto dialects. *University of Venice Working Papers in Linguistics* 13:127–154.
- Muysken, Pieter, and Henk van Riemsdijk. 1985. Projecting features and featuring projections. In *Features and projections*, ed. by Pieter Muysken and Henk van Riemsdijk, 1–30. Dordrecht: Foris.
- Osborne, Timothy, Michael Putnam, and Thomas M. Gross. 2011. Bare phrase structure, label-less trees, and specifier-less syntax: Is Minimalism becoming a dependency grammar? *The Linguistic Review* 28:315–364.
- Partee, Barbara, Alice ter Meulen, and Robert Wall. 1987. *Mathematical methods in linguistics*. Dordrecht: Kluwer.
- Postma, Gertjan, and Johan Rooryck. 2009. When spell-out is vocalization—and when it is not. Paper presented at GLOW 32, Université de Nantes, 15–18 April.
- Pylkkänen, Liina. 2002. Introducing arguments. Doctoral dissertation, MIT, Cambridge, MA.
- Starke, Michal. 2001. Move dissolves into Merge: A theory of locality. Doctoral dissertation, University of Geneva.
- Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36.1:1–6.
- Starke, Michal. 2011. Towards elegant parameters: Language variation reduces to the size of lexically stored trees. Ms., University of Tromsø. Available at <http://ling.auf.net/lingbuzz/001183>.
- Svenonius, Peter. 2004. On the edge. In *Peripheries: Syntactic edges and their effects*, ed. by David Adger, Cécile De Cat, and George Tsoulas, 259–287. Dordrecht: Kluwer.
- Uriagereka, Juan. 1997. Multiple Spell-Out. *Groninger Arbeiten zur germanistischen Linguistik* 40:109–135.
- Williams, Edwin. 2009. There is no alternative to cartography. In *Alternatives to cartography*, ed. by Jeroen van Craenenbroeck, 361–373. Berlin: Mouton de Gruyter.
- Zwart, C. Jan-Wouter. 2009a. Prospects for top-down derivation. *Catalan Journal of Linguistics* 8:161–187.

- Zwart, C. Jan-Wouter. 2009b. Uncharted territory? Towards a non-cartographic account of Germanic syntax. In *Advances in comparative Germanic syntax*, ed. by Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, Justin Nüger, and Florian Schäfer, 59–84. Amsterdam: John Benjamins.
- Zwart, C. Jan-Wouter. 2011. Structure and order: Asymmetric Merge. In *Oxford handbook of linguistic Minimalism*, ed. by Cedric Boeckx, 96–118. Oxford: Oxford University Press.

(De Belder)

Faculty of Arts

KU Leuven

Warmoesberg 26

1000 Brussels

Belgium

marijke.debelder@kuleuven.be

(Van Craenenbroeck)

Faculty of Arts

KU Leuven

Blijde-Inkomststraat 21

3000 Leuven

Belgium

jeroen.vancraenenbroeck@kuleuven.be

