How to merge a root

Marijke De Belder
CRISSP/HUB/UiL-OTS/Utrecht
marijkedebelder@yahoo.com
www.marijkedebelder.com

&

Jeroen van Craenenbroeck
CRISSP/HUB/FUSL/KUL
van.craenenbroeck.jeroen@gmail.com
www.jeroenvancraenenbroeck.com

abstract

Roots have been argued to be (a) devoid of grammatical features, (b) categoryless, (c) structurally defined, and (d) dominated by—rather than dominating—functional material (see among others Halle and Marantz 1993, Borer 2005a, 2005b, 2009a, 2009b). We show that these four properties follow straightforwardly from the theory of Merge. Adopting and adapting ideas from Jaspers (1998), Fortuny (2008), and Zwart (2009a, 2009b, 2010), we argue that Merge is inherently asymmetric, and that the first merge operation in each (sub)derivation—called Primary Merge here—combines an element from the vocabulary inventory with the null derivation. At the level of Vocabulary Insertion, this null derivation serves as the insertion site for roots. In addition, the proposal outlined here allows us to formulate a vocabulary insertion mechanism that treats functional terminal nodes and root terminal nodes alike, and it leads to a model of the grammar in which derivations proceed left-to-right across cyclic domains, but bottom-up within each cyclic domain (cf. Uriagereka 1999).

keywords

roots, asymmetric Merge, primary Merge, derivational model, Distributed Morphology, Exo-Skeletal Model

1. Introduction*

The main goal of this paper is to show that four central stipulations—or axioms if you will—of currently prevalent theories of roots can be derived in a principled manner from the theory of Merge. The four stipulations 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 Jaspers (1998), Fortuny (2008), Zwart (2009a, 2009b, 2010), 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 domain[1]—called Primary Merge in this paper[2]—creates a radically empty structural position at the bottom of

* We would like to thank Hagit Borer, Guglielmo Cinque, Marcel Den Dikken, Heidi Harley, Dany Jaspers, Guido Vanden Wyngaerd, Masaya Yoshida, and the members of CRISSP (Brussels) and UiL-OTS (Utrecht) for their very insightful questions, comments and suggestions. The ususal 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 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 a merely descriptive label and that no theoretical significance should be attached to it.

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 paper is organized as follows. The next section introduces—and where necessary, provides 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 and hence essentially have to be stipulated. Section three introduces the theory of asymmetric Merge and discusses the special position occupied by Primary Merge in such a system, while section four shows how the four central properties of roots follow without any additional assumptions from the theory outlined in section three. Section five explores 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 derivations proceed from left to right across cyclic domains, but bottom-up within each individual cyclic domain. Moreover, our analysis paves the way for a mechanism of Vocabulary Insertion in which functional and lexical terminal nodes are treated alike. Section six reviews and addresses a number of potential threats for our account, and section seven sums up and concludes.

2. Four axioms about roots

2.1 Introduction

This section introduces the four axioms about roots under discussion here (see in particular Halle and Marantz 1993, Harley & Noyer 1999, Borer 2005a, 2005b, 2009a, 2009b). About three of them we are fairly brief, as the arguments in support of these claims are well-known (see the references just mentioned) and hence need not be reiterated here. The fourth one, however, i.e. the fact that roots have to be defined structurally rather than lexically (see section 2.4), we discuss in more detail, as this is one of the areas where existing theories of morphology take a different stance. We present new evidence from word formation in Dutch supporting the hypothesis that roots have to be 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. A handful of them is illustrated in (1).

(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 ((1)a) and as a mass noun ((1)b), while in (1)c it is a transitive verb and in (1)d one that is obligatorily passive. As soon as functional material is added to this root, however, the interpretation becomes completely fixed. 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 neither function as a mass DP, nor as a verbal element. More generally, while substantive formatives are

[3] See Borer (2005a:119) on the functional nature of cardinals.

extremely malleable and can 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 both marked 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. In the context of this paper, 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 (see also below, section 2.4 for more discussion); that is, it remains a stipulation.

2.3 Roots have no syntactic category

The second property of roots follows more or less directly from the first one. If a root has no grammatical features, the logical consequence is that it has no syntactic category either (cf. the examples in (1)). 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—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 (cf. Marantz 1997, Borer 2005a,b, 2009b). 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 this operation 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 projection (see Doetjes 1997 and Grimshaw 2009 for a proposal 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 considerable 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 which neither Distributed Morphology nor Borer's Exo-Skeletal Model achieve at this point.

2.4 Roots are defined structurally, not lexically

As pointed out in section 2.1, this is an issue with respect to which there is disagreement among current formal theories of morphology. We show that a crucial factor in this debate is whether vocabulary items—including roots—are inserted early or late. In the former case roots have to be defined lexically, while in the latter a structural definition is called for. This section points out the correlation between the way roots are defined and

---

[4] Needless to say, we do not mean to imply that words such as *stone* are completely devoid of content in general. For one, it is listed in Vocabulary and thus has a phonological matrix (and hence phonological features). Secondly, 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 Encyclopedia do not affect the syntactic derivation; they only become relevant post-syntactically (Marantz 1995, Borer 2005a).

their point of insertion and further provides 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, but that 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 is only one way to account for its extreme malleability (see above, section 2.2) and that is to stipulate that this lexical item itself has no grammatical features.[5] In this respect it would differ from functional items such as *many* or *the*, which are inherently marked for [+count] or [+definite] respectively. More generally, the distinction between roots and non-roots is one that is made in the lexicon: roots are lexical items without grammatical features, whereas non-roots 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 non-root is merged, its grammatical features determine whether it can be legitimately inserted or not—and if not, the result is ungrammaticality, not uninterpretability.

The most consistent 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)[6] 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[7] of vocabulary items (henceforth VIs) are inserted early, i.e. in Narrow Syntax. 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—i.e. an LVI—is merged, a root is created. This means that roots are defined lexically in the XSM; it is the absence of grammatical features on the LVIs that makes them into roots.

Summing up, under a lexical definition the relevant dividing line between roots and non-roots is situated in the lexicon. What distinguishes a root from its functional structure is a lexical property; the former is the result of merging a featureless VI, the latter results from merging a VI with grammatical features. As such, 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 compared to preceding morphological theories is the so-called Late Insertion Hypothesis (cf. Harley & Noyer 1999). It concerns the idea that the phonological expression of syntactic terminals is only provided post-syntactically, in the mapping to PF. In other words, syntax only manipulates abstract bundles of grammatical features, not actual vocabulary items. However, given that roots do not introduce any such grammatical features, Late Insertion-based models make a distinction

[5] 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 (to appear) for arguments against this alternative.
[6] 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.
[7] A phonological index is an abstract index that refers to all 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 name *phonological* index.

between root and non-root (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 as such they do not play any active role in the syntactic derivation. The way in which this intuition is usually implemented is by merging the feature [Root] into the structure (see Halle & Marantz 1993, Harley & Noyer 1998, 1999). This feature can be seen as a (syntactically inactive) placeholder or diacritic that will signal to the relevant post-syntactic mechanism that a root should be inserted in this position. Technical details aside, though, it should be clear that late insertion of vocabulary items forces one to define roots structurally rather than lexically.

2.4.3 Interim summary

The previous two subsections have made clear that early insertion of vocabulary items leads to a lexical definition of roots, while late insertion is only compatible 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. The next section explores 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 models 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, i.e. functional vocabulary items (such as *the*) being merged in functional nodes and lexical vocabulary items (e.g. *book*) in root 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 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 particular type of VI in a root terminal node.[8] Given that this position is devoid of features, it is immaterial whether the VI realizing this position post-syntactically bears any grammatical features. The use of FVIs as roots thus constitutes a potential testing ground for distinguishing between the lexical and the structural definition of roots.[9] 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.'

[8] Needless to say, a lot will depend on the precise insertion mechanism for VIs. This is an issue we return to in section 5.3. A related question concerns the role of the [Root]-feature in determining which VIs can be inserted in an root terminal node: if FVIs do not carry such a feature, then they arguably cannot be inserted in this position under a structural definition of roots either. Given that our analysis will do away with the [Root]-feature altogether, this question will become obsolete in the next section. See note 12 for some discussion, though.

[9] We return to the opposite state of affairs, i.e. LVIs that are merged in functional nodes, in section 5.3 below. See also Borer 2005a:10n4 for some discussion.

(3)

In	een	krantenartikel	komt	het	**wat/hoe/wie/waar**
in	a	newspaper.article	comes	the	what/how/who/where
altijd	voor	het	**waarom**.		
always	before	the	why		

'In a newspaper the what/how/who/where always precedes the why.'

(4)

De studenten	**jij**-en	onderling.
the students	you-INFINITIVE	amongst.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	$the_{neuter.def}$-word

'*Paard* takes a neuter article.'

Each of these examples exemplifies the use of an FVI in a root position. In (2) and (3) a *wh*-pronoun merges under a nominal structure, while in (5) the personal pronoun *ik* 'I' does. Examples (4) and (6) illustrate that a personal pronoun and a conjunction can be inserted under a verbal structure, while (7) illustrates the use of a definite article as the left-hand 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 rather a root which 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).

(8) a.

het getik	van	de	klok
the GE-tick	of	the	clock

'the ticking of the clock.'

b.

het gefluit	van	de	vogeltjes
the GE-whistle	of	the	birds

'the whistling of the birds.'

Dutch has a derivational word-formation process to form nouns which refer to a pluractional event by means of *ge*-prefixation. As is illustrated in (9), this type of word-formation productively allows FVIs to occur in root position.[10]

[10] The translations given here are only indicative. The precise interpretation of these examples may vary according to the context.

(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.[11]
I need all that GE-PRT not
'I don't like the constant need for confirmation.'

e. Ik hoef al dat ge-**waarom** niet
I need all that GE-why not
'I don't like the constant need for justification.'

f. Ik hoef al dat ge-**nooit** niet
I need all that GE-never not
'I don't like the 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 assume 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 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 the next section for discussion—turn whatever is merged in that slot into a root.

This conclusion is further corroborated by an aspect of the data in (2)-(9) left undiscussed so far. 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 (cf. 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, but rather (default) third person singular (see (10)), suggesting that its inherent [ϕ]-features are not seen by the syntactic derivation.

[11] We follow Munaro & Poletto (2003) (among others) in assuming that sentential particles are FVIs realizing a functional head in the clausal left periphery.

(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 is shown in (11). This means that its selectional features are inactive in the example 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 paper. We have shown that roots should be defined structurally rather than lexically, i.e. there are designated positions in the syntactic structure whereby whatever is merged in that position starts functioning as a root. Just as was the case in the previous two sections, 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.[12]

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, etc.[13] In traditional, pre-DM/XSM-days, this hierarchical asymmetry could be made sense of by arguing that in a bottom-up derivation, the lexical projections first introduced the conceptual notions which the functional material then subsequently tied 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 whatsoever 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 acategorial and featureless (see above, sections 2.2 and 2.3), how can any functional head select for (and hence be merged with) a root?[14] The next section shows how these problems dissolve under a Merge-based definition of roots.

---

[12] 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 also note 8 for discussion.
[13] This statement should of course be relativized to cyclic domains: within one cyclic domain functional projections dominate lexical ones. This is an issue we return to in detail in section 5.2.
[14] Note in this respect that Borer (2009b:1) has to state as an axiom that "Extended projections must have an L core".

2.5 Summary: desiderata for a theory of roots

The past four sections 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 of them have in common is that they have to be stated as primitives of the theory. They do not follow from any independent properties of our model of the grammar, and hence, they essentially constitute a list of four separate axioms or stipulations (depending on one's perspective). Formulated more positively, what we would like our grammar to generate is a structural position at the beginning of the derivation that is acategorial and radically featureless, where root material can be inserted during (late) Vocabulary Insertion. In the next section 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

This section introduces a theory of Merge which the next section will use to derive the four basic properties of roots outlined above. We proceed in two steps. Section 3.2 argues, following Jaspers (1998), Langendoen (2003), Zwart (2009a, 2009b, 2010) and others, that Merge is inherently asymmetric and that pair merge rather than set merge is the default—and arguably 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. Section 3.3 sums 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, i.e. $\{\alpha, \beta\}$ (see e.g. Chomsky 1995:243). When considering only 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.[15] More generally, Merge (α, β) = Merge (β, α). In the remainder of this section, however, we present a number of arguments—taken from Chomsky (1995), Jaspers (1998), Langendoen (2003), and Zwart (2009a, 2009b, 2010)—suggesting that the picture just sketched is too simple, and that there is an inherent asymmetry to Merge. Accordingly, we will adopt Zwart's (2010:7) conclusion that pair Merge rather than set Merge is the basic structure-building principle of natural language.

The first complication is highlighted by Chomsky himself. After introducing the definition given above, he points out that output conditions dictate that mere set formation does not suffice (Chomsky 1995:243). Given that different types of constituents (e.g. verbal and nominal ones) are interpreted differently at LF and PF, 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 insight by assuming that either α or β functions as the label of the

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

newly formed constituent, i.e. Merge (α, β) = $\{\alpha, \{\alpha, \beta\}\}$ (with α the label of the complex constituent). As pointed out by Langendoen (2003:3), however, $\{\alpha, \{\alpha, \beta\}\}$ is set-theoretically equivalent to the ordered pair $<\alpha, \beta>$.[16] In other words, by admitting that one of the two elements combined by Merge projects, we are led to the conclusion 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, i.e. the fact that for every Merge operation, one element is 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. It is clear that an implementation of Merge in terms of pure set formation does not succeed in capturing Derivational Asymmetry as set merge is a strictly symmetrical operation.

Thirdly and finally, Zwart (2009b: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 (2006:5) 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 paper we adopt and adapt Zwart's (2009a, 2010) definition. It is given in (12).

(12) **Unary Merge** (pre-final version)
Merge selects a single element from a resource[17] and includes it in the object under construction.

Zwart 2009a:62, 2010:7 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 $<\alpha, \delta>$.

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 notion "element" refers to in this definition. It is an assumption rarely made explicit in current (morpho)syntactic literature (though see Chomsky 1995:383n27, Drury 1998:76n20, Fortuny 2008:18) that the objects combined by Merge are *sets* of features.[18] For example, under the (common) assumption that T° is a combination of person, number, gender en 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:

---

[16] The most standard definition of ordered pairs is that of Kuratowski (1921) given in (i) below (see also Partee, Ter Meulen & Wall 1987:27). The definition referred to by Langendoen (2003) 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—i.e. $<a, b> = <c, d>$ if and only if a=c and c=d—see http://us.metamath.org/mpegif/opthreg.html.
(i) $<a, b> =_{def} \{\{a\}, \{a, b\}\}$

[17] 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.

[18] In DM the operation the operation responsible for creating these sets is called fusion or bundling, see Marantz 2006, Bobaljik & Thraínsson 1998, Pylkkänen 2002, De Belder to appear for discussion and examples.

(13) **Unary Merge** (final version)
Merge selects a single subset from a resource (e.g. {α}), includes it in the derivation under construction (δ), and yields an ordered pair (e.g. <{α}, δ>, assuming {α} projects).

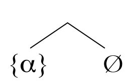
With this discussion as background, we are now ready to turn to the first application of Merge.

3.3 Primary Merge

An application of Merge that usually does not get a lot of attention in the literature 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 (see also Zwart 2010:8). 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 25?—thus straying from the Strong Minimalist Thesis (Chomsky 2000). Moreover, it is not clear how the selection of two objects can be limited to only the first step of the derivation. One 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 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. Just like the first scenario, this one has little appeal to it, as we are now introducing into the system a substantial degree of look-ahead and a concomitant increase in computational workload.

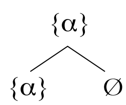
Part of this problem is alleviated under the perspective sketched in the previous section. As is clear from the definition of Unary Merge in (13), this operation always and without exception targets a single element from the resource. What remains to be determined, then, is what it means to be 'included 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 {α} is the first one to be taken from the resource by Unary Merge, it is included into an empty derivation, i.e. the object under construction is the empty set Ø (see also Zwart 2010:8). The output of this instance of Merge is no different from any other: it yields an ordered pair, in this case <{α}, Ø>. 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 = {α, β} as the input for a derivation. The first instance of Unary Merge takes a subset from R, say {α}, and includes it in the object under construction, i.e. Ø, yielding the structure in (14).

(14)


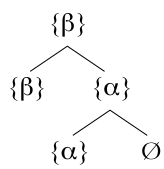
The question at this point is what the label of this complex constituent is. So far we have been proceeding under the assumption that merging an element from the resource to the derivation under construction yields an ordered pair of which that newly-merged element is the left-hand member. However, recall that in the previous section we converted Chomsky's labeled structure {α, {α, β}} into the ordered pair <α, β>. 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 (1995:244f) discerns two classes of situations. In the first, the structural configuration unambiguously identifies one of the two elements undergoing merge as the projecting member. These configurations include a head merging with a complex constituent (the head projects) and movement/internal Merge (the moving element does not project—though see Donati 2006 for a different view). In the other case—i.e. a head merging with a head or an XP externally merging with an XP—it is the featural specification (and possibly the concomitant checking relation) that determines which of the two will project. With respect to the structure in (14), we contend that it falls into the first category. That is, it is always unambiguously clear that {α} will project, regardless of its feature specification. The reason for this is twofold. First of all, Ø 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 on or copying information from a daughter node onto its mother, this radical emptiness makes Ø inherently incapable of projecting. Secondly, if Ø were to project, this would incorrectly identify the constituent in (14) as an empty derivation, thus obscuring the fact that (the non-empty element) {α} has already been merged into that derivation. Summing up, then, the complete representation of the Primary Merge operation described above is as in (15).

(15)



At this point, Unary Merge takes {β} from R and merges it with the structure in (15). Supposing that {β} is the projecting element, this Merge operation yields the ordered pair <{β}, <{α}, Ø>>, which can be graphically represented as in (16).

(16)



In short, the theory of Merge developed in this and the preceding section entails that every derivation begins with the merger of a radically empty element, which accordingly sits at the very bottom 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 (2009b, 2010) also assumes that Primary Merge involves "merger with nothing/the empty set" (Zwart 2010:10). His implementation, however, differs from ours in two ways. Firstly, his system is strictly top down—or rather, left-to-right: what Merge does, is split 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 (see e.g. Zwart 2009b:164). As has become clear from the above discussion (and see also the next section), in our proposal the position created by merger with the empty set is real and plays a central role in natural language.[19]

The idea that Primary Merge involves merger with the empty set is also found in 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, in order for Merge to be successive, 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 Ø. When Ø 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 singleton 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 was the case in Zwart's system, the impact of the initial empty set on the remaining derivation is non-existent. This is exactly where Fortuny's analysis differs from ours: while we agree that Primary Merge involves the null set as one of its members, we contend that this empty position is a syntactic terminal that receives a phonological exponence in the post-syntactic morphological module (see below, section 4).

3.4 Summary

In this section we have introduced and discussed our theory of Merge. Following Jaspers (1998), Langendoen (2003), Zwart (2009a, 2009b, 2010) and others, we have 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. In the next section we argue that this is where roots are inserted in the post-syntactic module.

4. Deriving the properties of roots

4.1 Introduction

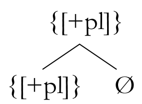
This section combines the insights of the previous two. Section two introduced four central properties of roots and showed that while empirically well supported, they do not follow from anything in the theory and hence have to be stipulated as axioms. The previous section led to the conclusion that Merge 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. We first (in section 4.2) go through a sample derivation of a nominal constituent to illustrate how exactly the theory of Merge interacts with that of roots, while section 4.3 returns to the basic properties of roots outlined in the first half of this paper.

---

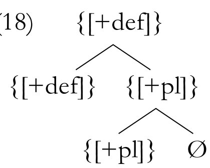
[19] As Zwart (2010:6) points out, the LCA might necessitate positing an empty position independently of any considerations related to Merge. Given that this principle depends on asymmetric c-command in order to convert hierarchical structure into linear order, the merger of two non-branching nodes yields a non-linearizable structure, i.e. precisely the kind of structure that is created by Primary Merge. Chomsky proposes that such configurations must be rendered asymmetric (e.g. via movement) before they reach the PF-interface (see also Moro 2000 for related discussion), but more generally, one might require that every right-branching structure ends in an empty position. The fact that the perspective on Merge developed in this paper yields precisely this result on independent grounds thus serves as additional support for our proposal.

4.2 Asymmetric Primary Merge of roots: a sample derivation

This section presents a sample derivation of the nominal constituent *the books* as an illustration of how our view on Merge meshes with the syntax of roots. Recall that we adopt Late Insertion of vocabulary items. 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 (see also notes 8 and 12). For the example at hand, this means that the resource is a set containing a definiteness feature and a plural feature, i.e. R = {[+def], [+pl]}.[20] Based on this resource, the derivation proceeds as follows. Unary Merge first selects the singleton containing the plural feature from R 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).

(17)


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

(18)


At this point, the syntactic derivation is finished and the structure is handed over to PF. One of the operations on the way to the interface with the articulatory-perceptual system is Vocabulary Insertion (see Harley & Noyer 1999 for discussion). When confronted with the structure in (18) the grammar searches its Vocabulary for matching vocabulary items and it encounters the following VIs:[21]

(19)	a.	/ðə/	↔	[+def]
	b.	/s/	↔	[+pl]
	c.	/buk/	↔	Ø

The phonological exponents on the left-hand side of the equivalences in (19) are inserted into the terminal nodes of the structure in (18), and the derivation converges as the nominal constituent *the books*.[22]

---

[20] For the sake of exposition, we are abstracting away from any other grammatical features that might underlie the DP *the books*. For example, most DM-analyses would include a *n*-head in R.

[21] We return in detail to the properties of the vocabulary insertion mechanism in section 5.3.

[22] 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 & Noyer 2001 for extensive 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 have glossed over it for expository purposes.

4.3 Returning to the four root axioms

The derivation in (17)-(19) demonstrates how the empty position created by asymmetric Primary Merge can serve as an insertion site for roots. This section returns to the four root axioms introduced in section two and shows how they follow straightforwardly 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 (20)c 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 marking (featural or otherwise) on a specific subset of lexical items. 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 (cf. (20)a). Regardless of which vocabulary item is inserted in this position in the morphological component (and see below, section 5.3 for extensive discussion of the insertion mechanism), during the syntactic derivation this position will be completely inert.

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

Finally, the fact that roots are merged lower than functional material is also an integral part of our analysis. The only stage at which the derivation is null is the very beginning and so the only Merge operation that can involve Ø as one of its members is Primary Merge. Given that the occurrence of roots is directly dependent on merger with Ø, this implies that roots are never inserted in mid-derivation, i.e. dominating previously merged functional material (see below, section 5.2, on derivations containing multiple roots).

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

Section 4 has shown how the basic properties of roots can be reduced to theorems of 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 (cf. also Zwart 2009b, 2010). We explore the consequences of this analysis and conclude that derivations proceed left-to-right across cyclic domains, but bottom-up within each cyclic domain. In section 5.3 we examine the mechanism underlying Vocabulary Insertion. We argue for a single competition principle regulating the insertion of vocabulary items both in root terminal nodes and functional terminal nodes and show how this principle allows for the insertion of both lexical and functional material in root terminal nodes. Section 5.4 summarizes.

5.2 Derivations with multiple roots: layered derivations

5.2.1 Introduction

Recall from section three that Primary Merge merges a subset of the resource to the null state of the derivation—or more technically, to the empty set. This empty syntactic terminal then serves as the insertion site for a root at the post-syntactic level of Vocabulary Insertion. Given that a derivation 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 can host and the number of derivations it is the output of. We can formulate this as in (21).

(21) **One Derivation One Root (ODOR)**
For every derivation there is exactly one root, and for every root there is exactly one derivation.

Needless to say, this principle raises analytical questions with respect to strings containing multiple roots such as the VP in (22)a or the clause in (22)b. The ODOR-principle predicts that (22)a should be the result of two independent derivations (one for each root), while (22)b should contain three. More generally, the approach developed so far forces us to adopt the concept of layered derivations (in the sense of Zwart 2009b, 2010), whereby the output of one derivation can appear as an atom in the next one.

(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 vocabulary 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 (1999) multiple spell-out model whereby all non-complements—including, in our view, the *v*-V-complex—are spelled out top-down—or rather, left-to-right—in separate derivations, until finally the root derivation is constructed. Section 5.2.4 sums up.

5.2.2 Layered derivations as readmittance to the resource

Zwart (2009b: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 one derivation to be readmitted to the resource from which its members were originally drawn. In order to see how this would work, we will go through a sample derivation of the VP in (22)a. 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 the resource in our analysis contains no information about roots.)

(23) R = {*v*, [+def]}

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

(24) {[+def]}
{[+def]} Ø

(25) R = {*v*}

At this point the initial derivation is stuck. In particular, Merge cannot append {*v*} to the structure in (24) as that would cause a feature conflict between *v* and [+def]—or alternatively, a violation of the selectional requirements of *v*. 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).[23] The solution lies in layering the derivation, i.e. in readmitting the object constructed in (24) to the resource. This yields the following representation for R:

(26) R = {*v*, <{[+def]}, Ø>}

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

(27) {*v*}
{*v*} Ø

(28) R = {<{[+def]}, Ø>}

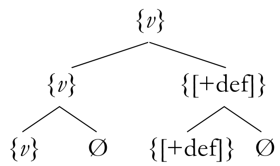
Merge now targets the one remaining (complex) element in R and appends it to the structure in (27), yielding the representation in (29).[24,25]

[23] More generally, what this illustrates is that our conception of Primary (and Unary) Merge prohibits multiple configurations to be built in tandem. For argumentation that a single tree workspace is to be preferred over a multiple tree workspace, cf. Postma & Rooryck in progress.

[24] Note that the representation in (29) glosses over one complication, i.e. the fact that Merge should not select the member of R in (28) as is, but rather the singleton set containing that member (cf. the definition in (13)). We return to this issue below.

[25] 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—i.e. readmitting it to the resource and then appending it to the null derivation via

(29)



Given that R is now empty, the derivation can be handed over to the post-syntactic module. At the level of Vocabulary Insertion, the vocabulary items *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 Agree-based analysis of roots can yield surface strings containing more than one root.

A key ingredient of the analysis just presented is the assumption that the vocabulary resource can contain elements that are the output of preceding derivations and hence syntactically complex. This assumption is corroborated by the discussion in Ackema & Neeleman (2004:122-129) (see also Zwart 2009b:172). Ackema & Neeleman present a host of data, all of which involve syntactically complex phrases as parts of words. Some relevant examples are given in (30).

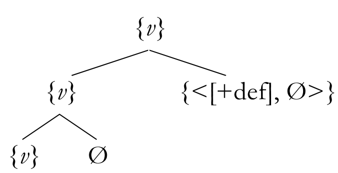
(30) a. a sit-on-the-guidelines Euro policy
b. animal-to-human transplant experiments
c. go-anywhere-at-any-time-access
d. I feel particularly sit-around-and-do-nothing-ish today.

Summing up, the ODOR-principle in (21), which is a logical consequence of our theory of roots, leads to the conclusion that a VP such as the one in (22)a contains two derivations. This subsection has provided a first, simple implementation of this idea, based on Zwart's (2009b, 2010) notion of layered derivations. The next subsection adds a complication to this picture, viz. the idea that layered derivations create opacity effects.

5.2.3 Layered derivations and opacity

As was pointed out in footnote 24, the structure in (29) glosses over the fact that the operation Select does not target elements from the resource directly, but rather subsets of such elements. Accordingly, a more accurate representation would be the following:

(31)



The transition from (29) to (31) is not innocuous. Given that {<[+def], Ø>} is not of the form {α, {α, β}}—i.e. the output of an (asymmetric) Merge operation—it is not recognized bij C_{HL} as a syntactically complex element, and instead is treated as an atom (see also Uriagereka 1999:6). Put differently, the internal structure of

Primary Merge—would only cause the object to project further, and would not yield a verbal structure. For extensive argumentation that roots indeed have no argument structure, see Borer 2009b.

this constituent has become opaque to any further syntactic operations (see also Zwart 2009b:173 on layered derivations creating opacity effects). Given that direct objects such as *the cookie* in the VP in (22)a are generally not islands for extraction, this implies that the derivation presented in the previous subsection cannot be correct. What we propose to change is the order in which the various subderivations are readmitted to the resource. In order to illustrate this, we turn to the more complex example in (22)b, repeated below.

(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 starts out as in (33).

(33) R = {*v*, [+def], [+def]}

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

(34) {[+def]}

{[+def]} Ø

(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 (2009b:175-178) argues explicitly 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 of the present paper, this means that *v* is merged with the null state of the derivation, yielding that structure in (36) and the modified resource in (37).

(36) {*v*}

{*v*} Ø

(37) R = {<{[+def]}, Ø>, [+def]}

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

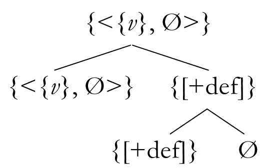
(38) {[+def]}

{[+def]} Ø

(39) R = {<{[+def]}, Ø>, <{*v*}, Ø>}

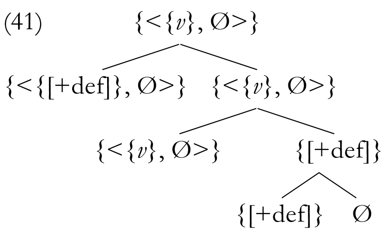
The DP in (38) is then merged with the *v*-V-complex, as in (40).

(40)



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

(41)



Given that R is now depleted, the syntactic derivation is finished and the structure can be spelled out. At the level of Vocabulary Insertion, the syntactic terminals of (41) are matched up with the appropriate 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 entails that everything but the most deeply embedded internal argument is spelled out in a separate derivation in a top-down or left-to-right fashion, but that within each (sub)derivation structure building proceeds strictly bottom-up. It is only when then most deeply embedded internal argument is reached that all subderivations are brought together into one syntactic representation.[26] It is worth pointing out that the model of the grammar just described bears a very close resemblance to the one Uriagereka (1999) and Drury (2005) arrive at on independent grounds.

Summing up, under the (plausible) assumption that layered derivations create opacity effects, the analysis proposed in the previous subsection had to be revised. We have shown that the set of assumptions adopted in this paper lead to a model of the grammar in which syntactic computation proceeds bottom-up within derivations, but left-to-right across derivations.

5.2.4 Conclusion

If the insertion of roots is a by-product of asymmetric Primary Merge, then each derivation contains precisely one root. Accordingly, strings containing more than one root have to be the result of multiple—or more accurately, layered—derivations. This subsection has provided a concrete implementation of how such a multi-root derivation proceeds. The next section turns to the insertion mechanism responsible for assigning a phonological exponence to syntactic terminals.

---

[26] 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 paper. See Uriagereka 1999, Dury 2005, Van Gelderen 2003, and Hoffman 1996, among others, for relevant discussion.

5.3 The insertion mechanism: competition for root positions

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 vocabulary items (VIs). This issue is complicated by the fact that unlike previous Late Insertion models, we have claimed that functional vocabulary items (FVIs) can be inserted in root positions (cf. the data in (2)-(11)). Most DM-accounts of Vocabulary Insertion assume a radically different insertion mechanism for functional and lexical vocabulary items (LVIs). While the former are subject to competition (the VI most closely matching the syntactic terminal in features being the preferred insertion candidate), insertion of the latter is determined by free choice. This dual insertion mechanism presupposes a strict division of labor between the functional and the lexical domain: FVIs always and only spell out functional terminal nodes, whereas LVIs always and only spell out root terminal nodes. As such, we cannot adopt it here. What we propose instead is a single, unified insertion mechanism in which all Vocabulary Insertion is regulated by competition.

5.3.2 Vocabulary Insertion in Distributed Morphology

As was pointed out above, DM assumes different modes of insertion for FVIs and LVIs. The insertion of FVIs is regulated by competition. More specifically, they are inserted into functional terminal nodes along the lines of the Subset Principle in (42) (Halle 1997, Kiparsky 1973, Anderson 1986).

(42) **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 (Halle 1997:428).[27]

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). If the Subset Principle does not determine the winner (e.g. because two VIs are an equally close match for a particular terminal node), one can simply stipulate one (aka extrinsic ordering, see Halle and Marantz 1993) or rely on a Universal Feature Hierarchy to choose one of the two (Harley 1994).

LVIs on the other hand do not compete. Their insertion is based on free choice (Harley and Noyer 1998). That said, the precise way in which this free choice is implemented tends to vary. For example, Harley and Noyer (1997) 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 & Noyer (1999) suggest that LVIs are endowed with the feature [Root] or with an index. What all of these options have in common however, is that all LVIs have some

[27] Note that Halle uses the term '(terminal) morpheme' for what we have been calling—and will continue to call—'(syntactic) terminal node'.

sort of marking that they share with root terminal nodes, and it is this marking that allows all of them to be inserted freely in such positions.

From the point of view of the current paper, this two-sided insertion mechanism is problematic for three reasons. First of all, we have shown, based on the data in (2)-(11), that FVIs can be inserted in root terminal nodes. This implies that whatever diacritic allows LVIs to be inserted in such positions—say, a [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 as such would no longer distinguish roots from non-roots. Secondly, the use of FVIs as roots considerably complicates the insertion mechanism itself. Note that the feature specification of an FVI—even if it were endowed with a [Root]-feature or other such diacritic—would never be a subset of the set of features present on root terminal nodes, as these positions do not contain any grammatical features.[28] Thirdly, if FVIs were endowed with a [Root]-feature, then by the Subset Principle this feature should be present on functional terminal nodes as well (thus further hollowing out the concept of a [Root]-feature). This would imply that LVIs should be able to be (freely) inserted in functional terminal nodes as well, contrary to fact.[29]

Summing up, while the DM-approach to Vocabulary Insertion works well in a world where the functional and the lexical realm are strictly separate, it faces considerable problems in light of data such as (2)-(11), where a functional element shows up in a typically lexical context. The next section shows how a unified insertion mechanism based on competition can overcome these problems.

5.3.3 Unified insertion through competition

This section introduces a mechanism of Vocabulary Insertion that applies both to FVIs and to LVIs. The key ingredient will be the assumption that all insertion—whether it be in functional or root terminal nodes—is driven by competition. In a nutshell, what we propose is that in functional terminal nodes, Vocabulary Insertion proceeds much as claimed in DM, while for root terminal nodes, all vocabulary items (both FVIs and LVIs) prove to be an equally close match, thus creating the illusion of free choice.

Let us take as our starting point the Subset Principle in (42) and see how it fares when applied to *all* cases of Vocabulary Insertion (rather than just the insertion of FVIs in functional terminal nodes). The Subset Principle proceeds in two steps. First, it selects all possible candidates, i.e. all VIs whose feature specification forms a subset of the set of features on the terminal node. Second, the most optimal candidate is selected from this group: the VI that provides the closest match for the features on the terminal node.

As far as insertion in functional terminal nodes is concerned, this mechanism yields the correct results, even if we include LVIs in the list of possible insertion candidates. A concrete example can serve to illustrate this. Suppose a functional terminal node F bears the following feature specification: [+3rd, +sg, +masc, +past]. Under the reasoning outlined above, not only FVIs such as the ones in (43)a-c are possible insertion candidates, but also featureless LVIs such as *book* in (43). In fact, given that all LVIs are featureless, they are all potential realizations of F. When it comes to selecting the actual vocabulary item to spell out F, however, all

[28] It is worth pointing out that Caha's (2009) Superset Principle does not fare any better in this respect. This principle chooses the insertion candidate from those VIs whose feature set is a *super*set of the set of features present on the terminal node. In the hypothetical situation discussed in the main text, this would mean all VIs are possible candidates for insertion in root terminal nodes. However, given that LVIs by definition are a closer match for root terminal nodes that FVIs (neither LVIs nor root terminal nodes contain grammatical features), the latter will never surface in root position, contrary to fact.

[29] Regarding so-called semi-lexical items such as classifiers or measure nouns, we follow De Belder (to appear), who argues they are merely a subtype of FVIs.

FVIs in the candidate set will be closer matches for F's feature set than the featureless LVIs. As a result, no LVIs are (correctly) predicted to surface in functional terminal nodes.

(43)	a.	/wɔz/	↔	[+3rd, +sg, +masc, +past]	(FVI)
	b.	/hɪm/	↔	[+3rd, +sg, +masc]	(FVI)
	c.	/t/	↔	[+past][30]	(FVI)
	d.	/buk/	↔	Ø	(LVI)

In short, the Subset Principle as outlined in (42) straightforwardly yields the correct result for Vocabulary Insertion in functional terminal nodes, even if both FVIs and LVIs are taken to be potential insertion candidates. In root terminal nodes, however, the principle is less successful. Recall that such positions are radically empty—they are the result of asymmetric Primary Merge with the null state of the derivation. This means that the only insertion candidates allowed by the Subset Principle are featureless LVIs (only the empty set is a subset of the empty set). As such, the principle incorrectly predicts data such as those in (2)-(11) to be ill-formed.

What we need, then, is an insertion mechanism that retains the effects of the Subset Principle for functional terminal nodes, but yields different results in the case of root terminal nodes. More specifically, the (possibly empty) set of features present on a terminal node should act as a *filter* for the VIs that can be inserted in that position: if the terminal node contains features (i.e. in the case of functional terminal nodes), only VIs matching those features are retained, while if the terminal is featureless (i.e. is a root terminal node), it imposes no restrictions on the VI that can be inserted there—the filter is vacuous—and any VI is a possible insertion candidate. The Revised Subset Principle in (44) has precisely this effect:

(44) **The Revised Subset Principle**

Given a terminal node A with feature set F_0 and vocabulary items (VIs) $/B_{1,2,\ldots,n}/ \leftrightarrow F_{1,2,\ldots,n}$:
$/B_i/$ is inserted in A if $F_0 \times F_0 \subseteq F_0 \times F_i$. When several VIs meet this condition, the one for which $F_0 \times F_i$ most closely matches $F_0 \times F_0$ is chosen.

This principle states that the phonological exponent of a VI is inserted in a terminal node if the Cartesian product of the feature set of the VI and the feature set of the terminal node matches all or a subset of the ordered pairs of the Cartesian product of the feature set of the terminal node with itself. Insertion does not take place if the Cartesian product of the feature set of the VI and the terminal node contains ordered pairs not present in the Cartesian product of the feature set of the terminal node with itself. Where several VIs meet the conditions for insertion, the VI that yields the greatest number of matching pairs must be chosen. As we will now illustrate, this principle selects the same VI in the case of functional terminal nodes, but leads to a universal tie in the case of root terminal nodes. Suppose F is a functional terminal node with the (abstract) feature specification F_0 = {a, b}, and suppose furthermore that the lexicon of this hypothetical language contains only the four following VIs listed in (45). As is clear from their feature specification, the first is an LVI, while the latter three are FVIs.

[30] Intended here is the past tense suffix *-ed*.

(45) a. /bik/ ↔ Ø (LVI)
b. /ta/ ↔ {a} (FVI)
c. /plo/ ↔ {a, b} (FVI)
d. /stu/ ↔ {a, b, c} (FVI)

Recall from the highly parallel example in (43) that the Subset Principle would select /plo/ as the VI that realizes the functional terminal node F_0, as this is the VI whose feature set is the most closely matching subset of the feature set of F. The Revised Subset Principle behaves identically. Given that it involves not just first-order sets, but Cartesian products of such sets, let us first make explicit what the terms of the comparison are. They are listed in (46) and (47).

(46) $F_0 \times F_0 = \{a, b\} \times \{a, b\} = \{<a,a>, <a,b>, <b,a>, <b,b>\}$

(47) a. $F_0 \times F_{/bik/} = \{a, b\} \times \varnothing = \varnothing$
b. $F_0 \times F_{/ta/} = \{a, b\} \times \{a\} = \{<a,a>, <b,a>\}$
c. $F_0 \times F_{/plo/} = \{a, b\} \times \{a, b\} = \{<a,a>, <a,b>, <b,a>, <b,b>\}$
d. $F_0 \times F_{/stu/} = \{a, b\} \times \{a, b, c\} = \{<a,a>, <a,b>, <b,a>, <b,b>, <a,c>, b,c>\}$

The sets in (47) each have to be compared with the one in (46). In a first step, only those that form a subset of $F_0 \times F_0$ are retained as possible insertion candidates. This procedure eliminates set $F_0 \times F_{/stu/}$ in (47)d. Secondly, the set matching $F_0 \times F_0$ most closely is chosen as the actual realization of F. Given that $F_0 \times F_0$ is identical to $F_0 \times F_{/plo/}$ in (47)c, this VI wins the competition. More generally, in the case of Vocabulary Insertion in functional terminal nodes the Subset Principle and the Revised Subset Principle yield the same output.

Root terminal nodes are a different story altogether, however. Recall that they are radically featureless. This means that in this case $F_0 = \varnothing$. The sets that are being compared by the Revised Subset Principle are listed in (49).

(48) $F_0 \times F_0 = \varnothing \times \varnothing = \varnothing$

(49) a. $F_0 \times F_{/bik/} = \varnothing \times \varnothing = \varnothing$
b. $F_0 \times F_{/ta/} = \varnothing \times \{a\} = \varnothing$
c. $F_0 \times F_{/plo/} = \varnothing \times \{a, b\} = \varnothing$
d. $F_0 \times F_{/stu/} = \varnothing \times \{a, b, c\} = \varnothing$

Given that the Cartesian product of any set with the empty set yields the empty set, all sets in (48)/(49) are identical and more importantly, all sets in (48) are not only (trivially) subsets of the one in (49), they are also all the most closely matching subset. In other words, the Revised Subset Principle predicts that in the case of root terminal nodes there is a universal tie between VIs, and all of them—LVIs and FVIs alike—are potential candidates for insertion, exactly as required.

Summing up, we have proposed a unified insertion mechanism for both LVIs and FVIs that is based on competition. In the case of functional terminal nodes this mechanism works exactly like the traditional Subset Principle, but for root terminal nodes it leads to a universal tie, thus allowing all VIs—i.e. not only LVIs but also FVIs—to be inserted in that position and essentially creating free choice. Our Revised Subset Principle thus captures the intuition that the feature specification of a terminal node acts as a filter on the type of VI that can be inserted there. If the filter is vacuous, all VIs meet the requirement.

5.4 Conclusion

This section has explored two important consequences of the theory outlined in the first four sections of this paper. 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 for a single Vocabulary Insertion mechanism based on competition that allows FVIs to be inserted in root positions, but not the other way around.[31] The next section addresses two potential threats for our account.

6. Potential threats for our analysis: arguments for early insertion

6.1 Introduction

As was pointed out in section 2.4.2 there is a close link between a structural definition of roots and the Late Insertion hypothesis. Given that our analysis of roots has been a structural one *par excellence*—root terminal nodes are the mechanical by-product of asymmetric Primary Merge—late insertion of vocabulary items forms an integral part of our account. In this respect it is worth pointing out that recent analyses tend to allow a limited type of early insertion, typically of an index. For example, if a root will be realized by the LVI *cat*, an index corresponding to that particular VI, say √283, is merged in narrow syntax (thus creating a root terminal node). This index bears no phonological or semantic relation to the actual VI, but is a mere placeholder signaling that this specific VI will be inserted in that position. Given that this line of reasoning is directly at odds with the approach developed in this paper, we want to address the two main arguments in support of early index insertion here. We do so in the next two sections (see also Siddiqi 2006 for discussion).

6.2 Harley 2009

Standard DM makes a distinction between readjustment rules and suppletion. The former alter the phonological exponence of a single VI after vocabulary insertion. For example, the fact that the final rime of *destroy* is replaced by *-uct* in nominal contexts is due to a readjustment rule. Suppletion, on the other hand, is the result of two distinct VIs competing for insertion into the same syntactic terminal. Suppletive verb forms such as *go* and *went* are thus not assumed to be exponents of a single VI. The crucial criterion for distinguishing between readjustment rules and suppletion is phonological similarity: when two (morphosyntactically related) lexical items share a portion of their phonology, the relation between them is assumed to be not suppletive, but due to readjustment. Hence, the *think-thought* alternation is due to a readjustment rule (because the two forms share an onset), but that between *am* and *is* is suppletive. Marantz (1997) further proposes that suppletion is an exclusive property of *functional* VIs. His reasoning goes as follows: given that suppletion involves competition between VIs and given that only the insertion of functional VIs is regulated by competition in standard DM (see above, section 5.3, for discussion), only functional VIs can be suppletive.

---

[31] There is one scenario in which an LVI could be inserted in a functional terminal node, and that is when the language does not contain any FVI that bears (a subset of) the features present on that terminal node. This might be a possible source for language change, in particular grammaticalization.

It is this last conclusion that Harley 2009 provides counterevidence against. She presents data from Uto-Aztecan languages showing that lexical verbs can be suppletive depending on the number specification of their arguments. For example, the Hiaki verb meaning 'to wander' can surface either as *weama* or *rehte*. Given that these forms do not share any phonological characteristics, the reasoning developed above would classify them under suppletion, not readjustment. Crucially, however, the verbs undergoing this process are not functional. For example, Hiaki suppletive verbs can denote 'to fall in water', 'to stampede', 'to make a netbag', etc. These data thus falsify Marantz's (1997) claim that suppletion is restricted to the functional domain.

Harley's findings are problematic for the traditional DM approach to Vocabulary Insertion. Recall that Marantz linked suppletion to functional VIs precisely because the latter are inserted through competition. Root terminal nodes on the other hand receive their phonetic exponence through free choice between lexical VIs—possibly mediated through a [Root]-feature, see above. As such, there is no mechanism to allow two suppletive VIs to compete for insertion into a single root terminal node.

Harley proposes to solve this problem via a limited form of early insertion. More specifically, she assumes that an abstract index is merged in syntax in root positions, an index that singles out a particular VI but bears no phonological or semantic relation to it. The advantage of this approach is that one can now specify insertion contexts for the various phonological exponents of this index. For example, supposing the index for the Hiaki root 'to wander' would be 361, its VI would look as follows:

(50) √361 ↔ *weama* / [___ DP_{sg}]
↔ *rehte* elsewhere

The VI states that the root bearing the index 361 will be realized *weama* in the context of a singular DP-argument, and as *rehte* in all other contexts. More generally, by adopting early insertion of indices, Harley succeeds in introducing a limited form of competition in the vocabulary insertion of root terminal nodes. Given that suppletive contexts seem to require such competition, this is a welcome result.

Note, however, that Harley's analysis faces the same problem that other DM-accounts of Vocabulary Insertion do (see above, section 5.3.2), i.e. it draws too strong a dividing line between the lexical and the functional domain. In particular, while Vocabulary Insertion in the lexical realm is driven by indices, functional terminal nodes specify morphosyntactic features that the VIs inserted in them have to express. In order to make this more concrete, compare the lexical VI in (50) with the functional one of the demonstrative *those* in (51).

(51) [+D, +def, +distal, +pl] ↔ /δoz/

Vocabulary insertion into root terminal nodes is thus completely separated from vocabulary insertion into functional terminal nodes: while the former is based on and driven by indices, the latter is regulated by formal features. As we have shown in the beginning of this paper, however, such a strict separation is empirically untenable, as functional VIs can be inserted productively into root terminal nodes (see the data in (2)-(11) and surrounding discussion).

What we propose instead is to give up the distinction between readjustment and suppletion, and to allow non-phonologically related forms—such as the Hiaki verb forms *weama* and *rehte*—to be derived via readjustment. After all, if such rules are able to derive *thought* from *think* and *destruct* from *destroy*, why not *was* from *is* or *weama* from *rehte*? The assumption that the criterion distinguishing suppletion from readjustment is phonological resemblance is purely stipulative, and the only attempt we are aware of of attaching theoretical significance to the distinction—Marantz' claim that suppletion is the prerogative of the functional domain—has been clearly falsified by Harley 2009. Under the assumption that the *weama-rehte* alternation is the result of readjustment, Vocabulary Insertion can proceed exactly as outlined in the preceding sections: Asymmetric Primary Merge yields an empty syntactic terminal at the bottom of the syntactic representation, into which any VI can be inserted (see above, the Revised Subset Principle in section 5.3.3). When the VI meaning 'to wander' is selected, it surfaces as *rehte*, unless when selecting a singular DP. In the latter context it undergoes readjustment to *rehte* (just like *think* undergoes readjustment to *thought* in a past tense context).

Summing up, then, the Uto-Aztecan data discussed by Harley 2009 might be problematic on traditional DM-assumptions about Vocabulary Insertion, but they do not threaten the current proposal, nor do they not necessitate early insertion of root indices.

6.3 Pfau 2009

6.3.1 Introduction

Pfau (2009) argues for the early insertion of indices based on psycholinguistic research into speech errors. The kind of data he focuses on involve cases where a VI different from the intended one is used, as in (52) (Pfau 2009:87).

(52)	In	welcher	Höhe,	äh, Tiefe	haben	sie	gegraben?	[German]
	in	what	height	er depth	have	they	digged	
	'In what height, er, depth, did they dig?"							

In this example the speaker intended to talk about the depth (*Tiefe*) of the dig, but erroneously uses the VI *Höhe* 'height'. Pfau uses such data to establish the relative order of certain grammatical operations. Two of his findings are relevant in the context of this paper. The first concerns his claim that conceptual information must be present prior to Vocabulary Insertion, the second involves cases of gender agreement with the intended (rather than the actually inserted) VI. We discuss these two cases in turn in the next two subsections. Our line of argumentation will be the same in both. While we agree with Pfau's conclusion about the relative ordering of grammatical operations, we do not agree with this conclusion about absolute ordering. In particular, we see no compelling arguments in his data for endowing root positions with featural content in or prior to narrow syntax. As as a result, Pfau's findings will not prove to be problematic for our approach.

6.3.2 Meaning-based speech errors and conceptual information

Meaning-based speech errors involve cases whereby the wrongly inserted VI stands in a certain semantic relation—hyponymy, hyperonymy, antonymy, synonymy, etc.—to the intended VI. The error in (52) was a first illustration of this—height and depth being antonyms—another one can be found in (53) (Pfau 2009:87). Here, the intended noun is *Spitzer* 'pencil sharpener', which is a cohyponym of the wrongly inserted *Radiergummi* 'eraser'.

(53) Hast du einen Radiergummi da? (INTENDED: einen Spitzer)
have you an eraser there a pencil.sharpener
'Do you have an eraser?'

Pfau adopts from Levelt et al. (1999) the idea that a specific module, i.e. the Conceptualizer,[32] activates lexical concepts according to a preverbal message intention. This module contains a network of concepts. Pfau concludes that in order to account for meaning-based speech errors such as those in (52) and (53), the Conceptualizer must precede Vocabulary Insertion. He therefore assumes that the index of a root is part of the Numeration and hence present throughout the syntactic derivation.

As was pointed out above, there is a relative and an absolute claim in Pfau's argumentation, and while we agree with the former, we do not see any compelling evidence for the latter. Meaning-based speech errors indeed show that conceptual information is activated before or during Vocabulary Insertion, or to put it in Pfau's terminology: the Conceptualizer must indeed precede—or coincide with—Vocabulary Insertion. From this, however, it does not follow that this conceptual information must also be present in narrow syntax. Pfau's data are perfectly compatible with a model of the grammar in which a conceptualization module is sandwiched in between syntax and vocabulary insertion, feeding the latter but independent from the former. Meaning-based speech errors do not, then, provide a compelling argument for early index insertion and hence are not problematic from the perspective of this paper.

6.3.3 Form-based speech errors and gender accommodation

Pfau observes that there is a difference in gender accommodation between meaning and form-based speech errors. In the former case it is the wrongly inserted noun that controls gender agreement, while in the case of form-based errors it is the intended (yet initially unpronounced) noun that determines gender agreement on the determiner. The pair in (54)-(55) illustrates this dichotomy.

(54) Du musst die Tür dann festhalten, Quatsch, das Fenster
you must the$_{\text{FEM}}$ door$_{\text{FEM}}$ then hold rubbish the$_{\text{NEUT}}$ window$_{\text{NEUT}}$
'You'll have to hold the window then.' (Pfau 2009:344)

(55) Das ist immer der gleiche Chaos, äh, Kasus
that is always the$_{\text{MASC}}$ same chaos$_{\text{NEUT}}$ er case$_{\text{MASC}}$
'That's always the same chaos, er, case.' (Pfau 2009:125)

In (54) the intended noun *Fenster* 'window' is replaced by the cohyponym *Tür* 'door' and the determiner introducing the wrongly inserted noun (*die*) agrees with it in gender. In (55) on the other hand the intended noun *Kasus* 'case' is replaced by *Chaos* 'chaos' on purely formal grounds—both nouns are bisyllabic and have /ka/ as their first syllable—and the determiner agrees with the intended rather than the inserted noun.

Pfau proposes to analyze these data as follows. Meaning-based speech errors result from the insertion of a wrong, but semantically related, index in syntax, while in form-based errors syntax contains the correct index, but Vocabulary Insertion makes a mistake in selecting a (formally related, but) wrong VI. In both cases gender agreement between the determiner and (the index of) the noun takes place in narrow syntax. In meaning-based

[32] Pfau notes that the Conceptualizer may be identified as Encyclopedia in DM.

speech errors this will result in gender agreement with (the index of) the wrongly inserted noun, while form-based speech errors show agreement with (the index of) the intended noun.

Once again, there is a relative and an absolute side to Pfau's reasoning. On the one hand, he shows convincingly that while meaning-based speech errors have to precede (the mechanism regulating) gender agreement, form-based speech errors have to follow it. On the other hand, though, the conclusion that this ordering of operations implies that root indices have to be present in narrow syntax crucially depends on one's assumptions about where in the grammar (gender) agreement takes place. While for Pfau it is self-evident that such agreement takes place in narrow synax—in which case his absolute argument would indeed go through—we follow the line of thinking initiated by Bobaljik (2008), whereby agreement—or at the very least its precise morphological shape and value, see Van Koppen (2005)—is entirely a post-syntactic operation. This means that while the relative ordering of the two types of speech errors and (gender) agreement remains, its split-up into a pre- and a post-syntactic part is not warranted.

6.3.4 Conclusion

This subsection has discussed the data and conclusions from Pfau (2009) pertaining to Vocabulary Insertion and its purported arguments in favor of early index insertion. We have shown that while Pfau's claims about the relative ordering of grammatical operations are well-founded, there is no compelling reason to adopt their absolute counterparts.

6.4 Conclusion

This section has considered two potential threats for the account outlined in the first have of this paper. Both Harley (2009) and Pfau (2009) have presented arguments in favor of a limited type of early insertion, whereby the Numeration contains indices of the roots that will be inserted post-syntactically. We have examined these arguments in detail and have concluded that the data reviewed by Harley (2009) and Pfau (2009) are perfectly compatible with the theory outlined in the preceding sections, and hence, that they do not warrant any conclusions about early index insertion.

7. Conclusions and prospects

This paper has shown that four commonly assumed properties of roots follow without stipulation from an articulated theory of Merge. We have proposed to treat the very first instance of Merge as being identical to all other applications of this operation. As a side-effect of this uniformity, any and every syntactic derivation starts out with a structurally empty syntactic terminal in its most deeply embedded position. We have identified this position as being the insertion sites for roots, thus deriving their featurelessness and acategorial nature, while at the same time allowing for the insertion of functional vocabulary items in root terminal nodes. This last point was an important—and, as we have shown, empirically well-motivated—aspect setting our analysis apart from previous DM-based accounts of Vocabulary Insertion. It also necessitated a revision of the principle regulating such insertion (cf. our Revised Subset Principle), which had as a welcome side-effect that *all* vocabulary insertion is now subject to one and the same (competition-based) mechanism.

At the same time, our proposal has been shown to have far-reaching consequences in quite disparate portions of the grammar. It provides new evidence for the layered derivations approach to expressions containing more than root, for the post-syntactic nature of agreement, against the (stipulative) distinction between readjustment and suppletion, for the hypothesis that roots never directly take arguments, etc. Exploring some of these consequences in more detail is a task we look forward to taking up in future research.

8. References

Ackema, Peter & Ad Neeleman. 2004. *Beyond morphology. Interface conditions on word formation*. Oxford: OUP.

Anderson, Stephen R. 1986. Disjunctive ordering in inflectional morphology. *NLLT* 4:1-31.

Bobaljik, Jonathan. 2008. Where's Phi? Agreement as a Post-Syntactic Operation. In: Harbour, Daniel, David Adger & Susana Béjar (eds.) *Phi-Theory: Phi features across interfaces and modules*. Oxford: OUP, 295-328.

Bobaljik, Jonathan, and Höskuldur Thráinsson. 1998. Two heads aren't always better than one. *Syntax* 1: 37-71

Borer, Hagit. 2005a. *In Name Only: Structuring Sense vol I*. Oxford: OUP.

Borer, Hagit. 2005b. *The Normal Course of Events: Structuring Sense vol 2*. Oxford: OUP.

Borer, Hagit. 2009a. *Categorizing Roots*. Talk presented at Roots, Universität Stuttgart, June 10-12, 2009.

Borer, Hagit. 2009b. Roots and categories. Handout of a talk presented at the 19th Colloquium on Generative Grammar, Vitoria-Gasteiz, Faculty of Letters (Gradu Aretoa), University of the Basque Country.

Caha, Pavel. 2009. The Nanosyntax of Case. PhD Diss., University of Tromsø. lingBuzz/000956.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: The MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In: R. Martin, D. Michaels, & J.Uriagereka (eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: The MIT Press.

Chomsky, Noam. 2006. Approaching UG from below. Ms. MIT.

De Belder, Marijke. To appear. *On roots and categories*. PhD Diss., Utrecht University.

Doetjes, Jenny. 1997. *Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English*. PhD Diss., Leiden University.

Donati, Caterina. 2006. On WH-head movement. In: L. Cheng & N. Corver (eds.) *Wh-movement: moving on*. Cambridge, MA: The MIT Press.

Drury, John E. 1998. The promise of derivations: atomic Merge and Multiple Spell-Out. Ms. University of Maryland.

Drury, John E. 2005. *Alternative directions for minimalist inquiry: expanding and contracting phases of derivation*. PhD. Diss. University of Maryland.

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

Embick, David & Rolf Noyer. 2001. Movement Operations after Syntax. LI 32(4): 555-595.

Enderton, Herbert B. 1972. *A mathematical introduction to logic*. New York: Academic Press.

Epstein, Samuel D. 1999. Un-Principled Syntax and the Derivation of Syntactic Relations. In: S. Epstein and N. Hornstein (eds.) *Working Minimalism*. Cambridge, MA: The MIT Press.

Fortuny, Jordy 2008. *The emergence of order in syntax*. Amsterdam: John Benjamins.

Gallego, Ángel J. 2010. Locality and anti-locality in a Merge-based system. Handout of a talk presented at the *Antilocality and snowballing movement* workshop, GIST, Ghent, June 2010.

Gelderen, Véronique van. 2003. Scrambling unscrambled. PhD. Diss. Leiden University.

Grimshaw, Jane. 2009. The extended projection. Rutgers University, Ms.

Halle, Morris. 1997. Impoverishment and fission. In Papers at the Interface, edited by Benjamin Bruening, Y. Kang, and Martha McGinnis, no. 30 in MITWPL, pp. 425 – 449. MIT, Cambridge, MA.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In: Hale, Kenneth & Samuel Jay Keyser (eds.) *The view from building 20*. Cambridge: MIT Press. pp. 111-176.

Harley, Heidi. 1994. *Hug a Tree: Deriving the Morphosyntactic Feature Hierarchy*. MITWPL 21, Papers on Phonology and Morphology.

Harley, Heidi 2009. Roots: Identity, Insertion, Idiosyncracies. Talk presented at the Root Bound workshop, USC, February 21.

Harley, Heidi & Rolf Noyer. 1998. Licensing in the non-lexicalist lexicon: nominalizations, Vocabulary Items and the Encyclopedia. *MITWPL 32: Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, edited by Heidi Harley, 119-137. Cambridge: MITWPL.

Harley, Heidi and Rolf Noyer (1999) State-of-the-Article: Distributed Morphology, *Glot International* 4.4, 3-9

Hoffman, J. 1996. Syntactic and paratactic word order effects. PhD. Diss. University of Maryland.

Jaspers, Dany. 1998. Categories and recursion. *Journal of Applied Linguistics* 12:81-112.

Kiparsky, Paul. 1973. Elsewhere in phonology. In: Kiparsky, Paul and Stephen R. Anderson (eds.) *A Festschrift for Morris Halle*. 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.

Koppen, Marjo van. 2005. *One probe - two goals: Aspects of agreement in Dutch dialects*. PhD Diss. Leiden University.

Langendoen, D. Terence. 2003. Merge. In: A. Carnie, H. Harley and M. Willie (eds.) *Formal approaches to function in grammar: in honor of Eloise Jelinek*, 307-318. Amsterdam: John Benjamins.

Levelt, Claartje et al. 1999. *A developmental grammar for syllable structure in the production of child language.* Brain and Language 68, 291-299.

Marantz 1995. Cat as a phrasal idiom: Consequences of late insertion in Distributed Morphology. MIT, Cambridge MA, Ms.

Marantz, Alec 1997. No escape from syntax. In A. Dimitriadis, L. Siegel, et al., eds., University of Pennsylvania Working Papers in Linguistics, Vol. 4.2, Proceedings of the 21st Annual Penn Linguistics Colloquium, 201-225.

Marantz, Alec. 2006. Morphology Lectures: Streaming Video. Lectures at the DEC of the ENS, Paris, October. http://www.diffusion.ens.fr/index.php?res=themes&idtheme=18

Moro, Andrea. 2000. Dynamic antisymmetry. Cambridge, MA: The MIT Press.

Munaro, Nicola & Cecilia Poletto 2003. *Sentential particles and clausal typing in the Veneto dialects*. University of Venice Working papers in linguistics 13.

Partee, Barbara, Alice Ter Meulen & Robert Wall. 1987. *Mathematical methods in linguistics*. Dordrecht: Kluwer Academic Publishers.

Pfau, Roland 2009. *Grammar as processor. A Distributed Morphology account of spontaneous speech errors*. Amsterdam: Benjamins.

Postma, Gertjan & Johan Rooryck. In progress. Redesigning C_{HL} as the Single Tree Workspace. Ms. in preparation Leiden University/Meertens Insitute.

Pylkkänen, Liina 2002. *Introducing Arguments*. PhD Diss., MIT.

Siddiqi, Daniel. 2006. Minimize Exponence: Economy effects on the morphosyntactic component of the grammar. PhD Diss., University of Arizona.

Uriagereka, Juan. 1999. Multiple Spell-Out. Ms. University of Maryland.

Zwart, Jan-Wouter. 2009a. Uncharted territory? Towards a non-cartographic account of Germanic syntax. In: Artemis Alexiadou e.a. (eds.) *Advances in Comparative Germanic Syntax*. Amsterdam: John Benjamins.

Zwart, Jan-Wouter. 2009b. Prospects for top-down derivation. *Catalan Journal of linguistics* 8:161-187.

Zwart, Jan-Wouter. 2010. Structure and order: asymmetric merge. To appear in Cedric Boeckx (ed.) *Oxford Handbook of Linguistic Minimalism*. Oxford: OUP.