A third structure building operation for Minimalist Grammars

Johannes Schneider
Universität Leipzig
johannes.schneider@uni-leipzig.de

Abstract

I propose a new structure-building operation for Minimalist Grammars (Stabler 1997) which allows the grammar formalism to grow trees with more than one root. I demonstrate that together with the assumption that this new long-distance dependency holds between nominal arguments and their selectors, one can generate Horn amalgams and parasitic gaps with a number of desired properties.

1 Introduction

I propose a new structure-building operation *3rd-merge*, formalized within the framework of Minimalist Grammars (Stabler 1997), that makes it possible to generate tree structures with more than one root where the two roots are structurally independent except at a single connecting phrase, the pivot. Within minimalist syntax, a number of proposals exists to extend the formalism beyond the operations *merge* and *move*, most notably sidewards movement (Nunes 1995), parallel merge (Citko 2005) and grafts (van Riemsdijk 2006, van Riemsdijk 2010). The effect of *3rd-merge* is to allow the selector to long-distance select its argument out of an otherwise structurally independent root (I discuss similarities and differences to the abovementioned extensions below). I propose that this new long-distance dependency underlies the phenomena of Horn amalgams and parasitic gaps.

Horn amalgams are constructions where two apparently independent clauses share a common element (Lakoff 1974):

(1) Joscha adores [I think it was <u>cats</u>].

cats appears to be both the argument of *adore* and *was*. The clause containing the latter verb is structurally independent from the matrix clause; *adore* does not select for the parenthesis-like clause but the noun *cats*. Neither of those clauses c-command the other. Evidence for this comes e.g. from the fact that binding (or any other syntactic operation) between elements from each clause is impossible (see Kluck 2011, ch.3 for an overview). The so-called <u>pivot</u> *cats* is therefore shared by two otherwise independent clauses and is the only element accessible to both clauses.

In parasitic gaps, an otherwise ungrammatical long-distance dependency (here: extraction out of an adjunct) becomes grammatical in certain configurations in the presence of a licit long-distance dependency (Engdahl 1983 i.a.):

(2) [<u>Which article</u>]$_1$ did you file t_1 [without reading pg_1]?

Both 'real' and parasitic gap refer to the same element. I argue that the matrix clause and the adjunct share the single element *which article* in the same manner as amalgams; the crucial difference is that the adjunct as additional root is reintroduced into the matrix root. When the pivot is moved, this creates the appearance of two gaps.

The structure of this article is as follows: I present the algebraic definition of Minimalist Grammars from Stabler and Keenan (2003) (Section 2). I then introduce the new operation and the rules describing its behaviour (Section 3), together with an application to the phenomena they are supposed to derive. Section 4 concludes with a comparison with other operations and a discussion of open issues.

2 Minimalist Grammars

Stabler and Keenan (2003) provide an algebraic definition of Minimalist Grammars (MGs). A Minimalist Grammar $G = \langle \Sigma, F, Types, Lex, \mathscr{F} \rangle$, with a non-empty alphabet Σ, the set of Features F consisting of *base* features (n,v,c,...), the respective selection features, as well as licensor and licensee features for movement, i.e. $F = base \cup \{=f | f \in$

Proceedings of the Society for Computation in Linguistics (SCiL) 2023, pages 6-14.
Amherst, Massachusetts, June 15-17, 2023

$base\} \cup \{+f|f \in base\} \cup \{-f|f \in base\}$[1], the $Types = \{::,:\}$, with $\cdot \in \{::,:\}$ as shorthand. They call Chains $C = \Sigma^* \times Types \times F^*$, and Expressions $E = C^+$, with the lexicon $Lex \subseteq C^+$ as a finite subset of $\Sigma^* \times \{::\} \times F^*$. There are two operations $\mathscr{F} = \{merge, move\}$, defined as partial functions as in Figure 1. I use $s, t, u, v, w \in \Sigma^*$; $\beta, \gamma \in F^*$; $\delta, \varepsilon, \zeta \in F^+$, chains $\alpha_1, ..., \alpha_k$ or $\iota_1, ..., \iota_\ell$ or $\mu_1, ..., \mu_m$ with $k, l, m \geq 0$.

$merge : (E \times E) \to E$ consists of three subcases, defining merge into complement position (*merge1*) and specifier position (*merge2*) and merge of a moving element (*merge3*). $move : E \to E$ is described by two functions for which the Shortest Move Constraint (SMC) holds: no member of $\alpha_1, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_k$ has $-f$ as first feature. There is either movement to a final (*move1*) or non-final (*move2*) landing site.

3 A new operation

The core intuition behind the new operation is to create a new type of long-distance dependency between nominal arguments and their selectors such as verbs that allows verbs to select an element from within an independent root. Crucially, the secondary root remains structurally independent except at the selected phrase. A possible visualization would be that of a form of 'long-distance' in-situ merge. I introduce a new type of positive/negative feature pair for the new operation *3rd-merge*: $\#f$ and $\nparallel f$[2]. By assumption, there is only a single feature of this type in a language ($\nparallel n$ or $\nparallel d$, depending on what the assumed highest projection in the nominal domain of that language is). This restriction is driven purely by empirical considerations, to restrict the phenomena of amalgams and parasitic gaps to nominals (for now).

In addition to the symbols in the standard MG definition, I use $\psi \in F^3$ where $F^3 = \{\nparallel f | f \in base\}$; ω is of the form $[t : \psi\gamma, \mu_1, ..., \mu_m]$ and $\varsigma \in \{n, c\}$. For the present purposes, I restrict the structure of potential lexical items as follows: $\{=f, \#f\}^*.f.(\nparallel f.) {-f}^*$, i.e. there is at most one $\nparallel f$ directly after the category symbol of any given lexical item; $\#f$ behaves like selector features. Figure 2 provides an overview of all rules.

Let us assume that the highest projection in the nominal domain is *n*, and that all nouns have both n and $\nparallel n$ in their feature string ($cat :: n.\nparallel n$). A consequence is that additional roots can only grow on top of nominals. The assumption is that nominals are always merged via an application of *3rd-merge*. Phrases with feature string $f.\nparallel f$ are *3rd-merged* into complement or specifier position (*3merge-1/2*) or as moving item (*3merge-3*). Nominals are therefore treated as a trivial case of an independent root, namely one where no additional structure has grown on top of it. The category of this trivial root n is treated as syntactically inert after the application of *3rd-merge*.

The system also allows a non-trivial root to grow on top of a nominal before it is *3rd-merged*. I call such a root the *secondary root*[3] (e.g. the bracketed '*I think it was*' in (1)) since *3rd-merge* creates an asymmetry between the roots, as will be discussed below.

The rule *merge4* governs the special case where growth of a non-trivial secondary root is initiated on top of a nominal. A head selects for a category feature of an expression that is followed by $\nparallel f$. The merge features of the argument are erased but it becomes part of the chains of the selector, akin to merging moving expressions, and becomes inaccessible for the rest of the derivation within the secondary root until it is selected via an application of *3rd-merge* out of a different root. The inaccessible pivot is indicated by square brackets. I denote bracketed elements of the form $[t : \psi\gamma, \mu_1, ..., \mu_m]$ as ω. Note that so far this operation is only defined to apply in complement position. Example derivations for an amalgam and a parasitic gap can be found in Figure 3. '*I think it was*' and '*without reading*' are treated as such secondary roots, and the first steps in both derivations (selection by *was* or *reading*) exemplifies an application of *merge4*.

The introduction of ω by *merge4* now requires an update to the former *merge* and *move* rules so that an expression can contain 0 or 1 ω ('ω' abbreviates '0 or 1 ω' for readability). *merge1* and *merge2* remain unaffected save a potential presence of an inert ω. The argument in *merge1* and the function in *merge2* can contain ω. Note that, similar to the complement-only restriction for *merge4*, I disallow merging an expression into specifier position that contains ω. This would allow a potentially unbounded number of ω in an expression, with potentially non-trivial nesting, something I want to

[1]Note that there is some redundancy in this definition since there are no base movement features as categories.

[2]'plus-equals' and 'minus-equals', for lack of better terms.

[3]I use 'root' here as pars pro toto for the whole single-rooted subtree in a multi-rooted tree.

$$\frac{s :: =f\delta \qquad t \cdot f, \alpha_1, ..., \alpha_k}{st : \delta, \alpha_1, ..., \alpha_k}\ \text{merge1} \qquad \frac{s \cdot =f\delta, \alpha_1, ..., \alpha_k \qquad t \cdot f\varepsilon, \iota_1, ..., \iota_\ell}{s : \delta, \alpha_1, ..., \alpha_k, t : \varepsilon, \iota_1, ..., \iota_\ell}\ \text{merge3}$$

$$\frac{s : =f\delta, \alpha_1, ..., \alpha_k \qquad t \cdot f, \iota_1, ..., \iota_\ell}{ts : \delta, \alpha_1, ..., \alpha_k, \iota_1, ..., \iota_\ell}\ \text{merge2} \qquad \frac{s : +f\delta, \alpha_1, ..., \alpha_{i-1}, t : -f, \alpha_{i+1}, ..., \alpha_k}{ts : \delta, \alpha_1, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_k}\ \text{move1}$$

$$\frac{s : +f\delta, \alpha_1, ..., \alpha_{i-1}, t : -f\varepsilon, \alpha_{i+1}, ..., \alpha_k}{s : \delta, \alpha_1, ..., \alpha_{i-1}, t : \varepsilon, \alpha_{i+1}, ..., \alpha_k}\ \text{move2}$$

Figure 1: Standard MG rules

$$\frac{s :: =f\delta \qquad t \cdot f, \omega, \alpha_1, ..., \alpha_k}{st : \delta, \omega, \alpha_1, ..., \alpha_k}\ \text{merge1}$$

$$\frac{s : =f\delta, \omega, \alpha_1, ..., \alpha_k \qquad t \cdot f, \iota_1, ..., \iota_\ell}{ts : \delta, \omega, \alpha_1, ..., \alpha_k, \iota_1, ..., \iota_\ell}\ \text{merge2}$$

$$\frac{s : =f\delta, \omega, \alpha_1, ..., \alpha_k \qquad t \cdot f\varepsilon, \mu_1, ..., \mu_m}{s : \delta, \omega, \alpha_1, ..., \alpha_k, t : \varepsilon, \mu_1, ..., \mu_m}\ \text{merge3 spec}$$

$$\frac{s :: =f\delta \qquad t \cdot f\psi\gamma, \alpha_1, ..., \alpha_k}{s : \delta, [t : \psi\gamma, \alpha_1, ..., \alpha_k]}\ \text{merge4}$$

$$\frac{s : +f\delta, \omega, \alpha_1, ..., \alpha_{i-1}, u : -f, \alpha_{i+1}, ..., \alpha_k}{us : \delta, \omega, \alpha_1, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_k}\ \text{move1}$$

$$\frac{s : +f\delta, \omega, \alpha_1, ..., \alpha_{i-1}, u : -f\varepsilon, \alpha_{i+1}, ..., \alpha_k}{s : \delta, \omega, \alpha_1, ..., \alpha_{i-1}, u : \varepsilon, \alpha_{i+1}, ..., \alpha_k}\ \text{move2}$$

$$\frac{s :: \#f\delta \qquad t \cdot f.\not\#f, \alpha_1, ..., \alpha_k}{st : \delta, \alpha_1, ..., \alpha_k}\ \text{3merge-1}$$

$$\frac{s : \#f\delta, \omega, \alpha_1, ..., \alpha_k \qquad t \cdot f.\not\#f, \iota_1, ..., \iota_\ell}{ts : \delta, \omega, \alpha_1, ..., \alpha_k, \iota_1, ..., \iota_\ell}\ \text{3merge-2}$$

$$\frac{s \cdot \#f\delta, \omega, \alpha_1, ..., \alpha_k \qquad t \cdot f.\not\#f\varepsilon, \iota_1, ..., \iota_\ell}{s : \delta, \omega, \alpha_1, ..., \alpha_k, t : \varepsilon, \iota_1, ..., \iota_\ell}\ \text{3merge-3}$$

$$\frac{s :: \#f\delta \qquad t : c, [u : \not\#f, \alpha_1, ..., \alpha_k]}{stu : \delta, \alpha_1, ..., \alpha_k}\ \text{3merge-1'}$$

$$\frac{s : \#f\delta, \omega, \alpha_1, ..., \alpha_k \qquad t : c, [u : \not\#f, \iota_1, ..., \iota_\ell]}{tus : \delta, \omega, \alpha_1, ..., \alpha_k, \iota_1, ..., \iota_\ell}\ \text{3merge-2'}$$

$$\frac{s \cdot \#f\delta, \omega, \alpha_1, ..., \alpha_k \qquad t : \varsigma\gamma, [u : \not\#f\varepsilon, \iota_1, ..., \iota_\ell]}{s : \delta, \omega, [t : \varsigma\gamma, u : \varepsilon], \iota_1, ..., \iota_\ell, \alpha_1, ..., \alpha_k}\ \text{3merge-4}$$

$$\frac{s : =\varsigma\delta, \omega, [t : \varsigma, u : \varepsilon], \alpha_1, ..., \alpha_k}{ts : \delta, \omega, \alpha_1, ..., \alpha_k, u : \varepsilon}\ \text{chain-merge1}$$

$$\frac{s : =\varsigma\delta, \omega, [t : \varsigma\varepsilon, u : \zeta], \alpha_1, ..., \alpha_k}{s : \delta, \omega, \alpha_1, ..., \alpha_k, t : \varepsilon, u : \zeta}\ \text{chain-merge2}$$

$$\frac{s : +f\delta, \omega, \alpha_1, ..., \alpha_{i-1}, [t : \varsigma\gamma, u : -f\varepsilon], \alpha_{i+1}, ..., \alpha_k}{s : \delta, \omega, \alpha_1, ..., \alpha_{i-1}, [t : \varsigma\gamma, u : \varepsilon], \alpha_{i+1}, ..., \alpha_k}\ \text{move3}$$

Figure 2: MGs with *3rd-merge*

exclude. There is thus a **ban on $ǂf$ in specifiers.**

The original *merge3* is joined by the additional *merge3 spec* that allows for the possibility of a selector containing ω when merging into specifier position. A selector cannot contain ω when merging its complement, but I also exclude the option that its argument contains ω in this case. Such a rule would be 'unphysical' in that it would lead to 'root-distributed remnant movement'. An example would be VP-movement inside an amalgam minus its inert object as ω which appears in its in-situ position in the matrix root. To the best of my knowledge there is no such phenomenon. Previous *move* rules, however, are updated trivially to allow for the presence of inaccessible ω. In sum, *merge4* initiates growth of a secondary root, but previous *merge* and *move* rules continue to build structure in the familiar way, largely unaffected by the presence/absence of a pivot.

3.1 Amalgams

We now turn to the cases where *3rd-merge* selects an argument from within a non-trivial secondary root, as in Horn amalgams such as (1) where '*I think it was cats*' grows on top of the cordoned off NP *cats* with a $ǂn$ feature. This clause, and the clause headed by *adore* are structurally independent. As mentioned, no element from either clause can bind an element from the other, i.e. there is no c-command relationship between these clauses[4]. Only the pivot is accessible for both clauses. In *3merge-1'/2'*, the selector selects (and therefore c-commands) only the element carrying $ǂf$ from the bracketed chain, but does not establish any syntactic relationship with the rest of the expression. This property of the rule effectively introduces multiply-rooted trees since there is an undominated complete root tree with a single position inside where it can 'dock' with another root.

The rule enforces that the expression from which the argument of the selector originates is a complete CP which has no features unchecked besides the pivot in the bracketed part. The *c* on the head of the secondary root is *not* selected and remains unchecked but vanishes in the resulting expression. This implements the idea that a syntactic object is complete if the only unchecked features in *all* its roots are of start category *c*. Another effect that this rule enforces is a derivational 'timing' in that secondary roots can only be connected with a primary root (by definition the root whose head carries $\#f$) after they are built completely, but not in an intermediate stage. This would yield an expression with multiple heads that need to check their features, with often non-trivial bracketing. The rules above avoid this complication.

In the result of the rule, amalgam plus pivot are linearized as a single unit with respect to the verb, yielding the correct *adores I think it was cats*. This step is also illustrated in Figure 3. Also note that I do allow completed amalgams in specifier position (*3merge-2'*), it is only *unchecked* $ǂf$-features that are disallowed, for the reasons discussed above.

As a last note, the rules allow for subextraction from pivots into the matrix root, as indicated by the presence of chains. This is empirically justified:

(3) Of which person does Daniel have [I think it was a painting *t*]?

There is also empirical justification for isolating from the secondary root not only the $ǂf$-carrying element but also its moving subparts, i.e. disallow subextraction into the secondary root. The outcome of such an extraction is ungrammatical. This is another way in which the asymmetry between selector and selectee in a *3rd-merge* dependency manifests itself.

(4) a. *Ville has [of his daughter, I think it was a painting *t*]
b. *Ørjan has [of which daughter, do I think it was a painting *t*].

I turn now to cases where there is additional structure on top of the pivot, but the pivot NP itself has a movement feature. I want to exclude movement of the pivot of amalgams. Movement of the pivot on its own is ungrammatical and would, metaphorically speaking, lead to the amalgam being a disconnected piece of structure; pied-piping of the whole amalgam also appears to be quite degraded:

(5) a. *Chicago, Peter went to [I think it was *t*].
b. ?*[I think it was Chicago], Peter went to *t*.

Instead, I want to reserve such cases for a different phenomenon. I propose the following empirical split. With additional roots, there are two possibilities: either that root remains free, which corresponds to amalgams, or that root is reintroduced into the matrix root again. I propose that this option

[4] For this reason, the idea, as suggested by a reviewer, to let *adore* select the cleft-structure and then select for *cats* via a step of covert movement leads to wrong predictions.

only occurs when the pivot connecting both parts of the resulting cyclic graph carries a movement feature. This movement of the connecting element leads - at least from a derived tree perspective - to the breakup of the cycle. In a 1D-string, such a movement gives rise to what appears to be two distinct gaps. The phenomenon that this corresponds to is parasitic gaps (such a multidominance account of parasitic gaps has been proposed in Kasai 2010).

3.2 Parasitic gaps

3merge-4 is the rule that governs the behaviour of moving pivots[5]. Its effect is that of selection for the moving pivot without erasing the category feature of the root that hosts it - in contrast to amalgams (this can be seen in Figure 3 where *file* selects *which article* out of the adjunct). That root, ς, can either be *c* or *n/d* since parasitic gaps can occur not only in clausal elements like relative clauses or adjuncts but also in NPs as in subject parasitic gaps. A small difference to amalgams is that the host phrase of the pivot can carry a movement feature. I allow this possibility for potential movement of adjuncts or subject movement. Other than that, there are no unchecked features besides ω.

There are a number of side issues that this rule addresses as well. Upon *3rd-merging* the pivot, sub-movers $\iota_1, ..., \iota_\ell$ are 'released' so that they become accessible parts of the chain in the outcome. This appears to be empirically justified since e.g. complements of nouns that are pivots in parasitic gap constructions can be scrambled to a position lower than the pivot (though, again, only into the matrix root, not the secondary one):

(6) ?[Welche Bücher t_1]$_2$ hat [über Potsdam]$_1$
which books has about potsdam
jeder gekauft t_2 ohne je zu lesen pg_2?
everyone bought without ever to read
Which books about Potsdam did everyone buy without ever reading?

I also disallow movement of the pivot itself within the secondary root. Such a step appears unnecessary for Horn amalgams where the pivot occurs in base position within the amalgam. This also fits well with the observation that parasitic gaps in subject position are usually ungrammatical (see e.g. Mayo 1997). Under standard assumptions, subjects move to receive case. The impossibility of moving in the secondary root would exclude this for independent reasons. This ties in with another property of the pivot that I have assumed throughout the definitions in which they appeared: they always occur in complement position in the secondary root. For the *it*-cleft(-like) constructions in amalgams, this appears to be correct, as well as for parasitic gap hosts. As mentioned, subject gaps are excluded. Indirect object gaps appear to be degraded as well:

(7) ?*Which person did you send out after giving pg_1 an article?

Until a convincing need for non-complement pivots arises, I restrict *3merge-4* to complement position.

I have allowed for the presence of ω in the selector in *3merge-4* which in principle allows for amalgams in amalgams or parasitic gaps in parasitic gap hosts. Whether this is empirically justified is beyond the scope of this paper.

Let us return to the core issues. The expression that is introduced into the chain of the selector ($[t : \varsigma\gamma, u : \varepsilon]$) as a result of *3merge-4* differs from ω in that it does not contain a $\nparallel f$-feature. There are three rules that govern the behaviour of such a bracketed expression. Either the pivot can move to a non-final landing site (*move3*), 'pied-piping' the whole expression with it. This would be the case if A-movement precedes $\bar{A}$-movement of the pivot.

The other two rules (*chain-merge1/2*) govern the reintroduction of the secondary root into the main root. This amounts to 'chain-internal' merge, akin to move rules, with the difference being that the expression carries an unchecked *category* feature, not a movement feature. *chain-merge1/2* describe the point in the derivation where the parasitic gap host (e.g. an adjunct like *[without reading:c, which book:-wh]*) is merged into its position in the matrix root (an application for *chain-merge1* occurs in Figure 3 where an empty vP-adjunction head ϵ selects for the adjunct). As a result, the parasitic gap host is either linearized in its final position (*chain-merge1*) or becomes a moving chain (*chain-merge2*); in both cases, the moving element that corresponds to both real and parasitic gap (*u:ε*) is released and becomes part of the chains. From there it moves to a position higher than the reintroduction site of its host, deriving the anti-c-command property of parasitic gaps. The last steps of the derivation of (2) can also be found in Figure 3[6].

[5]The rule as it stands is an abbreviation for the specifier and complement merge cases - in the latter case, $\alpha_1, ..., \alpha_k$ and ω is missing in the selector.

[6]I abstract away both from how adjunction works (treating it as normal merge) and the rightward dislocation of the

3.3 Further issues and applications

So far I have assumed that only nominals carry a $ǂf$-feature. It is considered a core property of parasitic gaps that only NPs can correspond to them (see e.g. Culicover 2001). NPs are also prototypical pivots in amalgams; predicative adjectives e.g. are degraded as pivots in Horn amalgams (Kluck 2011, 74):

(8) ?*Bea is [I think it's <u>blond</u>].

Note though that Engdahl 1983 cites AP parasitic gaps as acceptable in Swedish and that amalgams in NPs on top of attributive adjectives are acceptable in some contexts ('*an [I think you can call it simple <u>solution</u>]*', see Kluck 2011). Further research is necessary to determine whether the restriction of $ǂf$ to nominals is correct or needs to be relaxed at least for adjectives.

In the more restrictive system, there is only a single category that can carry the additional negative feature, i.e. either $n.ǂn$ or $d.ǂd$, depending on one's stance in the DP/NP-debate and/or whether one describes a DP- or NP-language (see Bošković 2008 for such a split). With the new operation, however, it is possible to propose a new solution to the structure of the DP: one assumes that NPs universally have the feature sequence $n.ǂn$ and that verbs select for $ǂn$, even in DP-languages, explaining why verbs can 'long-distance' select for types of NPs even in those languages (an argument against DP-structure by Bruening et al. 2018). DP-languages would differ from the system presented so far in that every DP is a '*mini-amalgam*': they require *d* to select NPs as in a *merge4* application, and it is only the resulting DP that can be the argument in a *3merge-1* rule.

$$\frac{s :: {=}n.d\gamma \qquad t \cdot n.ǂn, \alpha_1, ..., \alpha_k}{[st : d.ǂn.\gamma, \alpha_1, ..., \alpha_k]} \text{ DP-merge}$$

$$\frac{s :: \#n\delta \qquad [t \cdot d.ǂn, \alpha_1, ..., \alpha_k]}{st : \delta, \alpha_1, ..., \alpha_k} \text{ 3merge-1}^{DP}$$

The first *d* selecting $n.ǂn$ thus has a special status, and it is only further merge with that DP that leads to amalgams proper or parasitic gap hosts.

The only purpose of *3rd-merge*, then, is to connect the two major spines, the nominal and the verbal/clausal one. In such a system, the fact that the additional root can grow further and either remain independent (amalgams) or get reconnected (parasitic gaps) is a simple consequence of the way clausal and nominal spine are merged. The three apparently distinct phenomena share a common core, and the fact that parasitic gaps and amalgams are restricted to nominals falls out as a consequence of the assumption that *3rd-merge* connects verbs and nominals and does not need to be stipulated separately.

Showing the full rule set for this system is beyond the scope of this paper, however. There are a number of empirical and theoretical issues that need to be considered. Possibilities like NP extraction out of DP as e.g. in German complicate the rules. One also needs to ensure that it is the first D selecting an NP that is turned into a bracketed ω, not a higher one. This is easier if one assumes that all additional material in NPs like adjectives and numerals are adjoined by category preserving operations and D always selects something of type $n.ǂn$. Not all approaches assume this and would need to be dealt with differently. I therefore leave a full exploration of a system that unifies DP/NPs, amalgams and parasitic gaps for future research.

As a last point, there is also the issue that when growing amalgams or parasitic gap hosts (in a *merge4* step), the verb needs to select via $=n$ or $=d$. Thus one would need to allow optionality in the way verbs select arguments ($ǂn/{=}d$) which appears to unnecessarily bloat their lexicon entry. However, this is independently necessary if one assumes that (weak) pronouns are just a single head d [7]. A stronger but related argument for the variable nature of selection comes from a number of verbs that disallow weak pronouns, Postal's 1994 so-called anti-pronominal contexts (9-a). Strikingly, it is exactly those verbs that cannot occur in parasitic gap hosts (9-c) even though they do allow wh-extraction (9-b). Both apparently unrelated facts are derived together by the assumption that the lexicon entries of this class of verbs lack the $=d/{=}n$ option:

(9) a. *She likes the colour black, so she painted the door it.
b. What colour$_1$ did she paint the door t_1?
c. *What colour$_1$ did she grow to hate t_1 after painting her door pg_1?

This is also yet another example of an asymmetry

adjunct in this example.

[7] Verbs still cannot select via $=d$ for 'full' DPs with lexical content in matrix roots since they would then contain unchecked $ǂn$, preventing the derivation to converge.

between the roots since the mode of selection for the pivot appears to differ. An in-depth investigation of the empirical facts concerning this asymmetry is beyond the scope of this paper, however.

4 Discussion

To summarize, I introduced a new operation, *3rd-merge*, to Minimalist Grammars. By postulate, only nominals can carry $\sharp f$ so the new long-distance dependency holds between nominal arguments and their selectors. The main effect of the new operation is to allow selection of an item from within an additional root without establishing a syntactic relation to any other part of that root. Additional roots can either remain independent, which corresponds to the phenomenon of Horn amalgams, or they can be reintroduced into the matrix root, which results in parasitic gap constructions after movement of the pivot in the resulting cyclic structure.

I want to discuss a number of commonalities and differences between *3rd-merge* and other extensions of minimalist grammars. What *3rd-merge* and sidewards movement (Nunes 1995, especially in the formalization by Stabler 2006) have in common is the general idea of further relaxing resource sensitivity. In the sidewards movement system in Stabler (2006), however, a category feature e.g. can be re-used a potentially unbounded number of times. The system set up here does not give up resource sensitivity completely but only allows one further type of re-use of an expression, besides being merged and moved, thereby stipulating a third dependency type. The third type of re-use leads to the growth of an additional root which is distinct from the possibilities of sidewards movement. Just as movement is not 'just' a reuse of category features but a dependency (related to but nonetheless) distinct from merge with its own properties and restrictions, it is important in my opinion to equally separate this third reuse of expressions. This way one can investigate the properties and restrictions of this new dependency in their own right.

If resource-sensitivity needs to be relaxed further, e.g. for multiple parasitic gap constructions, one has more control over which features exactly are to be changed in that way. Whether it is *merge*, *move* or *3rd-merge* features that can be reused might have different empirical consequences.

Torr and Stabler (2016), building on Kobele (2008), extend MGs to deal with ATB-movement (among other things). The idea behind these approaches is the unification of the identical but distinct movers of both conjuncts. Those approaches are then extended to cover other one-to-many dependencies such as control and parasitic gaps. Parasitic gaps, however, are markedly different from ATB-constructions as demonstrated in Postal (1993). They are not confined to coordinations but are subject to a number of restrictions irrelevant for ATB, such as a restriction to $\bar{A}$-movement, a categorial restriction to NPs, the anti-pronominal condition shown in (9) and many others. Parasitic gaps are optional and their position can be filled, contrary to ATB-patterns where all gaps are obligatory and mutually depend on each other, i.e. a mutual symbiosis compared to an asymmetric parasitism. There is also the asymmetry in subextraction (see (6)) that is non-trivial in a system that treats the origin of unified movers on equal footing. For these reasons I treat the asymmetries of parasitic gaps as a different phenomenon, not a subtype of ATB-movement.

The properties of amalgams are another central reason to adopt a system as presented here. There is convincing evidence that amalgams contain an undominated, independent secondary root (Kluck 2011, ch.3), a structure the above approaches cannot currently generate. In the present approach, a head can select for an element from within a different root without, however, connecting with the rest of the root. Amalgams also exhibit restrictions and asymmetries that are shared by parasitic gaps, such as a putative categorial restriction to NPs and (sub)extraction asymmetries (see (4)). Since these phenomena pattern together and they can both be derived by a system that allows multiple roots, it is useful to derive them with the same mechanism while treating the more symmetrical ATB-phenomenon as distinct.

What distinguishes *3rd-merge* from parallel merge (Citko 2005, Citko 2011), apart from a formal implementation, is that it is not 'parallel' or symmetric. In parallel merge, a head A and a head C that merge with phrase B both stand on equal footing. *3rd-merge* introduces an asymmetry between selector and selectee. This property is shared by grafts (van Riemsdijk 2006, van Riemsdijk 2010), the operation that is its closest match. van Riemsdijk uses this operation mainly to derive properties of free relatives and transparent free relatives ('*She ate what she called egg fried rice.*') but also Horn amalgams. van Riemsdijk notes empirical asymmetries but remains agnostic as to how they come about (see e.g. van Riemsdijk 2006, fn.8, 'all trees are equal'). For *3rd-merge*, the asymmetry is built into the definition of the operation: the selector is always part of the matrix root while additional structure on top of the NP is always a secondary root that is integrated into the main structure. Further asymmetries are part of the definition, such as the impossibility of extraction and subextraction of the pivot into the secondary root. Graft can apply at any stage but must do so long before the whole clause is built for phase considerations (van Riemsdijk 2006, ch. 4.3). I proposed the opposite for *3rd-merge* since merge of an intermediate stage of a secondary root would lead to a proliferation of unchecked features that are difficult to track in the algebraic definition presented here.

Before closing this article, I want to mention two further issues that need to be addressed. One is the linearization of amalgams. Not only were the pivots considered so far the most deeply embedded complement, they were also the most rightward element in the string. This would be different in SOV amalgams or with extraposed adverbials:

(10) ?Peter hat [ich glaub es war die Katze gewesen] gestreichelt.
Peter has I think it was the cat been petted.
Peter petted I think it was the cat.

This cannot be derived in the system set up so far. One reason for this is that the algebraic definition used here does two things at once: regulate the feature calculus *and* linearize the string. A more fine-grained approach should be able to treat those matters separately.

The last question concerns the expressive power of the grammar presented so far. Though I assume it to be the case, showing the equivalence to MCFGs would be reassuring. Apart from empirical considerations, it might be relevant for that purpose to determine whether to allow ω in *3merge-2'/4*, i.e. whether it is safe to allow unbounded nested amalgams/parasitic gaps. The same goes for the question whether there should be an SMC equivalent for the structure $[t : \varsigma\gamma, u : \varepsilon]$. Occurrence of more than one such element in an expression might lead to unwanted indeterminacies. As a last point, it would be of interest to know whether MGs with *3rd-merge* but without (remnant) movement allow generation of non-context free patterns.

References

Željko Bošković. 2008. What will you have, DP or NP? In *Proceedings of NELS 37*, volume 1, pages 101–114.

Benjamin Bruening, Xuyen Dinh, and Lan Kim. 2018. Selection, idioms, and the structure of nominal phrases with and without classifiers. *Glossa: a journal of general linguistics*, 3(1).

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

Barbara Citko. 2011. Multidominance. In Cedric Boeckx, editor, *The Oxford Handbook of Linguistic Minimalism*, chapter 6, pages 119–142. Oxford University Press.

Peter W. Culicover. 2001. Parasitic Gaps: A History. In *Parasitic Gaps*, pages 3–68. Oxford University Press.

Elisabet Engdahl. 1983. Parasitic gaps. *Linguistics and Philosophy*, 6(1):5–34.

Hironobu Kasai. 2010. Parasitic Gaps Under Multiple Dominance. *English Linguistics*, 27(2):235–269.

Marina Elisabeth Kluck. 2011. *Sentence amalgamation*. Ph.D. thesis, University of Groningen.

Gregory M. Kobele. 2008. Across-the-board extraction in Minimalist Grammars. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+9)*, pages 113–120, Tübingen, Germany. Association for Computational Linguistics.

George Lakoff. 1974. Syntactic Amalgams. In *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, volume 10, pages 321–344. Chicago Linguistic Society.

Pilar García Mayo. 1997. Non-Occurrence of Subject and Adjunct Parasitic Gaps. *Atlantis*, 19(2):125–133.

Jairo Nunes. 1995. *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program*. Ph.D. thesis, University of Maryland.

Paul M. Postal. 1993. Parasitic gaps and the across-the-board phenomenon. *Linguistic Inquiry*, 24(4):735–754.

Paul M. Postal. 1994. Parasitic and pseudoparasitic gaps. *Linguistic Inquiry*, 25(1):63–117.

Edward Stabler. 1997. Derivational minimalism. In *Logical Aspects of Computational Linguistics*, pages 68–95. Springer Berlin Heidelberg.

Edward P. Stabler. 2006. Sidewards without copying. In *Proceedings of the 11th conference on Formal Grammar*, pages 157–170.

$$
\dfrac{\dfrac{\text{was} :: {=}n.{=}n.v \qquad \text{cats} : n.{\neq}n}{\text{was} : {=}n.v, [\text{cats} : {\neq}n]}\,\text{merge4} \qquad \mathit{it} :: n}{\text{it was} : v, [\text{cats} : {\neq}n]}\,\text{merge2}
$$

$$
\vdots
$$

$$
\dfrac{\text{adore} : \#n.\#n.v \qquad \text{I think it was} : c, [\text{cats} : {\neq}n]}{\text{adore I think it was cats} : \#n.v}\,\text{3merge-1'}
$$

$$
\vdots
$$

$$
\text{Joscha adores I think it was cats} : c
$$

(a) partial derivation of an amalgam as in (1)

$$
\dfrac{\text{reading} :: {=}n.v \qquad \text{which article} : n.{\neq}n.{-}wh}{\text{reading} : v, [\text{which article} : {\neq}n.{-}wh]}\,\text{merge4}
$$

$$
\vdots
$$

$$
\dfrac{\text{file} :: \#n.\#n.v \qquad \text{without reading} : c, [\text{which article} : {\neq}n.{-}wh]}{\text{file} : \#n.v, [\text{without reading} : c, \text{which article} : -wh]}\,\text{3merge-4}
$$

$$
\vdots
$$

$$
\dfrac{\epsilon\ \text{the manager file} := c.v, [\text{without reading} : c, \text{which article} : -wh]}{\text{the manager file without reading} : v, \text{which article} : -wh}\,\text{chain-merge1}
$$

$$
\vdots
$$

$$
\dfrac{\text{did the manager file without reading} : +wh.c, \text{which article} : -wh}{\text{Which article did the manager file without reading} : c}\,\text{move1}
$$

(b) partial derivation of a parasitic gap as in (2)

Figure 3: Example derivations

Edward P. Stabler and Edward L. Keenan. 2003. Structural similarity within and among languages. *Theoretical Computer Science*, 293(2):345–363.

John Torr and Edward P. Stabler. 2016. Coordination in Minimalist Grammars: Excorporation and Across the Board (Head) Movement. In *Proceedings of the 12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12)*, pages 1–17.

Henk van Riemsdijk. 2006. Grafts follow from merge. In *Phases of Interpretation*, pages 17–44. Mouton de Gruyter.

Henk van Riemsdijk. 2010. Grappling with graft. In *Structure Preserved*, Linguistik Aktuell, pages 289–298. John Benjamins Publishing Company.