

Deriving parasitic gaps by fission und fusion

Anke Assmann

2011

This paper presents a new approach to parasitic gaps which is able to implement the observation that two gaps have only one antecedent by transferring the concepts of fission and fusion to the syntactic component. I argue that the antecedent of a parasitic gap construction first fissions into two independent parts—thereby allowing it to occupy two different base positions—and is then fused again into one item which occupies the surface position. I further show that this analysis makes the correct predictions concerning the properties of parasitic gaps, setting it apart from previous approaches, which all had to struggle with capturing the behavior of parasitic gaps. In particular, I will present solutions to the following questions: (i) why are parasitic gaps only selectively sensitive to islands? (ii) Why do the real and the parasitic gap exhibit different properties concerning reconstruction and extractability from weak islands? (iii) Why can parasitic gaps only be licensed by \bar{A} -movement but not by A-movement?

1 Introduction

The answer to the question of how *one* dislocated syntactic category can be traced back to *two* different base positions in two different clauses has posed a challenge to linguistic theory for almost fifty years now. Discovered by Ross (1967), this phenomenon was thoroughly described by Taraldsen (1981) and Engdahl (1983) who introduced the term *parasitic gap* referring to the idea that there is actually only one real base position whereas the second putative base position is a gap that parasitically attaches to the antecedent-gap dependency.

In spite of many attempts to find an adequate analysis for this phenomenon—whereby scholars have put special emphasis on developing approaches which classify parasitic gap constructions as instances of other syntactic constructions—none of the solutions could succeed in fully capturing the idiosyncratic properties of parasitic gap constructions.¹

This paper presents a new account of parasitic gaps, analyzing the puzzling fact that one item seems to be able to occur simultaneously in two different positions by allowing it to fission into two parts. Under this approach, no use of concepts like interarboreal movement (Nunes 1995, 2001, 2004) or empty operator movement

¹For an overview of properties and analyses of parasitic gaps, see Culicover (2001).

(e.g. Chomsky 1986a; Nissenbaum 2000) is necessary. Theories which make use of these syntactic means are shown to make wrong empirical predictions.

In what follows, parasitic gaps are gaps which are covered by the definition in (1) (cf. Taraldsen 1981; Engdahl 1983).

- (1) *Parasitic gaps*
A parasitic gap (pg) is a gap which is licensed by the antecedent of another real gap (t) created by movement.

Constructions in several languages meet this definition. Consider the examples in (2) and (3) from English and German, respectively.

- (2) Which article did you file *t* [without reading *pg*]?
 (3) Welchen Artikel hast du [ohne *pg* zu lesen] *t* abgeheftet?
 which article have you without to read filed
 ‘Which article did you file without reading?’

In both these sentences, the *wh*-phrase *which article* has moved from its base position in the matrix clause—which is the complement position of the verb *file*—to the left periphery of the sentence, thereby licensing a second gap in the complement position of the verb *read* inside the adverbial clause.

Despite the fact that the behaviors of parasitic gaps in English and German are not fully identical, parasitic gap constructions in both languages show certain conspicuous properties.

The aim of this paper is to provide answers to major questions concerning parasitic gaps. First and foremost, we have to clarify what parasitic gaps are (4-a), i.e., how a 2-1 dependency can be established, and which syntactic operation they originate in if not movement (4-b).

- (4) a. What are parasitic gaps?
 b. Which syntactic operation do parasitic gaps originate in?

Then, we have to find explanations for the puzzling properties of parasitic gaps based on the answers to the two general questions. In particular, I will discuss and solve the following four conundrums parasitic gap constructions pose:

- (5) a. Why are parasitic gaps only selectively sensitive to islands?
 b. Why can complex antecedents only be reconstructed into the real gap but not the parasitic gap?
 c. Why are certain properties of the antecedent only important for its relation to the real gap but not to the parasitic gap?
 d. Why can parasitic gaps only be licensed by \bar{A} -movement but not by A-movement?

In order to answer these questions, I will start in section 2 by presenting a new idea of why parasitic gap constructions exist and how they come into existence. This new approach will then provide natural explanations for the puzzling properties summarized in the questions in (5) (section 3). Afterwards, in section 4, I will discuss empirical and theoretical questions concerning this new approach in comparison to

previous analyses of parasitic gaps. Section 5 concludes.

2 The fission approach to parasitic gaps

In order to answer the general questions about parasitic gaps I suggest a new approach, its main idea being that one lexical item can fission into two parts, making use of the concept of the morphological operation of fission (Halle and Marantz 1993). Fission will be construed as an operation that precedes the syntactic derivation. As a result of presyntactic fission, two unstable objects are created that may be targeted separately by syntactic operations for some time but will eventually have to be recombined, or fused, just like two morphemes may undergo morphological fusion (see Halle and Marantz 1993). Thus, fission and fusion explain how it is possible that in parasitic gap constructions there is only one overtly realized antecedent but two gaps, using concepts which are already established to derive completely different linguistic phenomena in morphology.

In this section, I will show that such syntactic fission integrated into a standard minimalist framework has major impact on derivations which in turn naturally explain the idiosyncratic properties of parasitic gap constructions.

In order to do so, I will begin by thoroughly clarifying the assumptions the analysis is based on before I introduce the syntactic fission operation and go through a detailed derivation of a parasitic gap sentence.

2.1 Background assumptions

Syntactic structure:

The syntactic structure I assume is shown in (6). Following standard minimalist assumptions, there are three functional heads above VP: v, T and C. Nominal categories are headed by a functional head D (Abney 1987).

$$(6) \quad [_{CP} C [_{TP} DP_{Subj} T [_{vP} DP_{Subj} v [_{VP} V DP_{Obj}]]]]$$

Assumptions about derivation (Chomsky 1995, 2000, 2001, 2008):

I assume that syntax is strictly derivational and obeys the Strict Cycle Condition (Chomsky 1973), and that derivations allow no look-ahead.²

The structure-building operation is Merge triggered by features [$\bullet F \bullet$] and the probe operation is Agree triggered by features [$*F*$] (Sternefeld 2006; Heck and Müller 2007; Müller 2010a).³ Agree is possible if a probe feature [$*F*$] c-commands a matching feature [F] or vice versa (Koopman 2006; Baker 2008; Riedel 2009) as long as [$*F*$] and [F] belong to the same phase.

Spell-out of syntactic structure is cyclic, proceeding in phases with designated phase heads v, C and D (Chomsky 2000, 2001; see Svenonius 2004, Heck and Zimmermann 2004 for D) and successive-cyclic movement is enabled by

²For a discussion of whether derivations which involve fission require look-ahead, see section 4.2.

³There is no operation *Move* per se. Rather, Move is understood as re-Merge. For reasons of notational simplicity and the fact that the exact implementation of Move does not play a role here, re-Merge will be depicted in the derivations below by letting the same item appear in different positions.

edge features [$\bullet X \bullet$] (Chomsky 2000) which can be added to phase heads whenever the edge feature condition is met.

CED-based islands:

The analysis of CED-based islands is adapted from the theory of Müller (2010a), which predicts that categories become islands if they are last-merged categories in a phase. This prediction results from the following assumptions: probe ($[*F*]$) and structure-building ($[\bullet F \bullet]$) features on heads are ordered on stacks, i.e., only one feature is accessible to the derivation at a time. Assuming that Merge only deletes features $[\bullet F \bullet]$ and Agree only features $[*F*]$ and that every instance of these two operations must delete a feature (cf. *Last Resort*, Müller 2010a, 42), it follows that all Merge and Agree operations triggered by a head are strictly ordered. Now, in order to derive the result that only last-merged categories become islands, the following restrictions on edge feature insertion are imposed⁴:

- (7) *Edge feature condition (adapted from Müller 2010a, 42)*
 An edge feature [$\bullet X \bullet$] can be assigned to the head γ of a phase only if (a) and (b) hold:
- a. there are still operation-triggering features in the edge domain of the γ -phase.
 - b. [$\bullet X \bullet$] ends up on top of γ 's stack of operation triggering features.

As long as a phase head or a specifier of the phase head still has operation-triggering features, the phase head can host edge features, which allow extraction. But if a category α with no operation-triggering features is the last-merged specifier in a phase, it deletes the last operation-triggering feature of the phase head and with it the possibility for further edge feature insertion. Possessing no edge feature, the phase head can no longer trigger movement, such that no categories in α can move out of α . Hence, last-merged categories are expected to be islands for movement as long as they do not host operation-triggering features themselves.

Let us have a look at an example to show how the theory works in detail. The sentence in (8) contains an adjunct clause which is an island for movement preventing the *wh*-phrase *who* from leaving the adjunct clause.

- (8) *Who did Mary cry after John hit *t*? Huang (1982, 503)

The ungrammaticality of (8) can be derived as follows: the derivation starts by building the embedded clause *after John hit who*. Then, the matrix vP is derived based on the feature specification of *v* in (9-a): first, *v* has to merge with the matrix VP. This operation deletes the feature $[\bullet V \bullet]$ as shown in (9-b). After this step, the subject *Mary* is merged, deleting $[\bullet D \bullet]$ but not introducing new operation-

⁴The edge feature condition in Müller (2010a) differs from the one in (7) in that only operation-triggering features on the phase head γ itself may allow edge feature insertion. In (7), however, also features on specifiers of γ may do so. This extension of the edge feature condition is important for parasitic gap constructions where (adjunct) islands can be circumvented under certain circumstances, in contrast to (adjunct) islands in non-parasitic gap constructions. See sections 2.3 and 3.1 for derivations of parasitic gap constructions and how specifiers of a phase head can keep it active for edge feature insertion.

triggering features to the edge domain of the matrix vP (cf. (9-c)). Assuming that adjunction does not delete operation-triggering features, the phase head's feature stack is empty at the point when the temporal adverbial clause is adjoined to the matrix vP, and since there are no further operation-triggering features in the edge domain of the vP, no edge feature can be inserted on v.⁵ Without the insertion of an edge feature, however, *who* cannot escape the embedded clause.

- (9) a. *Feature specification of v*
 $v [\bullet V \bullet < \bullet D \bullet]$
 b. *Merge of VP deletes $[\bullet V \bullet]$*
 $[_v v [\bullet V \bullet < \bullet D \bullet] [_{VP} cry]]$
 c. *Merge of subject deletes $[\bullet D \bullet] \rightarrow v$ becomes inactive*
 $[_v Mary [_v v [\bullet D \bullet] [_{VP} cry]]]$
 d. *Merge of adjunct clause $\rightarrow v$ is already inactive*
 $[_{VP} [_v Mary [_v v [] [_{VP} cry]]]] [_{CP} \text{ after John hit who}]]$

2.2 Prederivational fission

In what follows now, I will introduce a fission mechanism that manipulates the feature specification of certain lexical items in the numeration. Because of this feature manipulation, the behavior of fissioned items and their interaction with other items in the derivation is expected to differ from the behavior and interaction properties of items that have not undergone fission. Afterwards, I will show how this fission operation combined with the background assumptions presented in section 2.1 derives parasitic gap constructions.⁶

As has been widely discussed, parasitic gaps show a behavior that differs drastically from other non-parasitic gaps.⁷ I propose that this special behavior of parasitic gaps and the fact that something like a parasitic gap exists at all both come from the possibility of having imperfect numerations. Imagine the following scenario: instead of forming a numeration which leads to a converging derivation, we have a numeration where there are fewer items than needed to guarantee a successful derivation. I would like to suggest that these defective numerations lead to parasitic gap constructions.

⁵Note that even if Merge of adjunct clauses were triggered by a structure-building feature $[\bullet C \bullet]$ (Alexiadou 1994, 1997; Cinque 1999, 2004), the feature would be deleted when the adjunct clause is merged, so that no edge feature can be inserted on v afterwards.

⁶The following analysis is adapted from Assmann (2010), where a version of the present theory was first developed. The approach outlined here differs from Assmann (2010) in two points. First, the prederivational operation that applies in parasitic gap constructions was called ‘duplication’ and was described as a copy operation. Although the idea and the effects of the two operations are the same, the term ‘fission’ is more adequate to describe the operation and is supposed to make the idea of the process clearer. Second, *duplication* introduced a new type of feature $[\triangleright F \triangleleft]$ into the grammar. This additional feature type has been dismissed and replaced by normal probe features $[*F*]$, thereby simplifying the theory. I would like to thank Fabian Heck and Gereon Müller for this suggestions.

⁷See e.g. Postal (1993) about the differences between parasitic gap constructions and across-the-board movement. See also section 3.

(10) *Working hypothesis*

Sentences with parasitic gaps are based on defective numerations.

Specifically, I claim that a parasitic gap sentence as in (11) has a numeration as shown in (12).

(11) Which article did you file *t* [before reading *pg*]?

(12) *Numeration of (11):*

N = [<i>which</i>	[D,	ϕ ,	wh, •N•, ..., PHON ⁸ , SEM],
	<i>article</i>	[N, ...],
	C	[C, •wh•, •T•, ...],
	<i>you</i>	[D,	ϕ ,	...],
	T	[* ϕ *,	T, EPP, •v•, ...],
	v	[•D•,	* ϕ *,	v, •V•, ...],
	<i>file</i>	[•D•,		V, ...],
	Op _{temp} ⁹	[...],
	<i>before</i>	[C, •T•, ...],
	T	[* ϕ *,	T, EPP, •v•, ...],
	PRO	[D,	ϕ	...],
	v	[•D•,	* ϕ *,	v, •V•, ...],
	<i>read</i>	[•D•,		V, ...]]

If we look at the feature specifications of the lexical items in (12), we can see that we have four features [•D•] and four features [* ϕ *].¹⁰ These features can only be checked by their respective counterparts [D] and [ϕ]. Of these counterparts, however, we only have three features each. So, what renders this numeration defective is that exactly one lexical item with the features [D, ϕ] is missing. No matter how the derivation proceeds otherwise, it is bound to crash since the uninterpretable features [•D•, * ϕ *] cannot be deleted (cf. *Full Interpretation*, Chomsky 1986b). Put differently, an additional instance of one of the D items is the only possibility to save the numeration. This possibility is enabled by a fission operation as defined in (13).¹¹

⁸ The features PHON and SEM are labels for feature sets. ‘PHON’ encompasses all phonological features and ‘SEM’ all semantic features. Assuming a realizational morphology like Distributed Morphology (Halle and Marantz 1993) which adopts the principle of late insertion, the phonological features will be realized postsyntactically.

⁹ Following Larson (1988), I assume that temporal adverbial clauses contain a temporal operator. For reasons of simplicity, I take this operator to be base-generated in Spec,CP even though it is actually moved from a lower position. See Larson (1988); Dubinsky and Williams (1995).

¹⁰ Following Chomsky (2000), I assume that case assignment is a by-product of ϕ -Agree. Thus, case is not a feature per se and therefore not part of the feature specification of the nominal categories in (12).

¹¹ The idea of fission as an operation that conditions syntactic derivations is not new but has already been explored in depth by Agbayani (1998) where fission is an operation that applies inside the derivation. Agbayani and Ochi (2007) discuss how fission can be used to describe parasitic gaps. However, the fission operation I am proposing differs from the one of Agbayani and Ochi (2007) in important respects, and it makes different predictions (cf. section 3).

- (13) *Fission* ($N = [L, \dots]$)
- a. There are structure-building and probe features $[\bullet F_1 \bullet], \dots [\bullet F_i \bullet], [*F_j *], \dots [*F_n *]$ in the numeration N that do not have matching features $[F_1], \dots [F_n]$.
 - b. There is a lexical item L in N that has features $[F_1, \dots, F_n, G_1, \dots, G_m]$.
 - c. L fissions into two items L_1 and L_2 , whereby L_2 has the features $[F_1, \dots, F_n]$ and L_1 has the features $[*F_1 *, \dots, *F_n *, G_1, \dots, G_m]$.

The first condition (13-a) under which fission applies is exactly the situation found in the numeration in (12): there are operation-triggering features which cannot be checked. The second condition (13-b) demands that the numeration has matching features available at all. If these two conditions are met, fission applies with the result that there is an additional item which enables a convergent derivation. Furthermore, the definition of the fission operation in (13) forces the two parts to fuse again in the derivation by inserting probe features on one part which must be checked by its counterpart.

Assume that, in our example numeration in (12), fission will affect the lexical item *which*, creating the items *which*₁ and *which*₂.

- (14)
- a. *Before fission*
which $[D, \phi, wh, \bullet N \bullet, \dots, PHON, SEM]$
 - b. *After fission*
*which*₁ $[*D *, *\phi *, wh, \bullet N \bullet, \dots, PHON, SEM]$
*which*₂ $[D, \phi]$

Now, having two instances of *which*, the defective numeration is repaired and the derivation of the sentence can start.

But before I continue with the derivation of the sentence in (11), three notes are in order. First, note that the phonological features remain with *which*₁, i.e., *which*₂ is phonologically empty. The fission operation in (13) demands that only features which are necessary for the convergence of the derivation are fissioned. Since phonological features are not needed—otherwise, empty categories should not be possible at all—all phonological features remain with one part of a fissioned item with the result that the other part is phonologically empty. In fact, if the phonological features fissioned as well, they would become a part of the checking configuration between the two parts of this item so that a part of the phonological features would be deleted. This would result in a deviant realization of the item (cf. footnote 8). Hence, we would falsely predict that the pronunciation of lexical items depends on whether it is the antecedent of a parasitic gap or not.

Second, fission in (14) is simplified regarding semantic features. Whatever is necessary to guarantee a successful interpretation, e.g. a referential index, also takes part in fission.¹²

¹²See Browning (1989); Řezáč (2004) for theories which assume that the ϕ -features of a nominal category are linked to the referential index. Under this analysis, the referential index together with the ϕ -features would be present on both items after fission which furthermore ensures that the probe features on *which*₁ can only be checked by *which*₂.

Third, the fission operation in (13) is defined as a last-resort operation. Alternatively, a new numeration could be computed which contains an additional pronoun that must be bound by *which*. Such a numeration leads to a sentence where the position of the parasitic gap is occupied by a pronoun.

2.3 A sample derivation

In this section, I will integrate the fission operation introduced above in (13) into the general framework outlined in section 2.1. In doing so, I will go through a detailed derivation of (11), repeated in (15), which is based on the numeration after fission of the lexical item *which* has applied.

(15) Which article did you file *t* [before reading *pg*]?

Assuming that numerations consist of subarrays of lexical items which are activated one after the other (Chomsky 2000), the first subarray of the numeration of (15) contains all items that belong to the vP phase of the embedded adverbial clause (16). At first, *read* merges with *which*₂, i.e. the item that only has a categorial feature and ϕ -features. Then, *v* merges with VP which deletes the feature [$\bullet V \bullet$] on *v* followed by ϕ -Agree with *which*₂ deleting the feature [$\ast \phi \ast$]. Now, an edge feature is inserted on *v* which allows movement of *which*₂ to the edge domain of the vP. Finally, *v* merges with PRO whereby its last operation-triggering feature [$\bullet D \bullet$] is deleted.¹³ The development of the feature stack of the phase head *v* during these steps of the derivation is shown in (17) and the structure in (18).

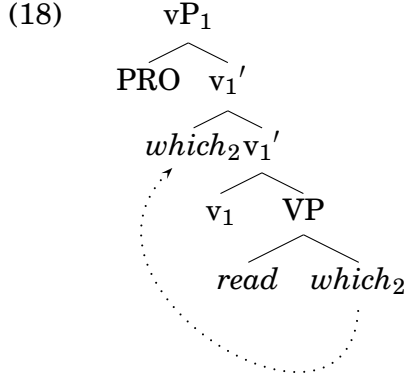
(16) Subarray vP_1
 $\{v_{P_1} \text{ read}, \text{which}_2, v_1, \text{PRO}\}$

(17) Feature stack of v_1

a.	$v_1 [\bullet V \bullet < \ast \phi \ast < \bullet D \bullet]$	→ (Merge of VP)
b.	$v_1 [\bullet \cancel{V} \bullet < \ast \phi \ast < \bullet D \bullet]$	→ (ϕ -Agree)
c.	$v_1 [\ast \phi \ast < \bullet D \bullet]$	→ (edge feature insertion)
d.	$v_1 [\bullet X \bullet < \bullet D \bullet]$	→ (movement of <i>which</i> ₂ to Spec, vP)
e.	$v_1 [\bullet \cancel{X} \bullet < \bullet D \bullet]$	→ (Merge of PRO)
f.	$v_1 [\bullet \cancel{D} \bullet]$	

¹³The fact that intermediate movement steps precede external Merge of the outermost specifier, i.e., that in this case *which*₂ ‘tucks in’ below the subject PRO, follows directly from the *Intermediate Step Corollary* (Müller 2010b, 5).

(i) *Intermediate Step Corollary:*
Intermediate movement steps to specifiers of X (as required by the PIC) must take place before a final specifier is merged in XP.

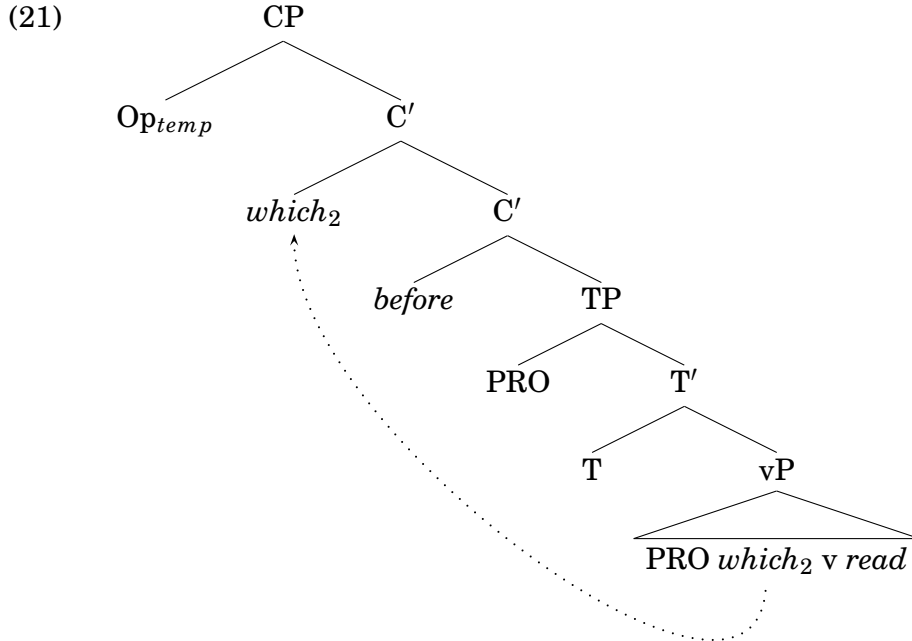


Next, the subarray containing the items of the embedded CP is selected (19). Here, the first step is Merge of T and vP₁. PRO then ϕ -agrees with T and moves to Spec,TP in order to satisfy T's EPP property. After this step, the complementizer *before* is merged with TP. Since *before* has only one operation-triggering feature left, namely [\bullet Op \bullet], an edge feature can be inserted, which again enables movement of *which*₂ to the edge of the CP phase. Finally, the temporal operator is merged and the CP phase is completed. The feature stack of *before* is shown in (20) and the final structure of the adjunct clause in (21).

- (19) Subarray CP₁
 {CP₁ T, *before*, Op}

- (20) Feature Matrix of C

- | | | |
|----|--|--|
| a. | <i>before</i> [\bullet T \bullet < \bullet Op \bullet] | |
| b. | <i>before</i> [\bullet T \bullet < \bullet Op \bullet] | → (Merge of TP) |
| c. | <i>before</i> [\bullet X \bullet < \bullet Op \bullet] | → (edge feature insertion) |
| d. | <i>before</i> [\bullet X \bullet < \bullet Op \bullet] | → (movement of <i>which</i> ₂ to Spec,CP) |
| e. | <i>before</i> [\bullet Op \bullet] | → (Merge of Op _{temp}) |



Afterwards, the subarray of the matrix vP is activated (22). Before this step, the subarray of the DP has been activated and *which*₁ has merged with *article*. The

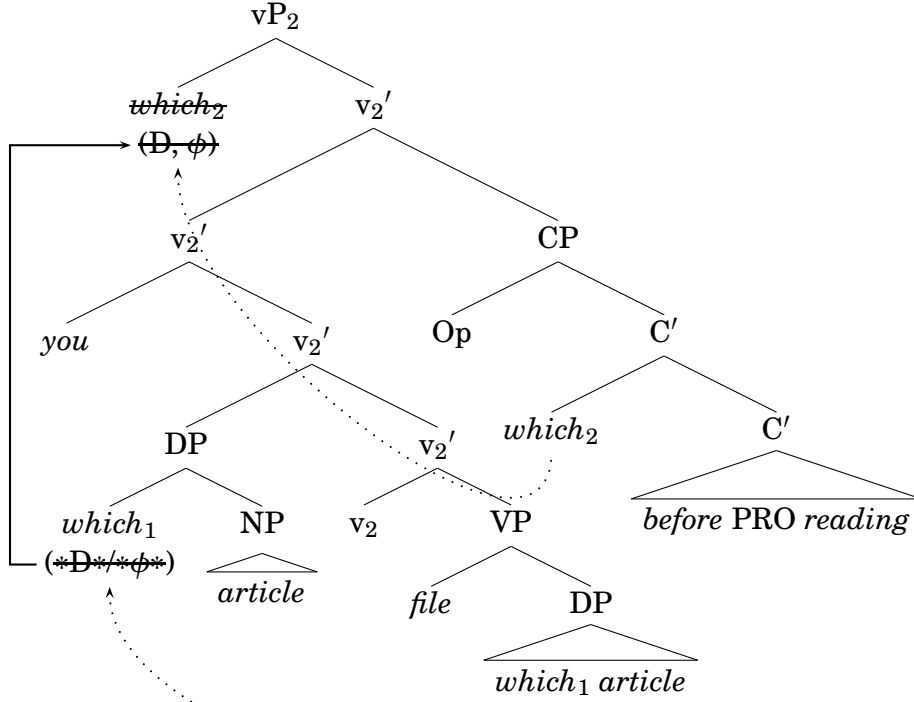
resulting DP *which*₁ *article* is now merged with *file*. Then, *v* merges with VP and agrees with *which*₁ *article*. Now, an edge feature is inserted which is deleted by movement of *which*₁ *article* to Spec,vP. Next, *you* merges with *v'* and the adjunct clause adjoins to *v'*. Although at this point, *v* itself has no operation-triggering features anymore, there is one item in its edge domain that has such features, namely *which*₁. Due to fission, *which*₁ has probe features that must be checked by *which*₂. Adopting the edge feature condition in (7), these probe features on *which*₁ now allow edge feature insertion on *v* (see (23-g)) and thus enable movement of *which*₂ to Spec,vP, creating a configuration where *which*₁ can agree with *which*₂ in categorial and ϕ -features, i.e., Fusion applies.¹⁴ Assuming that elements can be deleted under recoverability (Chomsky 1980)—which is a given when all features of one element, here *which*₂, have entered into Agree with another designated item, here *which*₁—*which*₂ can be deleted (see Roberts 2010 for related ideas). The development of the feature stacks after each derivational step is shown in (23). The final structure of the matrix vP is given in (24).

- (22) *Subarray vP*₂
 $\{vP_2 \text{ file, } which_1, v_2, you\}$

- (23) *Feature Matrix of v*
- a. $v_2 [\bullet V \bullet < * \phi * < \bullet D \bullet]$ → (Merge of VP)
 - b. $v_2 [\bullet \bar{V} \bullet < * \phi * < \bullet D \bullet]$ → (Agree with *which*₁ *article*)
 - c. $v_2 [* \phi * < \bullet D \bullet]$ → (edge feature insertion)
 - d. $v_2 [\bullet X \bullet < \bullet D \bullet]$ → (movement of *which*₁ *article* to Spec,vP)
new features in the edge domain of *v* (*D*/ ϕ *)
 - e. $v_2 [\bullet \bar{X} \bullet < \bullet D \bullet], which_1 [*D*/\phi*]$ → (Merge of *you*)
 - f. $v_2 [\bullet \bar{D} \bullet], which_1 [*D*/\phi*]$ → (Adjunction of adjunct clause)
 - g. $v_2 [], which_1 [*D*/\phi*]$ → (edge feature insertion)
 - h. $v_2 [\bullet X \bullet], which_1 [*D*/\phi*]$ → (movement of *which*₂ to Spec,vP)
 - i. $v_2 [\bullet \bar{X} \bullet], which_1 [*D*/\phi*]$ → (Agree between *which*₁ and *which*₂)
 - j. $v_2 [], which_1 [*D*/\phi*]$

¹⁴It might be argued that Agree between *which*₁ and *which*₂ is impossible due to minimality since the subject is a closer potential goal (cf. (24)). Note, however, that the subject bears a different referential index than *which*₁. Since *which*₁ can only check its probe features with a referentially identical goal (see section 2.2 and footnote 12), the subject cannot undermine Agree in this case.

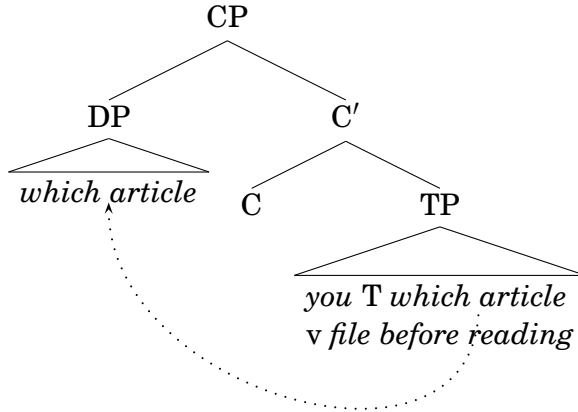
(24)



Note that Agree between *which*₁ and *which*₂ is possible since *which*₂ c-commands *which*₁ and both are specifiers of the same phase (cf. the conditions of Agree in section 2.1). If *which*₂ had not moved out of the embedded CP, the probe features on *which*₁ would not be deleted and the derivation would crash.¹⁵

Finally, the last subarray containing the items of the matrix CP phase is selected. The derivation continues until C is merged, carrying a feature [\bullet wh \bullet] that is checked by *which article*. The final structure of (11) is given in (25).

(25)



2.4 Interim summary

In this section, I developed a new approach to parasitic gaps which is based upon a prederivational fission operation. I have proposed that sentences with parasitic

¹⁵Note that if Merge of the two parts of *which* were vice versa, i.e., *which*₁ had been base-generated in the embedded clause and *which*₂ in the matrix clause, the derivation would crash as well due the fact that the probe features on *which*₁ could not be deleted. This issue is discussed in detail in section 4.2.

gaps originate from numerations which do not have enough lexical items to enable a converging derivation. If this is the case, one item of the numeration has to fission into two in order to enter the derivation twice, thereby preventing the numeration from crashing. Nevertheless, the fission of a lexical item is not without costs: only if the two parts can fuse back together into one item is the derivation successful. This fusion of categories has been implemented as full Agree with deletion under recoverability.

Thus, the main result so far is that we can answer the general questions about parasitic gaps in (4) at the end of section 1:

1. *Which syntactic operation do parasitic gaps originate in if not movement?*

Parasitic gaps are the result of a prederivational fission operation which, on the one hand, allows a lexical item to occur in two different base positions by fissioning the item but imposes a condition of fusion on the other hand. Fusion is implemented as full Agree of the two parts where one part can be deleted under recoverability. Since the configuration for this Agree operation is only given when the two parts belong to the same phase, both parts have to move to higher positions which are specifier positions of *v* in the case above.

2. *What are parasitic gaps?*

A parasitic gap is a gap whose antecedent is the deleted part of a formerly fissioned lexical item.

With this background, we are now able to answer further questions about parasitic gaps concerning their apparently idiosyncratic behavior.

3 Answering more questions

One of the main goals of this paper is to find solutions to four important conundrums parasitic gap constructions pose, which were listed in (5). In what follows, I will give a detailed introduction to each problem and show how the fission approach to parasitic gaps can solve it.

3.1 Conundrum I: Why are parasitic gaps ‘selectively’ sensitive to islands?

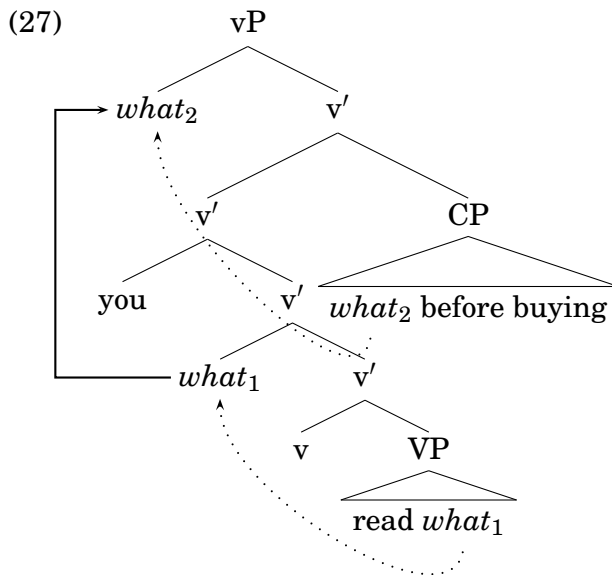
The paradoxical property of parasitic gaps to be separable from their antecedent by one island but not by more than one island (Kayne 1983, 1984; Bennis and Hoekstra 1985; Longobardi 1984; Koster 1986; Chomsky 1986*a*; Manzini 1994; Nunes 1995, 2001, 2004) is shown in (26).¹⁶

- (26) a. What did you read *t* [_{isl} before buying *pg*]?
 b. *What did you read *t* [_{isl} after expecting me to call the editor [_{isl} before buying *pg*]]?

¹⁶This property of parasitic gap constructions can be observed in other languages as well, e.g. German.

In (26-a), *what* is separated from the parasitic gap by one adjunct island while in (26-b) there are two islands between them. Why do islands matter in (26-a) but not in (26-b)? The solution to this puzzle within the fission approach of section 2 falls out from the conditions imposed on fusion of the two parts of the antecedent.

Note that the derivations of the two sentences in (26) are identical up to a certain point. In both cases, *what* fissions into *what*₁ and *what*₂, leaving *what*₁ with additional probe features that must be checked by *what*₂. Since *what*₂ is merged in an adjunct clause while *what*₁ is merged in the matrix clause and since fusion can apply only locally, *what*₂ must leave the adjunct clause, i.e., it must leave an island. In (26-a), this is possible since the following configuration is given:

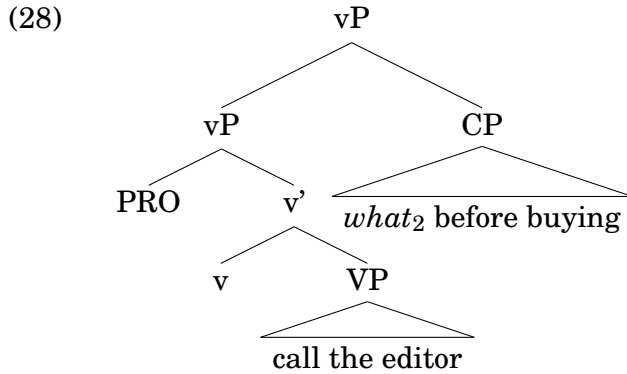


Remember from section 2.1 that the adjunct clause would be an island if there were no operation-triggering features in the edge domain of the vP phase at the point where the adjunct clause adjoins to v'. The reason is that movement from the adjunct clause is only possible if v has edge features which in turn can only be inserted in presence of other operation-triggering features (cf. (7)). Indeed, the v head has no such features since it has already checked its agreement features with *what*₁ and its selection feature [\bullet D \bullet] with the subject *you*. Due to fission, however, *what*₁ has operation-triggering features which are present in the edge domain of the vP phase as a result of an intermediary movement step to Spec,vP, which is needed for independent reasons given that all movement proceeds from phase edge to phase edge (PIC, see Chomsky (2000)). These features can keep the phase head active, allowing edge feature insertion on v (cf. the edge feature condition in (7)).

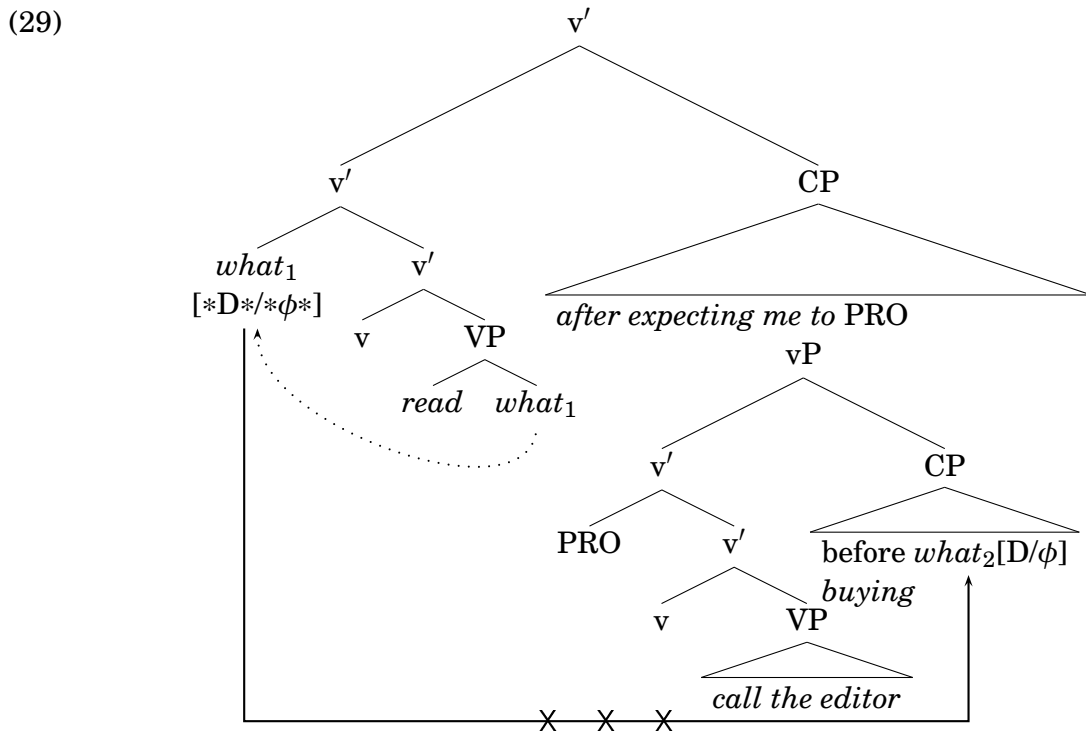
-
- (i) a. Welchen Artikel hast du [_{isl} ohne *pg* zu lesen] *t* abgeheftet?
 which article have you without to read filed
 'Which article did you file without reading?'
 b. *Welchen Artikel hast du [_{isl} ohne den Autor zu kennen [_{isl} der *pg* geschrieben
 which article have you without the author to know who written
 hat]] *t* abgeheftet?
 has filed
 'Which article did you file without knowing the author who wrote?'

With an edge feature comes the possibility for *what*₂ to leave the adjunct clause. Thus, the adjunct island in (26-a) can be circumvented which in turn gives the impression that parasitic gaps are not sensitive to islands.

Now, the derivation of (26-b) is basically the same as that of (26-a) with only one major difference: *what*₂ is further embedded than in (26-a). This difference, however, leads to a crash of the derivation. The critical point in the derivation is shown in (28).



Again, *v* has agreement features which are checked by *the editor*. The structure-building feature [$\bullet D \bullet$] of *v* is deleted by Merge of the covert subject PRO. But in contrast to the derivation of (26-a), there are no further operation-triggering features in the vP phase at the point when the lowest adjunct is adjoined. Hence no edge feature can be inserted and *what*₂ is stuck in the adjunct clause; any movement of *what*₂ would fatally violate the PIC. Thus, *what*₁ and *what*₂ can never be in a configuration where they can fuse, and the probe features on *what*₁ remain unchecked, leading to a violation of *Full Interpretation* (Chomsky 1986b) as shown in (29).



In summary, parasitic gaps are always sensitive to islands. But in the derivation of a parasitic gap sentence, the highest potential island is rendered accessible by the additional probe features on the antecedent of the parasitic gap, which are originally the result of fission, and which now show up in the very edge domain of the phase whose specifier needs to be targeted by the antecedent of the parasitic gap, given the PIC.

3.2 Conundrum II: Why can complex antecedents only be re-constructed into the real gap but not the parasitic gap?

Binding data as in (30) suggest that the position of the parasitic gap cannot host a complex DP in contrast to the real gap.

- (30) a. [Which books about himself₁]₂ did John₁ file t_2 before Mary read pg_2 ?
 b. *[Which books about herself₁]₂ did John file t_2 before Mary₁ read pg_2 ?
Kearney (1983)

In (30), the complex *wh*-phrase contains an anaphor which must be bound by an antecedent that agrees with the anaphor in gender. Having two potential base positions in two different clauses, the anaphor inside the *wh*-phrase has available two potential antecedents, namely the subjects of the respective clauses, of which, however, only one agrees with the anaphor in gender. In this context, the following asymmetry can be noticed: in (30-a), binding of the anaphor *himself* is possible but in (30-b), binding of the anaphor *herself* fails.

The fission approach offers a principled explanation for this binding asymmetry. In (30), fission of *which* creates the following two items.

- (31) *which* [D, ϕ , acc, wh, $\bullet N \bullet$, SEM, PHON, ...]
 a. *which*₁ [$\bullet D \bullet$, $\bullet \phi \bullet$, $\bullet acc \bullet$, wh, $\bullet N \bullet$, SEM, PHON, ...]
 b. *which*₂ [D, ϕ , acc]

Note that only one part, namely *which*₁, has the feature [$\bullet N \bullet$] which enables it to merge with the NP *books about him/herself*. The resulting DP *which*₁ *books about him/herself* is merged in the matrix clause where the anaphor can be bound by the subject *John*. But the subject of the adjunct clause *Mary* cannot bind the anaphor since it does not c-command it from its position inside the adjunct clause. Thus, the only antecedent available for the anaphor is the matrix subject but not the embedded subject. The configuration is shown in (32).

(32)

John

before Mary read

DP *which*₁...

v

VP

file *which*₁ book
about *himself*/**herself*

3.3 Conundrum III: Why are certain properties of the antecedent only important for its relation to the real gap but not to the parasitic gap?

Another asymmetry between the real and the parasitic gap shows up in contexts where gap and antecedent are separated by weak islands such as *wh*-islands.¹⁷ In German, for example, *wh*-movement is known to be sensitive to *wh*-islands, in contrast to other types of movement such as (argument) topicalization (Fanselow 1987; Müller and Sternefeld 1993).

-
- ¹⁷The asymmetries between real and parasitic gap which are illustrated in (34) and below in (36) have not been described in the literature so far. I would like to thank Gereon Müller for bringing these data to my attention.

In (33), the *wh*-phrase *wie* creates a weak island from which only non-*wh*-phrases can be extracted. Thus, the *wh*-phrase *welche Radios* cannot move out of the embedded clause while the topicalized non-*wh*-phrase *Radios* can do so.

If *wh*-islands are combined with parasitic gap constructions, we can observe that sentences are strongly deviant if the *wh*-island intervenes between the antecedent and the real gap (34-a) but not if it intervenes between the antecedent and the parasitic gap (34-b).¹⁸

- (34) a. *Welche Radios weißt du [CP wie man [ohne *pg* zu reparieren] *t*
 which radios know you how one without to repair
 verkauft]?
 sells
 lit.: ‘Which radios do you know how to sell without repairing?’
 b. ?Welche Radios hast du [ohne zu wissen [CP wie man *pg*
 which radios have you without to know how one
 repariert]] *t* verkauft?
 repairs sold
 lit.: ‘Which radios did you sell without knowing how to repair?’

Thus, it seems that the *wh*-property of the antecedent only matters for its relation to the real but not the parasitic gap. This is exactly what we expect, given that according to (14-b), *welche*₂ merged in the position of *pg* does not bear a [wh] feature after fission (and thus does not qualify as a *wh*-phrase), whereas *welche*₁ merged in the position of *t* is equipped with a [wh] feature after prederivational fission.

3.3.2 Long-distance scrambling

Another example for this behavioral asymmetry comes from scrambling. In German, scrambling is strictly clause-bound (Ross 1967) as shown in the contrast between (35-b) and (35-c).

- (35) a. dass ich glaube [CP dass der Mann_j das Radio_i reparieren muss]
 that I believe that the man the radio repair must
 ‘that I believe that the man has to repair the radio’
 b. dass ich glaube [CP dass das Radio_i der Mann_j *t*_i reparieren muss]
 that I believe that the radio the man repair must
 ‘that I believe that the man has to repair the radio’
 c. *dass ich das Radio_i glaube [CP dass der Mann_j *t*_i reparieren muss]
 that I the radio believe that the man repair must
 ‘that I believe that the man has to repair the radio’

Since scrambling is able to license parasitic gaps in German (see (36-a); Felix 1985; Bennis and Hoekstra 1985; Webelhuth 1992; Mahajan 1990; also see Assmann and Heck 2011 vs. Fanselow 2001), we can test whether the property of clause-boundedness is given in parasitic gap contexts.

¹⁸Note that the grammaticality judgements here are independent of the presence or absence of the *ohne*-clause including the parasitic gap.

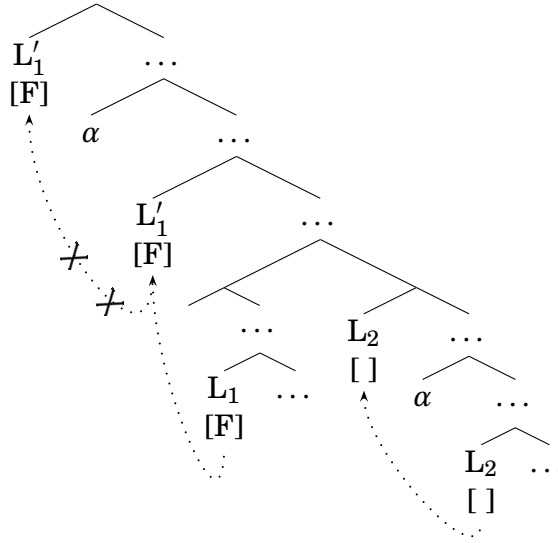
- (36) a. dass ich das Radio [ohne *pg* zu reparieren] *t* verkauft habe
 that I the radio without to repair sold have
 ‘that I have sold the radio without repairing it.’
- b. *dass ich das Radio glaube [_{CP} dass man [ohne *pg* zu reparieren] *t*
 that I the radio believe that one without to repair
 verkaufen kann]
 sell can
 ‘that I believe that one can sell the radio without repairing it’
- c. ?dass ich das Radio [ohne zu sagen [_{CP} dass man noch *pg*
 that I the radio without to say that one still
 reparieren muss]] *t* verkauft habe
 repair must sold have
 ‘that I have sold the radio without saying that one still has to repair it’

Obviously, the finite clause boundary does intervene between the antecedent *das Radio* and the real gap (36-b) but not between the antecedent and the parasitic gap (36-c). Assuming that scrambling is triggered by a scrambling feature [Σ] and movement of categories which bear a [Σ] feature may not cross a finite CP, we can explain the asymmetry illustrated in (36) as follows: the D head *das* is fissioned into two parts *das*₁ and *das*₂, of which only *das*₁ has the scrambling feature [Σ] (in analogy to the [wh] feature of *welche* in (34)). Then, movement of *das*₂ may freely cross finite CPs whereas *das*₁ bearing the [Σ] feature must not do so.

3.3.3 The pattern

The asymmetries we have seen so far seem to belong to a general pattern of asymmetries between the behavior of the two antecedent-gap dependencies. Assuming that movement types have certain properties because the item that is moved has a special feature [F] that encodes central aspects of the movement type—i.e., phrases to be *wh*-moved have a feature [wh]; phrases to be scrambled have a scrambling feature [Σ —the aforementioned pattern (which may include other cases as well) is predicted by the fission approach since the antecedent in a parasitic gap construction is fissioned into two parts whereby only one part has the feature [F] (L₁). This part, however, is merged in the position of the real gap. Hence, we end up with the following configuration.

(37)

(where α is an island for $[F]$)

Let α be a category that creates a weak island for the feature $[F]$. Then, only items that do not bear $[F]$ should be able to leave it. In other words, L_2 , which is the part of the antecedent L that is merged in the position of the parasitic gap, should be insensitive to the α -island since it has no feature $[F]$ while L_1 , which is the part merged in the position of the real gap, should be sensitive to α -islands. Hence, we predict that sentences such as (34-b) and (36-c) should be possible since neither does a *wh*-phrase cross a *wh*-island nor a scrambled phrase a finite clause boundary. On the other hand, sentences such as (34-a), (36-b) must be ungrammatical because the *wh*/ Σ -phrase L'_1 cannot move out of the α -island.

3.4 Conundrum IV: Why can parasitic gaps only be licensed by \bar{A} -movement but not by A-movement?

Finally, it is known that parasitic gaps can be licensed by movement types such as *wh*-movement, relativization, topicalization or scrambling, which are all \bar{A} -movement types, but not by movement types such as passive movement or raising, which are A-movement types (Engdahl 1983, 13).

- (38) a. *John was killed t by a tree falling on pg .
 b. *Mary seemed t to disapprove of John's talking to pg .

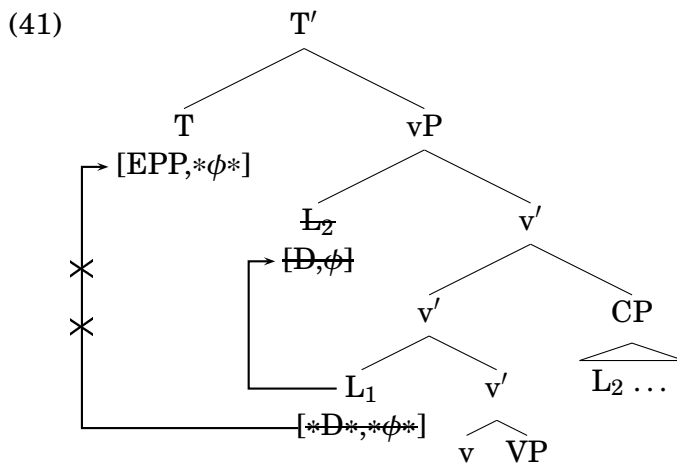
Interestingly, parasitic gaps cannot be licensed by a passive subject even if it moves further to an \bar{A} -position.

- (39) *Which house was sold t [before we could demolish pg]?
based on Legate (2003, 511)

Furthermore, parasitic gaps cannot be licensed from a subject position that c-commands it, whether it later moves to a legitimate position or not (*anti-command-condition*, see Taraldsen 1981; Chomsky 1982; Engdahl 1983; Safir 1987).

- (40) *Which spy t killed John before anybody could speak to pg ?
Safir (1987, 678)

The fission approach offers a simple solution to these puzzling data. In all the above cases, the antecedent of the parasitic gap has to land in or go through the subject position Spec,TP. To do so, the antecedent needs a categorial feature in order to satisfy T's EPP property, as well as ϕ -features which T must check. When the antecedent L of a parasitic gap construction is fissioned, these features are present on both parts L_1 and L_2 . On L_1 however, they are probe features which must be checked by the matching non-probe features on L_2 . Consequently, if L_1 and L_2 enter into full Agree, those features of L_1 which are necessary for occupying Spec,TP, are deleted. In other words, the derivations of the above sentences crash due to a timing problem in the derivation, with L_1 and L_2 agreeing before L_1 can move to Spec,TP. The configuration is shown in (41).



Thus, we expect that if movement to Spec,TP happens prior to fusion, subjects should be able to license parasitic gaps. In fact, this prediction is confirmed by data as in (42-a).

- (42) a. a note which [unless we send back *pg*] *t* will ruin our relationship
 b. *a note which *t* will ruin our relationship [unless we send back *pg*]
Haegeman (1984)

Assuming that the adjunct clause in (42-a) precedes the base position of the subject because it is merged higher in the tree, e.g., as an adjunct to the TP, it follows that the relative pronoun *which* moves to the subject position Spec,TP before the adjunct clause is even merged and thus before *which*'s feature specification is manipulated by fusion. In (42-b), however, we have the same configuration as in (41), with the result that *which* cannot license the parasitic gap in this case.

In summary, the fission approach provides straightforward answers to the four empirical questions discussed above without invoking additional stipulations. The fission operation itself, integrated into a standard minimalist framework, makes it possible not only to give a natural and intuitive explanation for the phenomenon of parasitic gaps but is, furthermore, able to predict the properties of parasitic gaps correctly.

4 Discussion

So far, we have seen how the fission approach to parasitic gaps works and how it answers to important questions concerning parasitic gaps and their properties. In this section, I will address more general empirical and conceptual aspects of the new approach in comparison to other analyses of parasitic gaps.

4.1 Discussion of empirical issues

What seems to be so puzzling about parasitic gaps is that most of their properties involve paradoxes or asymmetries. These paradoxes and asymmetries are resolved within the fission approach under the assumption that the antecedent of a parasitic gap occurs in different instances. In other words, the two feature manipulating operations fission and fusion have the effect that the antecedent L passes through three different stages in the derivation. The first two stages proceed independently from one another as L is fissioned into L_1 and L_2 (43-a,b). The third stage is the stage after fusion, i.e., feature deletion, has applied (43-c).

- (43) $L [F_1 \dots F_n, G_1 \dots G_n]$
 a. $L_1 [*F_1* \dots *F_n*, G_1 \dots G_n]$
 b. $L_2 [F_1 \dots F_n]$
 c. $L'_1 [G_1 \dots G_n]$

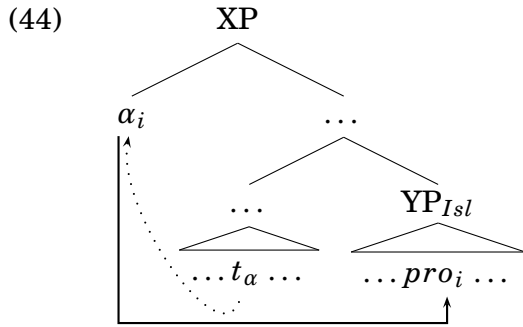
We thus even except that the properties of L differ according to the derivational stage L is in. So, whenever there are asymmetries such as the reconstruction asymmetries in section 3.2 or the island asymmetries in section 3.3, they can be traced back to the differences between the first and the second stage of L (L_1 vs. L_2).

Other asymmetries between antecedents of parasitic gap constructions and antecedents of simple non-parasitic gap constructions such as the \bar{A}/A -movement asymmetry in section 3.4 go back to the fact that only in parasitic gap constructions, the antecedent suffers from a drastic feature manipulation in the derivation that takes away its status as an argument and reduces it to a pure operator.

Thus, the asymmetries in the behavior of parasitic gaps discussed in section 3 provide strong support for the three derivational stages predicted by the fission approach.

Even though previous theories of parasitic gaps are to a certain degree empirically adequate as well, none of them is able to correctly predict all of the properties discussed in this paper. Basically, there are three types of parasitic gap analyses which have been argued for in the literature: the binding approach, the empty-operator-movement approach and the sideward movement approach.

Let us start by discussing the predictions made by the binding approach. Supporters of the binding approach claim that parasitic gaps are empty pronouns which are bound by the antecedent of the real gap (Chomsky 1982; Engdahl 1983, 1985; Cinque 1990; Postal 1993, 1994, 1998; Ouhalla 2001; Munn 2001). The configuration which shows this type of analysis is given in (44).

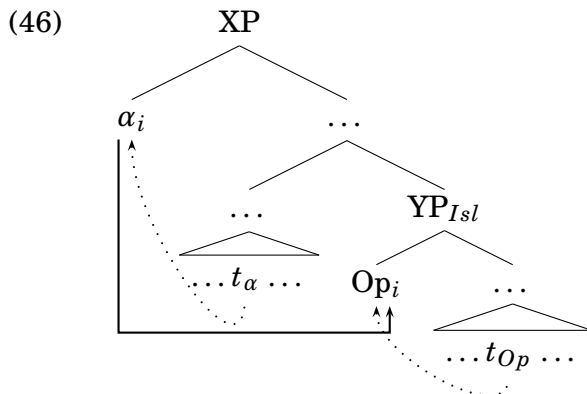


The binding approach is able to explain the reconstruction asymmetries (section 3.2) as well as the weak-island asymmetries (section 3.3): the only movement dependency is between the antecedent and the real gap. Hence, reconstruction can only be possible into the real gap and weak islands can only be crossed by the antecedent moving from the real gap position to its final position. Nevertheless, the binding approach cannot offer a principled solution to the other two puzzles discussed in section 3. First, there is no reason why an antecedent could not be able to bind a pronoun from an A-position, i.e., we expect that A-movement should license parasitic gaps, contrary to fact. Even if such configurational issues could be solved by stipulating that binding can only take place from an \bar{A} -position (e.g. Ouhalla 2001), data as in (45), repeated from (39), where the passivized subject moves further to an \bar{A} -position cannot be captured.

(45) *Which house was sold *t* [before we could demolish *pg*]?

Furthermore, the binding approach cannot solve the island paradox with parasitic gaps (cf. section 3.1) as binding is insensitive to islands in general (Chomsky 1977; Zaenen et al. 1981; Reuland 2001). For these reasons, the binding approach can be considered empirically inferior to the fission approach.

The empty-operator-movement approaches (see e.g. Contreras 1984; Kiss 1985; Chomsky 1986a; Lee 1998; Nissenbaum 2000) states that an empty operator is merged in the position of the parasitic gap which moves to a higher position where it is identified with the antecedent of the real gap; see (46).



Since this analysis involves actual movement of the operator from the parasitic gap position to a higher position, the selective island sensitivity of parasitic gaps can be explained. The highest island does not intervene between the antecedent

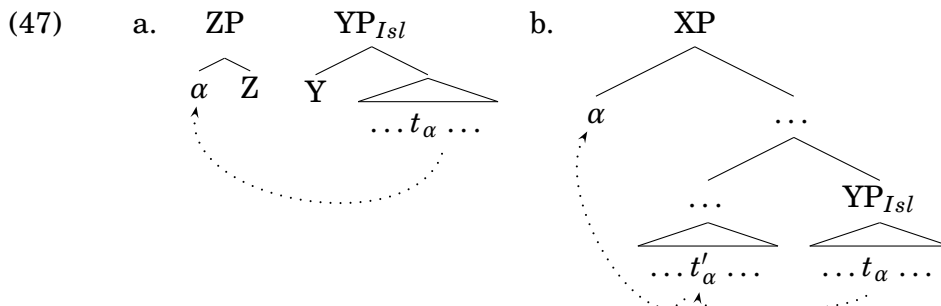
and the parasitic gap since there is no movement out of the island. Rather, the operator lands in the left periphery of this island. Nevertheless, all other islands would have to be crossed by movement.

Furthermore, the asymmetries between the real and the parasitic gap may in principle be captured since both gaps have, in fact, different antecedents. If one were to assume that the empty operator does not bear [wh] or [Σ], it may be insensitive to weak islands, predicting that *wh*-islands and finite clause boundaries should not fatally intervene between the antecedent of the real gap and the parasitic gap. And since the position of the parasitic gap is occupied by an empty operator, the antecedent of the real gap cannot be reconstructed into the parasitic gap, which provides an explanation for the reconstruction asymmetries in section 3.2.

However, data as in (45) which show that subjects cannot license parasitic gaps may be more problematic. In principle, the approach could involve configurational conditions that prohibit empty operators from being identified with categories in A-positions. Nonetheless, only ad hoc stipulations may prevent an antecedent which is in an otherwise legitimate \bar{A} -position but has moved through an A-position (cf. (45)) from licensing an empty operator.¹⁹

Thus, the fission approach has an advantage over the empty-operator-movement approach since it is able to explain these data without further ado.

Finally, the sideward movement approach (Nunes 1995, 2001, 2004) suggests that the antecedent of a parasitic gap construction is base-generated in the position of the parasitic gap and then moves interarborally to the position of the real gap, from where it moves on to its final position.



Within the framework Nunes (2001) assumes, the sideward movement approach can capture the selective island sensitivity by imposing conditions on the timing of operations in the derivation such that sideward movement itself can only apply in very restricted contexts: it may cross exactly one island but never more than one island. The fact that subjects cannot license parasitic gaps falls out from invoking additional stipulations about the licensing of traces by heads which define intervention effects whenever a subject licenses a parasitic gap.

¹⁹However, the main problem of the empty-operator-movement approach is that it cannot provide a principled explanation for the unexpected distribution of the empty operator. If parasitic gap constructions involved empty operator movement, we would expect that either overt operators can alternatively show up in the same position as the empty operator (similarly to the alternation between empty and overt operators in relative clauses), contrary to fact, or that empty operator movement is covert, which would render the explanation for the selective island sensitivity of parasitic gaps problematic since covert movement is known to be generally insensitive to islands (Huang 1982; Chomsky 1995).

The major problem for the sideward movement approach is, however, asymmetries between the real and the parasitic gap. Since the sideward movement approach claims that there is no difference between these two gaps, no asymmetries of any kind should exist. In particular, as the antecedent with all its properties and possible complexity is generated in both gaps, the antecedent should (i) be able to reconstruct in both positions to the same extent²⁰ and (ii) be sensitive to weak islands no matter where they occur. Hence, the data discussed in section 3.2 and 3.3 strongly argue against the sideward movement approach.

In conclusion, we have seen that the four properties of parasitic gaps which are directly predicted by the fission approach pose severe problems for other analyses of parasitic gaps. For each of the three types of accounts discussed here, there is at least one property which cannot be explained. Thus, on empirical grounds, the fission approach is superior to other theories of parasitic gaps.

4.2 Discussion of theoretical issues

Finally, I would like to discuss theoretical aspects of the fission approach in comparison to the other three approaches. In contrast to the approaches discussed above, the fission approach is based on a syntactic operation that seems at first sight not to be needed for independent reasons. Binding, empty operator movement and sideward movement are operations which are used to derive other syntactic phenomena such as resumption, relativization or across-the-board movement respectively. Note, however, that the extent to which the fission operation can derive phenomena besides parasitic gaps is not explored in full depth yet.²¹ I leave this issue to further research.

Another aspect that is worth discussing concerns the issue of look-ahead. There are two potential look-ahead problems in derivations which involve fission. The first one concerns the fission operation itself. It might be argued that in order to meet the conditions on fission, the course of the derivation must be known since only then, a crash of the derivation can be detected. However, this is not the case. For fission to apply, it is sufficient to run through the numeration and count operation-triggering and non-operation-triggering features. Assuming that numerations are nothing else but a set consisting of lexical items which are in turn sets of features, a scan runs through the numeration, building possible pairs of operation-triggering and matching features independently from what will be the correct pairing in the derivation. If look-ahead is defined as access to later stages of the derivation, the

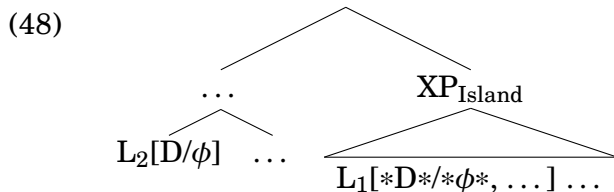
²⁰Nunes (2001) discusses data from Munn (1994) which show that reconstruction into the parasitic gap is possible when the category containing the parasitic gap precedes the real gap. Thus, he claims that reconstruction asymmetries do not pose a problem for his account. Note, however, that the sideward movement approach predicts that there should be *no* reconstruction asymmetries at all between the two gaps, contrary to fact.

(i) a. *Which picture of herself did every boy who saw *pg* say Mary liked *t*?
b. Which picture of himself did every boy who saw *pg* say Mary liked *t*? Munn (1994, 407)

²¹Agbayani and Ochi (2007), e.g., discuss how fission could also be used to analyze scrambling as base generation. The present fission approach may also be suited for analyzing other phenomena such as resumption and/or headless relative clauses.

prederivational scan sketched here is not look-ahead since the actual feature pairing in the derivation is immaterial for this scan, in which operation-triggering and non-operation-triggering features can be paired freely. Note, by the way, that such a prederivational scan does not increase the computational complexity of the derivation either as it has a worst-case complexity class $O(n^2)$, where n is the number of features in the numeration, just like the derivation. Since the prederivational scan as well as the derivation itself is a pairing of operation-triggering and matching features, it needs at most n^2 steps.²²

The second putative look-ahead problem concerns the question of whether it is necessary that the two parts L_1 and L_2 of a fissioned lexical item L be merged in designated positions. Put differently, one wonders what would happen if the merge positions of the two parts were reversed, putting L_2 in the position of the real gap and L_1 in the position of the parasitic gap. In this case the following scenario emerges:



In this configuration, the probe features of L_1 cannot be checked since Agree is only possible under c-command. Therefore, this derivation is excluded for independent reasons.

In conclusion, fission is a syntactic operation which needs further investigation but is conceptually and empirically appealing, since it leads to natural explanations for the properties of parasitic gaps without putting an additional computational burden on the derivation.

5 Conclusion

The aim of this paper was to find answers to the major questions concerning parasitic gaps. I have argued that parasitic gap constructions result from a prederivational fission operation which enables an item to enter the derivation twice starting in two different positions. This operation combined with an obligatory fusion of the two parts in the derivation leads to constructions in which one antecedent can be traced back to two different base positions. Furthermore, I have shown that the fission operation integrated into a minimalist framework can explain the idiosyncratic behavior of parasitic gaps, something which other approaches to parasitic gaps are not able to do to the same extent.

²²In fact, the actual complexity class will be lower than $O(n^2)$ due to the fact that features in the numerations are part of lexical items and that each feature can only interact in one operation.

Acknowledgements

I am indebted to Brian Agbayani, Petr Biskup, Doreen Georgi, Fabian Heck, Stefan Keine, Gereon Müller, Marc Richards, Martin Salzmann, Michael Sappir, Philipp Weisser, the participants of the Colloquium ‘Neuere Arbeiten zur Grammatiktheorie’ of the department of linguistics, at the University of Leipzig, the audience of the ConSOLE XIX in Groningen, 2011, and the 37th “Tagung zur Generativen Grammatik des Südens” (GGS) in Stuttgart, 2011, as well as an anonymous reviewer for helpful advice, comments, critique and discussion.

This research was partly funded by the DFG (Deutsche Gemeinschaft für Forschung) within the project *Lokale Modellierung nicht-lokaler Abhängigkeiten in der Syntax* (‘local modeling of non-local dependencies in syntax’).

<http://www.uni-leipzig.de/~lomo>

Anke Assmann

Universität Leipzig

Institut für Linguistik

anke.assmann@uni-leipzig.de

<http://www.uni-leipzig.de/~assmann>

References

- Abney, Steven (1987): The English Noun Phrase in Its Sentential Aspect. PhD thesis, MIT, Cambridge, Massachusetts.
- Agbayani, Brian (1998): Feature Attraction and Category Movement. PhD thesis, University of California, Irvine.
- Agbayani, Brian and Masao Ochi (2007): 'Split Lexical Insertion in Parasitic Gap Constructions', Ms. University of California, Fresno and University of Osaka.
- Alexiadou, Artemis (1994): Issues in the Syntax of Adverbs. PhD thesis, University of Potsdam.
- Alexiadou, Artemis (1997): *Adverb Placement. A case Study in Antisymmetric Syntax*. John Benjamins, Amsterdam.
- Assmann, Anke (2010): Parasitic Gaps in Derivational Grammar. Master's thesis, University of Leipzig.
- Assmann, Anke and Fabian Heck (2011): Einige Argumentasymmetrien im Deutschen und ihre Konsequenzen für die Theorie der freien Wortstellung. Ms. Universität Leipzig.
- Baker, Mark C. (2008): *The Syntax of Agreement and Concord*. Cambridge University Press, Cambridge.
- Bennis, Hans and Teun Hoekstra (1985): 'Gaps and Parasitic Gaps', *The Linguistic Review* 4, 29–87.
- Browning, Marguerite (1989): Null Operator Constructions. PhD thesis, MIT, Cambridge, Massachusetts.
- Chomsky, Noam (1973): Conditions on Transformations. In: S. Anderson and P. Kiparsky, eds, *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.
- Chomsky, Noam (1977): On Wh-Movement. In: P. W. Culicover, T. Wasow and A. Akmajian, eds, *Formal Syntax*. Academic Press, New York.
- Chomsky, Noam (1980): 'On Binding', *Linguistic Inquiry* 11(1), 1–46.
- Chomsky, Noam (1982): *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press, Cambridge, Massachusetts.
- Chomsky, Noam (1986a): *Barriers*. MIT Press, Cambridge, Mass.
- Chomsky, Noam (1986b): *Knowledge of Language: Its Nature, Origins, and Use*. Praeger, New York.
- Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Massachusetts.
- Chomsky, Noam (2000): Minimalist Inquiries: The Framework. In: R. Martin, D. Michaels and J. Uriagereka, eds, *Step By Step: Essays In Syntax in Honor of Howard Lasnik*. MIT Press, pp. 89–155.
- Chomsky, Noam (2001): Derivation by Phase. In: M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.
- Chomsky, Noam (2008): On Phases. In: R. Freidin, C. P. Otero and M. L. Zubizarreta, eds, *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*. MIT Press, Cambridge, MA, pp. 133–166.
- Cinque, Guglielmo (1990): *Types of A'-Dependencies*. MIT Press, Cambridge, Mass.
- Cinque, Guglielmo (1999): *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. Oxford University Press, New York and Oxford.
- Cinque, Guglielmo (2004): 'Issues in Adverbial Syntax', *Lingua* 114, 683–710.
- Contreras, Heles (1984): 'A Note on Parasitic Gaps', *Linguistic Inquiry* 15(4), 698–701.
- Culicover, Peter W. (2001): Parasitic Gaps: A History. In: P. D. Culicover and P. M. Postal, eds, *Parasitic Gaps*. MIT Press, Cambridge, Mass., chapter 1, pp. 3–68.
- Dubinsky, Stanley and Kemp Williams (1995): 'Recategorization of Prepositions as Complementizers: The Case of Temporal Prepositions in English', *Linguistic Inquiry* 26(1), 125–137.
- Engdahl, Elisabet (1983): 'Parasitic Gaps', *Linguistics and Philosophy* 6, 5–34.
- Engdahl, Elisabet (1985): 'Parasitic Gaps, Resumptive Pronouns, and Subject Extractions', *Linguistics* 23, 3–44.
- Fanselow, Gisbert (1987): *Konfiguralität*. Narr, Tübingen.
- Fanselow, Gisbert (2001): 'Features, Θ -Roles and Free Constituent Order', *Linguistic Inquiry* 32(3), 405–437.
- Felix, Sascha (1985): Parasitic Gaps in German. In: W. Abraham, ed., *Erklärende Syntax des Deutschen*. Stauffenburg, Tübingen, pp. 173–200.
- Haegeman, Liliane (1984): 'Parasitic Gaps and Adverbial Clauses', *Journal of Linguistics* 20, 229–

- Halle, Morris and Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. In: K. Hale and S. Keyser, eds, *The View from Building 20: Essays in Honor of Sylvain Bromberger*. MIT Press, Cambridge, Massachusetts, pp. 111–176.
- Heck, Fabian and Gereon Müller (2007): Extremely Local Optimization. In: E. Brainbridge and B. Agbayani, eds, *Proceedings of WECOL 26*. California State University, Fresno, pp. 170–183.
- Heck, Fabian and Malte Zimmermann (2004): DPs as Phases. Ms., Universität Leipzig and HU Berlin.
- Huang, C.T. James (1982): Logical Relations in Chinese and the Theory of Grammar. PhD thesis, MIT.
- Kayne, Richard (1983): ‘Connectedness’, *Linguistic Inquiry* **14**, 223–250.
- Kayne, Richard (1984): *Connectedness and Binary Branching*. Foris, Dordrecht.
- Kearney, Kevin (1983): *Governing Categories*. Storrs. Ms. University of Connecticut.
- Kiss, Katalin (1985): ‘Parasitic Chains’, *The Linguistic Review* **5**, 41–74.
- Koopman, Hilda (2006): Agreement Configurations: In Defense of Spec Head. In: C. Boeckx, ed., *Agreement Systems*. John Benjamins.
- Koster, Jan (1986): *Domains and Dynasties*. Foris, Dordrecht.
- Larson, Richard K. (1988): ‘Extraction and Multiple Selection in PP’, *MIT working papers in linguistics 9: The 25th anniversary of MIT linguistic* pp. 119–135.
- Lee, Kap-Hee (1998): ‘A Minimalist Approach to Parasitic Gap Constructions’, *Studies in Generative Grammar* **8**, 343–363.
- Legate, Julie Anne (2003): ‘Some Interface Properties of the Phase’, *Linguistic Inquiry* **34**, 506–515.
- Longobardi, Giuseppe (1984): Connectedness and Island Constraint. In: J. Guéron, H.-G. Obenauer and J.-Y. Pollock, eds, *Grammatical Representation*. Foris, Dordrecht, pp. 169–185.
- Mahajan, Anoop (1990): The A/A-bar Distinction and Movement Theory. PhD thesis, MIT, Cambridge, Massachusetts.
- Manzini, Rita (1994): ‘Locality, Minimalism and Parasitic Gaps’, *Linguistic Inquiry* **25**, 481–508.
- Müller, Gereon (2010a): ‘On Deriving CED Effects from the PIC’, *Linguistic Inquiry* **41**, 35–82.
- Müller, Gereon (2010b): Operator Islands, Maraudage, and the Intermediate Step Corollary. Ms. Universität Leipzig.
- Müller, Gereon and Wolfgang Sternefeld (1993): ‘Improper Movement and Unambiguous Binding’, *Linguistic Inquiry* **24**, 461–507.
- Munn, Alan (1994): A Minimalist Account of Reconstruction Asymmetries. In: M. Gonzalez, ed., *Proceedings of NELS 24*. pp. 397–410.
- Munn, Alan (2001): Explaining Parasitic Gap Restrictions. In: P. Culicover and P. Postal, eds, *Parasitic Gaps*. MIT Press, Cambridge, Massachusetts, chapter 11, pp. 369–392.
- Nissenbaum, Jonathan W. (2000): Investigations of Covert Phrase Movement. PhD thesis, MIT.
- Nunes, Jairo (1995): The Copy Theory of Movement and Linearization of Chains in the Minimalist Program. PhD thesis, University of Maryland.
- Nunes, Jairo (2001): ‘Sideward Movement’, *Linguistic Inquiry* **32**, 303–344.
- Nunes, Jairo (2004): *Linearization of Chains and Sideward Movement*. MIT Press, Cambridge, Mass.
- Ouhalla, Jamal (2001): Parasitic Gaps and Resumptive Pronouns. In: P. D. Culicover and P. M. Postal, eds, *Parasitic Gaps*. MIT Press, Cambridge, Mass., chapter 5, pp. 147–180.
- Postal, Paul (1993): ‘Parasitic Gaps and the Across-the-Board Phenomenon’, *Linguistic Inquiry* **24**, 735–754.
- Postal, Paul (1994): ‘Parasitic and Pseudoparasitic Gaps’, *Linguistic Inquiry* **25**, 347–364.
- Postal, Paul (1998): *Three Investigations of Extraction*. MIT Press, Cambridge, Massachusetts.
- Reuland, Eric (2001): ‘Primitives of Binding’, *Linguistic Inquiry* **32**(3), 439–492.
- Řezáč, Milan (2004): Agree and Merge. Ms. University of the Basque Country.
- Riedel, Kristina (2009): The Syntax of Object Marking in Sambia: A comparative Bantu perspective. PhD thesis, University Leiden.
- Roberts, Ian (2010): *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Vol. 59 of *Linguistic Inquiry Monographs*, MIT Press, Cambridge, Massachusetts.
- Ross, John R. (1967): Constraints on Variables in Syntax. PhD thesis, MIT.
- Safir, Ken (1987): ‘The Anti-C-Command Condition on Parasitic Gaps’, *Linguistic Inquiry* **18**, 678–

683.

Sternefeld, Wolfgang (2006): *Syntax*. Stauffenburg, Tübingen. Two volumes.

Svenonius, Peter (2004): On the Edge. In: Adger, D. and de Cat, C. and Tsoulas, G., ed., *Peripheries: Syntactic Edges and their Effects*. Kluwer, pp. 261–287.

Taraldsen, Knut (1981): The Theoretical Interpretation of a Class of Marked Extractions. In: A. Belletti, L. Brandi and L. Rizzi, eds, *Theory of Markedness in Generative Grammar*. Scuola Normale Superiore, pp. 475–516.

Webelhuth, Gert (1992): *Principles and Parameters of Syntactic Saturation*. Oxford University Press, Oxford.

Zaenen, Annie, Elisabet Engdahl and Joan M. Maling (1981): 'Resumptive Pronouns can be Syntactically Bound', *Linguistic Inquiry* pp. 679–682.