

The Island Paradox with Parasitic Gaps

A New Perspective on Parasitic Gaps

Anke Assmann¹
 (University of Leipzig)
 phi05eia@studserv.uni-leipzig.de

The Paradox:

Parasitic gaps are at the same time sensitive and insensitive to islands.

Claim:

The paradox can be resolved by implementing the hypotheses that in parasitic gap constructions, one lexical item has to occur simultaneously in two argument positions. In doing so, it has to split into two items. The result of splitting has to be redone during the derivation.

1 The Paradox

1.1 The Island Paradox with Parasitic Gaps

The Paradox:

Parasitic gaps can be separated from their antecedent by *one* barrier but *not by more than one* barrier. (cf. e.g. Kayne (1983); Chomsky (1986); Nunes (1995))

⇒ Parasitic gaps are at the same time sensitive and insensitive to islands.

English:

- (1) *Relative Clause Island + Adjunct Island*
- a. the article [which we filed *t* [without reading *pg*]] *Culicover (2001:27)*
 - b. *the article [which we filed *t* [without meeting the person [who wrote *pg*]]] *Culicover (2001:27)*

¹I am indebted to Brian Agbayani, Petr Biskup, Doreen Georgi, Fabian Heck, Stefan Keine, Gereon Müller, Martin Salzmann, Philipp Weisser, the participants of the Kolloquium 'Neuere Arbeiten zur Grammatiktheorie' of the department of linguistics in Leipzig as well as the audience of the ConSOLE XIX in Groningen 2011 for helpful advice, critique and discussion.

(2) *Adjunct Island + Adjunct Island*

- a. This is the man [Op John interviewed *t* [before meeting *pg*]]
 - b. *This is the man [Op John interviewed *t* [before expecting you to leave [without meeting *pg*]]]
- Chomsky (1986:55)*

German:

(3) *Relative Clause Island + Adjunct Island*

- a. Welchen Artikel hast du [ohne *pg* zu lesen] *t* abgeheftet?
 which article have you without to read filed
- b. *Welchen Artikel hast du [ohne den Autor zu kennen [der *pg* geschrieben hat]]
 which article have you without the author to know who written has
t abgeheftet?
 filed

Question:

Why is there a paradox?

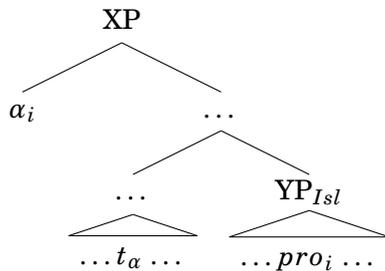
2 Criticism Against Previous Solutions

The Pro-Approach:

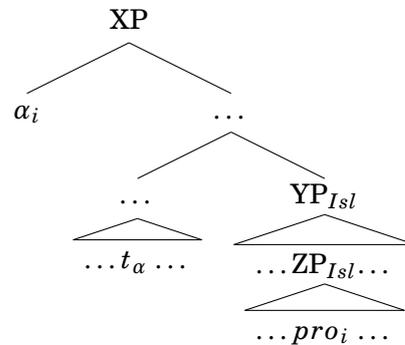
The place of the parasitic gap is occupied by an empty pronominal which is bound by the antecedent. (Chomsky (1982); Engdahl (1985); Cinque (1990); Postal (1993); Ouhalla (2001))

• *Configurations:*

(4) *One island*



(5) *More than one island*



• *Solution:*

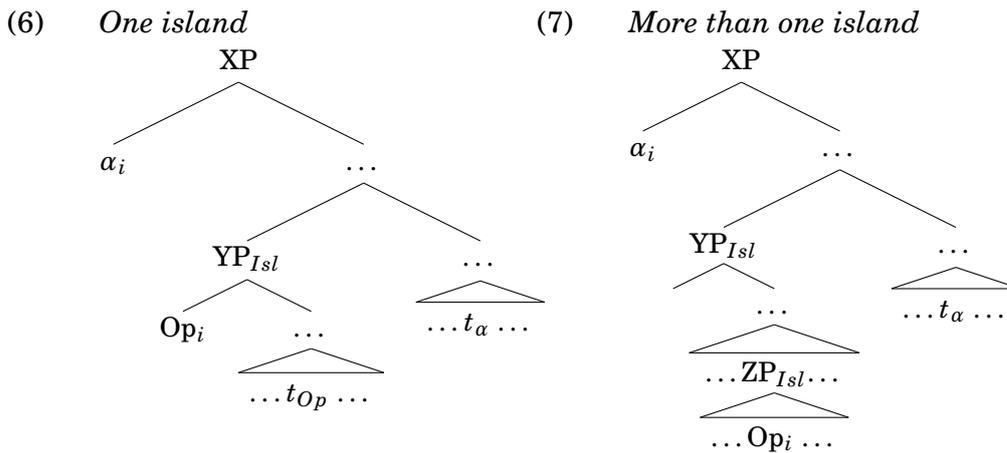
Since the licensing of parasitic gaps is pure binding, *pro*-approaches cannot solve the island paradox without further stipulations.²

²See e.g. Postal (1998) who assumes that the empty pronoun has to move to a higher position in order to be bound.

The Operator Approach:

The place of the parasitic gap is occupied by an empty operator which moves to a higher position inside the island but does not leave the island. In this position, the operator is identified with the antecedent of the real gap. (Chomsky (1986); Lee (1998); Nissenbaum (2000))

• *Configurations:*



• *Solution:*

In structures with only one island, the operator does not have to leave the only island. In structures with more than one island the operator has to move out of the lower islands but not out of the highest island

• *Problems with this account:*

In principle, operator approaches can capture the island paradox since they make use movement. However, there is empirical evidence against PG constructions involving operators.

- No overt operators are allowed in the position of the real operator (whether it is located higher or lower than the clause introducing preposition.)

- (8) a. *Which article₁ did you file t₁ [what₁ without reading t₁]?
 b. *Which article₁ did you file t₁ [without what₁ reading t₁]?

- Possible Solutions and their Flaws:

1. Prepositions like ‘without’, ‘before’ ... are actually complementizers. Overt Operators are not allowed in context of overt complementizers.
Objection: See Dubinsky and Williams (1995) for reasons why, prepositions like ‘without’ are in fact prepositions that select CPs with an empty C head.
2. Prepositions like ‘without’ select only non-interrogative CPs. Empty operators have a distinct feature specification.
Objection: If empty operators have a different formal make-up than their overt counterparts: how does it look like?; which feature makes empty operators ‘operators’?
Note: These objections against empty operators may also apply to other configurations that have been analyzed by means of empty operators (e.g tough movement).

3. Empty operator movement is covert.

Objection: Covert movement is usually not island sensitive (Huang (1982); Chomsky (1995)).

(9) who *t* likes books [that criticize whom]

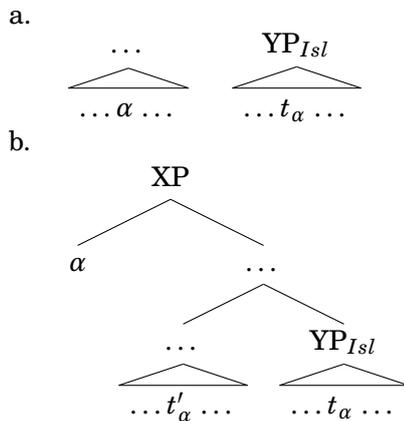
Chomsky (1995, 70)

The Sideward Movement Approach:

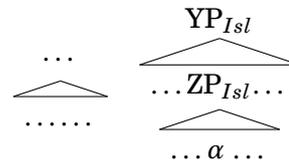
The place of the parasitic gap is occupied by the antecedent. From this position, the antecedent sideward-moves to the position of the real gap before the category becomes an island. (Nunes (2001, 2004))

• *Configurations:*

(10) *One island*



(11) *More than one island*



• *Solution:*

The antecedent is moved from the parasitic to the real gap position before the category is merged with some category of the matrix clause and so before it becomes an island. Only the highest island can be circumvented by sideward movement. If there is more than one island, the antecedent is stuck in its base position.

• *Problems with this account:*

The fact that only the highest island can be circumvented by sideward movement requires the following assumptions:

1. The lexical array is divided into subarrays Chomsky (2000).
2. Prepositions which introduce adverbial clauses are heads which select a CP.

Note: The first assumption is an essential part of the minimalist framework. The second assumption has been doubted since such prepositions are known not to make up a natural class.

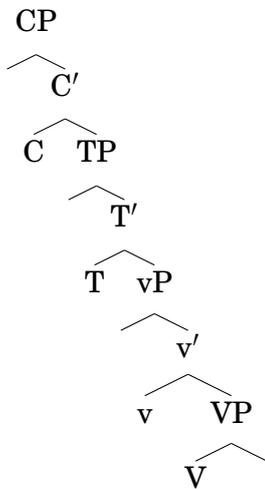
- Temporal prepositions like ‘before’, ‘after’ are actually complementizers. (Dubinsky and Williams (1995))
- Under the assumption that ‘before’, ‘after’ are complementizers, the sideward movement approach can be shown to overgenerate.

(12) *a man Op that John interviews *t* [after allowing you to leave [before meeting *pg*]]

3.1 Assumptions and Background

Minimalist Assumptions (Chomsky (1995, 2000, 2001, 2008)):

(15) *Syntactic Structure*



- Syntax is strictly derivational and obeys the Strict Cycle Condition.
- No look-ahead
- The structure-building operation is Merge triggered by features [$\bullet F \bullet$] and the probe operation is Agree triggered by features [$*F*$].
- Spell-out of syntactic structure is cyclic. Phases are CP, vP and DP.
- Cyclic movement is enabled by edge features [$\bullet X \bullet$] which can be added to phase heads obeying the Edge Feature Condition.

Island Model of Müller (2010):

Categories are islands if they are last-merged in a phase.

- Probe [$*F*$] and structure-building [$\bullet F \bullet$] features on heads are ordered on stacks.
- Merge deletes features [$\bullet F \bullet$] and Agree features [$*F*$].

(16) *Last Resort* (Müller (2010:40))

- Every syntactic operation must discharge either [$\bullet F \bullet$] or [$*F*$].
- Only features on the top of a feature list are accessible.

- Edge features are assigned to phase heads as long as they are active.:

(17) *Edge Feature Condition* (Müller (2010:42))

An edge feature [$\bullet X \bullet$] can be assigned to the head γ of a phase only if (a) and (b) hold:

- γ has not yet discharged all its structure-building or probe features.
- [$\bullet X \bullet$] ends up on top of γ 's list of structure-building features.

- Categories become islands for movement, i.e. no element can be extracted out of them, if they are the last-merged category in a phase.

Example: Adjunct Islands:(18) *Who did Mary cry after John hit *t*?

Huang (1982:503)

(19) a. Feature specification of *v* $v [\bullet V \bullet < \bullet D \bullet]$ b. Merge of VP deletes [$\bullet V \bullet$] $[_{VP} v [\bullet V \bullet < \bullet D \bullet] [_{VP} cry]]$ c. Merge of subject deletes [$\bullet D \bullet$] $[_{VP} Mary [_{VP} v [\bullet D \bullet] [_{VP} cry]]]$ d. Merge of adjunct clause $\rightarrow v$ is already inactive $[_{CP} [_{VP} you [_{VP} v [] [_{VP} cry]]] [_{CP} after John hit who]]$ **Modification of the Island Model of Müller (2010):**

Categories become islands if all operation-triggering features in the *edge domain* of the phase are deleted.

(20) *Edge Feature Condition (Revised)*An edge feature [$\bullet X \bullet$] can be assigned to the head γ of a phase only if (a) and (b) hold:

- there are still operation-triggering features in the *edge domain* of the γ -phase.
- [$\bullet X \bullet$] ends up on top of γ 's list of structure-building features.

3.2 Splitting³**Observation:**

Sentences with parasitic gaps are slightly marked or even impossible for most speakers in English and German (cf. e.g. Legate (2003)).

Note:

If parasitic gaps were indeed the result of binding or simply movement, the markedness or ungrammaticality of parasitic gap sentences cannot be explained.

Assumption:*Working Hypothesis*

Sentences with parasitic gaps are based on defective numerations.

³The following analysis is adapted from Assmann (2010) where this theory has been 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 to be a copying operation. Although the idea and the effects of the two operations are the same, the term 'Splitting' is more adequate for what is going on 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 stipulation has been dismissed since probe features [$*F*$] make exactly the same predictions. Thanks to Fabian Heck and Gereon Müller for pointing that out to me.

Main Idea:

Defective numerations need repair. A mechanism of repair is costly and not available to all speakers. Therefore, parasitic gap constructions are either marked (in case of applying the repair mechanism) or ungrammatical (in case of absence of a repair mechanism).

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

(22) *Numeration of (21):*

N = [<i>which</i>	[D, acc, φ , wh, •N•, *acc*, ..., PHON ⁴ , SEM],
	<i>article</i>	[N, acc, ...],
	C	[C, •wh•, •T•, ...],
	<i>you</i>	[D, φ , nom, ...],
	T	[* φ *, *nom*, T, EPP, •v•, ...],
	v	[•D•, *acc*, * φ *, v, •V•, ...],
	<i>file</i>	[•D•, V, ...],
	Op _{temp} ⁵	[...],
	<i>before</i>	[C, •T•, ...],
	T	[* φ *, *nom*, T, EPP, •v•, ...],
	PRO	[D, φ nom, ...],
	v	[•D•, *acc*, * φ *, v, •V•, ...],
	<i>read</i>	[•D•, V, ...]
]	

- The numeration of (22) has 4 [•D•] / 4 [* φ *] / 2 [*acc*] (v, v, *read*, *file*) features but only 3 matching [D] / 3 [φ] / 1 [acc] features (*which*, *you*, PRO).
- This numeration is about to crash.
- In order to save the numeration, *which* has to split.

(23) *Splitting (N= [L, ...])*

- a. There are structure-building and probe features [•F₁•], ... [•F_i•], [*F_j*], ... [*F_n*] in the numeration N that do not have matching features [F₁], ... [F_n].
- b. There is a lexical item L in N that has features [F₁, ..., F_n, G₁, ..., G_n].
- c. L splits into two items L₁ and L₂, whereby L₂ has the features [F₁, ..., F_n] and L₁ has the features [*F₁*], ..., [*F_n*], G₁, ..., G_n].

(24) a. *Before splitting*

which [D, acc, φ , wh, •N•, *acc*, ..., PHON, SEM]

b. *After splitting*

*which*₁ [*D•, *acc*, * φ *, wh, •N•, *acc*, ..., PHON, SEM]

*which*₂ [D, acc, φ]

Notes:

1. Only *which*₁ has phonological features.
2. Splitting does also effect semantic features. Whatever features are necessary to guarantee a success semantics appear on both elements as well (e.g. referential index).

⁴The features PHON and SEM are to be understood as feature sets. 'PHON' encompasses all phonological features and 'SEM' all semantic features.

⁵Following 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).

3.3 A Sample Derivation

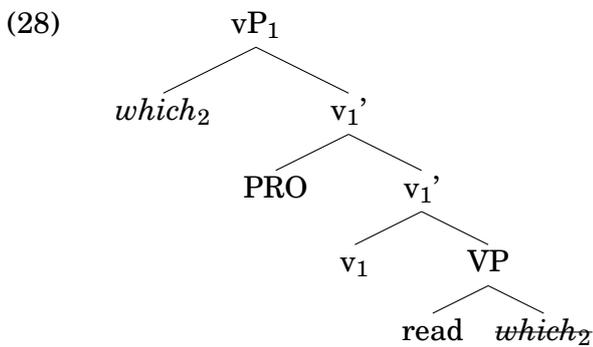
(25) Which article did you file t [before reading pg]?

Embedded vP:

(26) *Subarray vP₁*
 { v_{P_1} read, $which_2$, v_1 , PRO}

(27) *Feature Matrix of v₁*

- a. v_1 [$\bullet V \bullet < \bullet D \bullet < * \varphi * / * acc *$] → (Merge of VP)
- b. v_1 [$\bullet \bar{V} \bullet < \bullet D \bullet < * \varphi * / * acc *$] → (Merge of PRO)
- c. v_1 [$\bullet \bar{D} \bullet < * \varphi * / * acc *$] → (edge feature insertion)
- d. v_1 [$\bullet X \bullet < * \varphi * / * acc *$] → (Movement of $which_2$ to Spec, vP)
- e. v_1 [$\bullet X \bullet < * \varphi * / * acc *$] → (case and φ -Agree)
- f. v_1 [$* \varphi * / * acc *$]

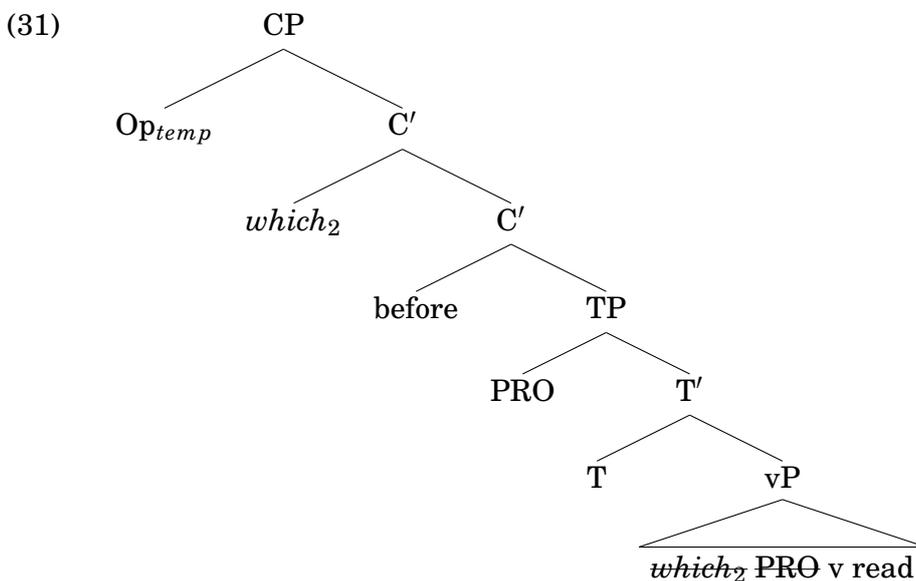


Embedded CP:

(29) *Subarray CP₁*
 { CP_1 T, before, Op}

(30) *Feature Matrix of C*

- a. before [$\bullet T \bullet < \bullet Op \bullet$] → (Merge of TP)
- b. before [$\bullet \bar{T} \bullet < \bullet Op \bullet$] → (edge feature insertion)
- c. before [$\bullet X \bullet < \bullet Op \bullet$] → (Movement of $which_2$ to Spec, CP)
- d. before [$\bullet X \bullet < \bullet Op \bullet$] → (Merge of Op_{temp})
- e. before [$\bullet Op \bullet$]



Matrix vP

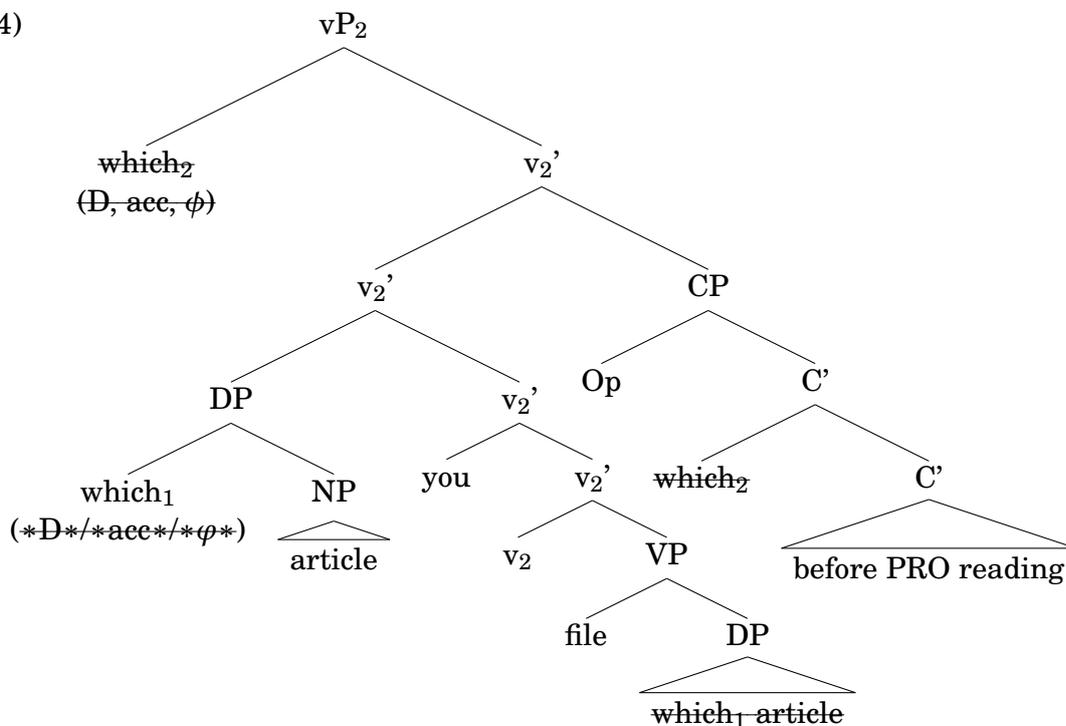
(32) Subarray vP₂

{_{vP₂} file, which₁, v₂, you}

(33) Feature Matrix of v

- a. v₂ [**•V•** < **•D•** < *φ*/*acc*] → (Merge of VP)
- b. v₂ [**•V•** < **•D•** < *φ*/*acc*] → (Merge of Subject)
- c. v₂ [**•D•** < *φ*/*acc*] → (edge feature insertion)
- d. v₂ [**•X•** < *φ*/*acc*] → (Movement of which article to Spec,vP)
 new features in the edge domain of v (*D* < *acc* / *φ*)
- e. v₂ [**•X•** < *φ*/*acc*], which₁ [*D*/*acc*/*φ*] → (case and φ-Agree)
- f. v₂ [*φ*/*acc*], which₁ [*D*/*acc*/*φ*] → (Merge of Adjunct Clause)
- g. v₂ [], which₁ [*D*/*acc*/*φ*] → (edge feature insertion)
- h. v₂ [**•X•**], which₁ [*D*/*acc*/*φ*] → (movement of which₂ to Spec,vP)
- i. v₂ [**•X•**], which₁ [*D*/*acc*/*φ*] → (Agree between which₁ and which₂)
- j. v₂ [], which [*D*/*acc*/*φ*]

(34)

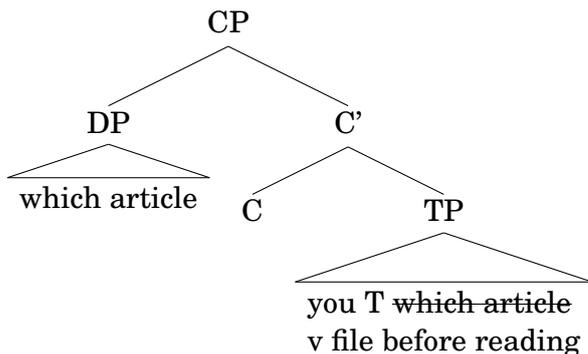


Matrix CP

(35) Subarray CP₂

{_{CP₂} T, C}

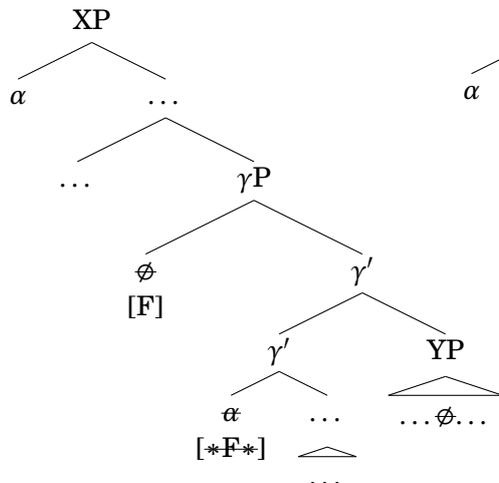
(36)



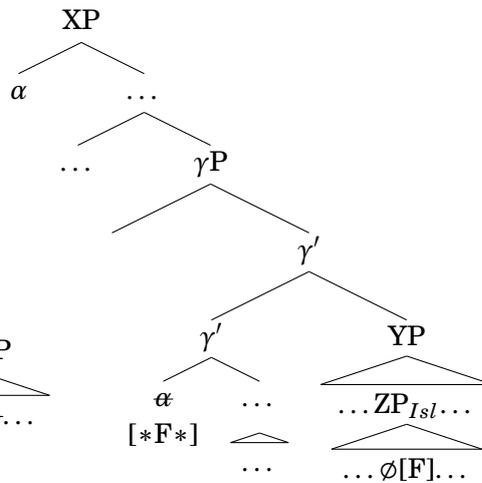
4 Paradox resolved

- Configurations:

(37) *One island*



(38) *More than one island*



- Solution:

Agree features are able to save only the highest adjunct clause from becoming an island by keeping the matrix v active. All last-merged categories inside the highest adjunct clause cannot be circumvented in this way.

5 Further Evidence – Three Stages of the Split Lexical Item

The splitting analysis predicts that a lexical item has three different stages due to three different feature specifications.

- (39) L [F₁ ... F_n, G₁ ... G_n]
 a. L₁ [*F₁* ... *F_n*, G₁ ... G_n]
 b. L₂ [F₁ ... F_n]
 c. L'₁ [G₁ ... G_n]

Example:

- (40) Which article did you file *t* before reading *pg*?
 (41) which [D, ϕ , acc, wh, •N•, SEM, PHON, ...]
 a. which₁ [*D*, * ϕ *, *acc*, wh, •N•, SEM, PHON, ...]
 b. which₂ [D, ϕ , acc]
 c. which'₁ [wh, •N•, SEM, PHON, ...]

(42)

Stage	take NP complement	occupy A-position	wh-properties
1. <i>which</i> ₁	+	+	+
2. <i>which</i> ₂	-	+	-
3. <i>which</i> ' ₁	-	-	+

Taking NP complements:

Only *which*₁ but not *which*₂ can take an NP complement.

Evidence from binding:

- (43) 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 (43-a), ‘books about himself’ is merged together with *which*₁ in the matrix clause. In this position, ‘himself’ can be bound by ‘John’.
- In (43-b) ‘books about herself’ is again merged with *which*₁ in the matrix clause. In this position, it cannot be bound by the possible antecedent ‘Mary’.

Occupying A-positions:

After Agree with *which*₂, *which*₁ loses its categorial, ϕ and case features. Therefore, it is incompatible with A-positions.

Evidence from the subject position in English:

- (44) *Derived subjects cannot license parasitic gaps*
 *Which house was sold *t* [before we could demolish *pg*]? *cf. Legate (2003, 511)*
- (45) *Base-generated subjects cannot license parasitic gaps*
 a. *Which spy *t* killed John before anybody could speak to *pg*?
 b. Which spy did John kill *t* before anybody could speak to *pg*?
 Safir (1987, 678)

- The English subject position Spec,TP requires an argument that has categorial, ϕ and case features.
- Arguments that have licensed a parasitic gap do not have such features anymore. (They have been checked via Agree.)

Sensitivity for *wh*-islands:

Only *which*₁ but not *which*₂ has a feature [wh].
Assumption: Only *wh*-elements are sensitive for *wh*-islands.

Evidence from German:

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

- In (46-a), *which*₁ has to cross a wh-island. This is impossible since it has a feature [wh].
- In (46-b) *which*₂ has to cross a wh-island. This is possible since it has no feature [wh].

6 Remarks

Applicability of Splitting:

- The splitting mechanism is a mechanism that is specifically developed for parasitic gap constructions
- The power of the splitting mechanism is not explored in full depth yet
- It might be that splitting can also be used to explain other phenomena (cf. Agbayani (1998); Agbayani and Ochi (2007)).

*Computational Complexity*⁶:

- Splitting involves a scan of the numeration.
- This scan, however, does not increase the running time of the derivation.
- It suffices to count operation-triggering and non-operation-triggering features. This process has a running time of $O(2n)$ (n = number of features), which is linear. The running time of the derivation is at least linear. (Every feature is checked by / checks another feature.)

Empirical Adequacy:

- Even though the splitting account provides a fully-fledged solution to the island paradox, there are some puzzling properties which may be problematic, e.g.:

- multiple parasitic gaps (Ross (1967); Engdahl (1983))

(47) The contract *which*₁ I want to peruse *t*₁ before damaging *pg*₁ while filing *pg*₁' is written on Peruvian papyrus Ross (1967:105).

- reversed wh-island sensitivity in English (Agbayani and Ochi (2007))

(48) *Who did John interview *t* before expecting us to ask you which job to give to *pg*? Agbayani and Ochi (2007:7).

- cases where parasitic gaps can be licensed by covert movement (Nissenbaum (2000))

(49) a. ?[Which senator]₁ did you persuade *t*₁ to borrow [which car]₂ [after getting [an opponent of *pg*₁] to put a bomb in *pg*₂]?

b. *[Which senator]₁ did you persuade *t*₁ to borrow [which car]₂ [after putting a bomb in *pg*₂]? Nissenbaum (2000:98)

⁶This question was pointed out to me by an anonymous reviewer.

References

- 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.
- Assmann, Anke (2010), Parasitic gaps in derivational grammar, Master's thesis, University of Leipzig.
- Chomsky, Noam (1982), *Some concepts and consequences of the theory of government and binding*, MIT Press, Cambridge, Massachusetts.
- Chomsky, Noam (1986), *Barriers*, MIT Press, Cambridge, Mass.
- 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.
- 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.
- Huang, C.T. James (1982), Logical relations in Chinese and the theory of grammar., PhD thesis, MIT.
- Kayne, Richard S. (1983), 'Connectedness', *Linguistic Inquiry* **14**, 223–250.
- Kearney, Kevin (1983), *Governing Categories*, University of Connecticut, Storrs.
- 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.
- Müller, Gereon (2010), 'On deriving ced effects from the pic', *Linguistic Inquiry* **41**, 35–82.
- 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 (1998), *Three investigations of extraction*, 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.