Opaque Interactions of Merge and Agree: On the Nature and Order of Elementary Operations

Dissertation

zur Erlangung des akademischen Grades
doctor philosophiae
(Dr. Phil.)

Eingereicht an der
Philologischen Fakultät der
Universität Leipzig
von

Doreen Georgi

Datum der Einreichung: 03.03.2014

Datum der Verleihung: 05.05.2014

Gutachter:

Prof. Dr. Gereon Müller
Institut für Linguistik, Universität Leipzig

Prof. Dr. Andrew Nevins
Linguistics Department, University College London

Abstract and outline

This thesis explores transparent and opaque interactions of operations in Minimalist syntax in order to gain insights into the nature of elementary operations and their mode of application. In particular, I investigate configurations in which a head H triggers more than one operation, viz. Merge and Agree. Assume the context for both operations is given after H is merged into the structure. A central question since the beginnings of transformational grammar has been whether rules apply simultaneously or sequentially in such a situation. The Earliness Principle seems to call for simultaneous application because this is the only way to apply both rules as soon as their context is met. However, simultaneous application can only produce opacity effects but no transparent interactions. Importantly, a number of researchers have shown in recent years that there are feeding and bleeding relations between Merge and Agree triggered by the same head. These can easily be derived if these operations are ordered relative to each other. The choice is taken to be language-specific, i.e. there is a parameter that can be set either to *Merge applies before Agree* (Merge > Agree) or to *Agree applies before Merge* (Agree > Merge).

The central claim of this thesis is that different types of Merge must be distinguished. Simply ordering Merge relative to Agree is not sufficient. There is empirical evidence suggesting that Merge operations, when triggered by the same head H, apply at different points of the derivation *within a single language*. We can see this once Merge is interleaved with Agree: Some Merge operations apply *before* Agree (feeding or bleeding it), while others apply *after* Agree (counter-feeding or counter-bleeding it):

(1)   Merge > Agree > Merge

But this is paradoxical if Merge is a uniform operation. Given this assumption, we cannot produce a symmetric order as in (1); either Merge applies before Agree or afterwards, but not both at the same time. This paradox can be resolved if the Merge operation that applies before Agree is actually of a different type than the one applying after Agree. Hence, the grammar does not simply order Agree relative to Merge, but rather to *subtypes of Merge*. In this way, distinct subtypes of Merge can apply at different points in the derivation. In (2), Merge type 1 applies before Agree and Merge type 2 applies after Agree.

(2) $Merge_1$ > Agree > $Merge_2$

I implement this order of operations by ordering the operation-inducing features on the head H that triggers them. The conclusion is that $Merge_1$ and $Merge_2$ must have different triggers in order for them to be able to get discharged at different points relative to the probe feature that triggers Agree. By looking at configurations as in (2) where Agree "splits up" Merge operations, we can investigate which subtypes of Merge need to be distinguished. Indeed, the split between types of Merge is not random; we find the same splits across languages over and over again.

A systematic split I want to investigate is the one between final and intermediate movement steps in a movement chain (where movement is internal Merge): They must have different triggers in order to allow for different ordering with respect to Agree. It has been assumed before that intermediate and final movement steps require different triggers. But the arguments for the split are either semantic and to some extent conceptual in nature, or they concern the PF-component. As for the first, feature-checking in final position goes hand in hand with semantic effects (i.e. scope taking) that are absent in intermediate positions; nevertheless, the idea has been pursued that all movement steps must be feature-driven, calling for a different type of trigger for non-final movement steps. As for the second, it is usually only the occurrence in the final position that is spelled out; this fact can be tied to the nature of the trigger for final vs. intermediate movement steps.

I will present morphosyntactic evidence for this hypothesis based on the observation that final and intermediate movement steps can have different consequences for the syntactic operation Agree. In some languages, we get the splitting configuration in (2): A single head H triggers both types of internal Merge and Agree, and one type of movement feeds Agree whereas the other counter-feeds it. Empirical evidence comes form cross-linguistic patterns of successive-cyclic movement. The literature on such reflexes is almost exclusively concerned with the pattern where the reflex occurs in every clause along the path of movement, thereby providing evidence for intermediate stop-overs of the moving element. However, there are more patterns, summarized in (3):

(3) *Reflexes of cross-clausal successive-cyclic movement*:
   a. Pattern I:
       The reflex occurs in every clause along the path of movement.
   b. Pattern II:
       The reflex occurs only in the clause in which the moved XP surfaces, but neither in clauses above nor below that clause.
   c. Pattern III:
       The reflex occurs only in clauses that are crossed by the moving XP, but not in the clause where the XP surfaces.

Let me provide an example for each pattern. Irish (Celtic) exhibits a pattern I reflex for complementizer selection. As shown in (4), the default form of the complementizer, *go*, changes to a form glossed *aL* in all clauses between the base position and the landing site of an Ā-moved XP (the base position is indicated by underline) .

(4) *Complementizer selection in Irish (McCloskey 2002: 3)*:
   a. Dúirt mé **gu**-r shíl   mé **go** meadh   sé ann
       said I *go*-PST thought I *go* would.be he there
       'I said that I thought that he would be there.'   *no extraction*
   b. an tainm **a** hinndeadh dúinn **a** bhí ___ ar an áit
       the name *aL* was.told   to.us *aL* was   on the place
       'the name that we were told was on the place'   *long relativization*

In Duala (Niger-Congo), we find a pattern II reflex: If an XP undergoes Ā-movement, the marker *no* obligatorily occurs after the first verbal element of the clause, compare (5-a) without overt extraction and (5-b) with wh-movement. Under long extraction, this marker can only appear in the in the clause where the wh-phrase lands, but not in the clause where the wh-phrase originates.

(5)   *No-marking in Duala (Epée 1976: 194-195)*:

    a.   Kuo a   po   njika ponda
        Kuo 3SG come which time
        'At what time will Kuo arrive?'   *wh-in-situ*

    b.   njika ponda Kuo a   po   **no** ___
        which time   Kuo 3SG come *no*
        'At what time will Kuo arrive?'   *short wh-movement*

    c.   njika buna o   ta **no** o   kwalane mba na o   mende timba ___
        which day   you PST *no* you tell   me that you FUT   return
        'When$_i$ did you tell me that you would return t$_i$?'

                                                long wh-movement

In Kitharaka (Niger-Congo), the preverbal focus marker *n-* has a pattern III distribution: It occurs only in those clauses along the path of movement where the moved XP does not surface, cf. (6-a). Therefore, we do not find it with clause-bound extraction as in (6-b).

(6)   *Preverbal focus marker in Kitharaka (Muriungi 2005: 67-69)*:

    a.   N-uu   u-ku-thugania ati John **n**-a-ug-ir-e   Lucy
        FOC-who 2SG-PRES-think that John FOC-SM-say-PERF-FV Lucy
        n-a-ring-ir-e   ___
        N-SM-beat-PERF-FV
        'Who do you think that John said Lucy beat?'   *long wh-movement*

    b.   I-mbi   Maria a-k-ir-e   ___
        FOC-what Maria SM-build-PERF-FV
        'What did Maria build?'   *short wh-movement*

(7) provides a more abstract representation of the patterns. Assume we find a reflex R on the C head, triggered by wh-movement of an XP from an embedded clause over two clause-boundaries to SpecCP of the topmost clause:

(7)   *Abstract patterns*:

    a.   PI: [$_{CP1}$ XP$_{wh}$ [$_{C'}$ **C-R** ... [$_{CP2}$ **C-R** ... [$_{CP3}$ **C-R** ... t$_{XP}$]]]]
    b.   PII: [$_{CP1}$ XP$_{wh}$ [$_{C'}$ **C-R** ... [$_{CP2}$ C ... [$_{CP3}$ C ... t$_{XP}$]]]]
    c.   PIII: [$_{CP1}$ XP$_{wh}$ [$_{C'}$ C ... [$_{CP2}$ **C-R** ... [$_{CP3}$ **C-R** ... t$_{XP}$]]]]

Given the standard assumption that the XP makes a stop-over in every SpecCP position on the way to its surface position, we see that the reflexes show a split between final and intermediate movement steps: They occur either on the C head that projects the final landing site of XP (pattern II), or only on C heads that project an intermediate landing site for XP (pattern III); or there is a reflex on C regardless of whether SpecCP is the final or an intermediate landing site (pattern I). I take the reflex to be the morphological realization of an Agree relation between a head H (here, C) and the XP moving through SpecHP. Furthermore, I assume that Agree applies upwards, in a Spec-Head-configuration. Thus, movement to SpecHP can feed Agree with H. In pattern I languages, all movement steps feed Agree. In pattern II and pattern III languages, we get opacity: In the former, only final movement steps feed Agree whereas intermediate steps counter-feed Agree; in the latter, the reverse holds, i.e. only intermediate movement steps feed Agree. This can be modeled if the features that trigger final and intermediate movement steps are distinct, such that they can apply at different points relative to Agree. The features are [•F•] for final and [•EF•] for intermediate steps, respectively; [∗F∗] triggers Agree. The transparent pattern I

arises if both types of Merge are triggered before Agree, cf. (8-a). Pattern II follows if final steps are triggered before Agree whereas non-final steps apply after Agree, cf. (8). The reverse order in (8-c) derives pattern III. In general, the movement step that applies before Agree feeds Agree, and a movement step that applies after Agree counter-feeds it.

(8)   a.   Pattern I: [•F•], [•EF•] > [∗F∗]   *all movement steps feed Agree*
      b.   Pattern II: [•F•] > [∗F∗] > [•EF•]   *only final steps feed Agree*
      c.   Pattern III: [•EF•] > [∗F∗] > [•F•]   *only intermediate steps feed Agree*

Thus, variation is accounted for by reordering of operation-inducing features. Previous analyses of the phenomenon usually treat only one of the patterns; a uniform account of all patterns is missing so far. The present thesis aims at closing this gap.

   I extend the approach by showing that Merge needs to be split into more subtypes than just final and intermediate internal Merge. There is also evidence for a distinction between external and internal Merge, as well as between different types of final and intermediate internal Merge (e.g. wh-movement vs. relativization). The logic of the argument is always the same: The Merge subtypes need to be distinguished because they have different consequences for Agree, and therefore have to apply at different points in the derivation relative to Agree.

   I will show that a split pattern of Merge and Agree as in (2) can be found on various heads in the clausal spine such as n, v, T, and C. Moreover, I follow Baker (2008) in assuming that there is a parameter on the direction of Agree (upward or downward) that can be independently determined by the behavior of Agree in the absence of any movement. I show that there are opaque interactions of Merge and Agree regardless of the direction of Agree; upward Agree as in (7) leads to feeding / counter-feeding; downward Agree can lead to bleeding / counter-bleeding or feeding / counter-feeding depending on whether the XP moved to SpecHP is the potential controller of Agree itself or not.

   Taken together, this thesis argues for a more fine-grained typology of Merge, for the need for extrinsic ordering of operation-inducing features, and – given that the approach crucially relies on timing – it provides strong evidence for a strictly derivational model of grammar.

This thesis is structured as follows: Chapter 1 provides an introduction to transparent and opaque rule interactions in syntax. Furthermore, it gives a brief overview over the treatment of these interactions in different stages of transformational grammar, with an emphasis on the challenges the phenomena pose for derivational and representational models of grammar, as well as the insights we gain from these interactions for the mode of rule application.

Chapter 2 motivates and introduces the ordering approach to opaque interactions of Merge and Agree. The chapter contains the set of data that provides evidence for a split between final and intermediate movement steps in an upward Agree context. First, I introduce the notion of successive-cyclicity. I then present the abstract patterns of morphological reflexes of movement and exemplify them with data from a number of languages. Afterwards, I show that my analysis with ordering of operations is superior to previous approaches for both empirical and conceptual reasons. The central piece of argument comes from languages which simultaneously display different patterns of reflexes, and optionality between patterns. Finally, I provide detailed analyses for the languages discussed in this chapter.

In chapter 3, I show that the same split between final and intermediate movement steps can also be found in languages where the Agree relation applies downwards. I present three seemingly unrelated phenomena (removal of defective intervention effects, movement-related case splits, the Anti-agreement effect) and argue that they all involve opaque interactions of Merge and Agree, though on different heads. Furthermore, I demonstrate that the (re)ordering approach provides a simple explanation for the cross-linguistic variation found with these phenomena, while it remains unaccounted for in previous approaches.

Chapter 4 extends the analysis to other subtypes of Merge: Opacity effects in the interaction of Merge and Agree also motivate a split between external and internal Merge as well as between different types of final and intermediate internal Merge.

In chapter 5 I explore the general consequences of the analysis. I first show that the present analysis provides an argument for the need of extrinsic ordering in addition to intrinsic ordering. I then address the central claim that final and intermediate movement steps must have different causes and illustrate how this requirement can be implemented in approaches with rather different characteristics than the Attract approach pursued here (Greed-based and non-feature-driven approaches to movement). Moreover, I argue that the present analysis has consequences for the timing of edge-feature discharge: Contrary to claims in the literature, the attested orderings show that edge features cannot generally be the first or the last feature a head discharges. Furthermore, I evaluate possible representational reconstructions of the ordering approach and show that a derivational system is much more simple and elegant. Finally, I recapitulate the parameters the analysis requires to derive the cross-linguistic variation.

Chapter 6 summarizes the main claims of the thesis and closes with an outlook on future research.

Contents

Acknowledgements

Writing this thesis was both a challenge and a great pleasure. I feel fortunate to have had the opportunity to meet so many interesting people during this time to whom I wish to express my gratitude.

First and foremost, I want to thank Gereon Müller for his support from the very beginning. His passion for linguistics has been infectious and I have benefited enormously from his vast knowledge and challenging questions. His ability to look at linguistic problems from a completely different perspective has been an invaluable inspiration to me. I thank him for his belief in me and for pushing me to make the steps towards a life in linguistics. Without his constant motivating pressure from early on I probably would not have dared to even think about writing a thesis.

A very special thanks to Andrew Nevins for agreeing to review this thesis – your efforts are much appreciated! Thank you for your interest in my work and for highly inspiring papers and talks (in Leipzig and elsewhere) on complex agreement patterns, which cannot but make you want to become a linguist.

I am indebted to Fabian Heck, Gereon Müller and Jochen Trommer for teaching me what I know about linguistics and for giving me the opportunity to do research in a highly creative and encouraging atmosphere. Their doors have always been open and I have profited enormously from their knowledge and advice.

My fellow PhD students Anke Assmann, Sebastian Bank, Daniela Henze, Timo Klein, Philipp Weisser, and Eva Zimmermann – that the last years in Leipzig have been such a great time for me is mainly due to you! Thank you so much for an intellectually and socially stimulating environment. It was great fun to travel to conferences and give talks with you!

In particular, thanks to Eva and Daniela for our "Projektbesprechungen" and relaxing board game nights; to Timo and Philipp (as well as Martin and Fabian) for teaching me foosball tricks and thereby also swearing in various German dialects; to Timo for reintroducing me to "Die 2"; and especially to Anke and Philipp, with whom I have been collaborating most. It has been very inspiring to discuss and develop analyses with you. I have learned a lot from our constructive discussions about the best analysis and I hope that we will continue to do joint work in the future. And to Katja, my future PhD fellow, for her refreshing pleasure in learning about syntax over a cup of coffee.

I also want to thank Petr Biskup, Antje Lahne, Stefan Keine, Andreas Opitz, Marc Richards, Barbara Stiebels, Sandhya Sundaresan, and Nina Topintzi for valuable input and discussions in our colloquium, during lunch or when having an after work beer.

I would also like to express my thanks to the people who answered all my data questions: Adrienn Jánosi, Peter Muriungi, Irina Nikolaeva, Milan Řezáč, Halldór Sigurðsson, Höskuldur Thráinsson, Harold Torrence, Coppe van Urk, and in particular Jason Zentz for sharing his knowledge about reflexes of movement in Bantu and for his comments on my work. In addition, I want to thank the audiences at the conferences where I had the

Meinen Großvätern
in liebevoller Erinnerung

Abbreviations and glosses

Abbreviations

AAE	Anti-agreement effect
ADJ	adjunct
ARG	argument
CNP	complex noun phrase
DM	Distributed Morphology
DO	direct object
DP_{ext}	external argument
DP_{int}	internal argument
DS	deep structure
GB	Government and Binding theory
IO	indirect object
IS	intermediate structure
ISD	imperative subject deletion
LF	logical form
MP	Minimalist Program
OBJ	object
OT	Optimality Theory
PI, PII, ... PV	pattern I, II, ... V
PIC	Phase Impenetrability Condition
PF	phonological form
PSR	phrase structure rules
SCC	Strict Cycle Condition
SCO	strong cross-over
SF	Stylistic Fronting
Spec	specifier
SS	surface structure
SU	subject
US	underlying structure
WCO	weak cross-over
VI	vocabulary item

Subscripts

k, n	movement index
i, j	referential index

Features

[F]	feature
[F:V]	valued feature
[•F•]	structure-building feature
[∗F:□∗]	unvalued probe feature
[∗F:V∗]	valued probe feature
~~[•F•]~~ / ~~[∗F:V∗]~~	discharged structure-building / probe feature
~~[F]~~	feature deleted by an impoverishment rule
#	number feature
π	person feature
φ	phi-features

Glosses

1	1st person		INF	infinitive
1excl	1st person exclusive		IRR	irrealis
1incl	1st person inclusive		L	linker
2	2nd person		LOC	locative
3	3rd person		NEG	negation
ABS	absolutive		NOM	nominative
ACC	accusative		NS	non-subject
AGR	agreement marker		OBJ	objective
BEN	benefactive		OBL	oblique
C	complementizer		PART	particle
CASE	case marker		PARTCP	participle
CL	class marker		PASS	passive
CLT	clitic		PERF	perfective
COMPL	completive		PL	plural
COND	conditional		POSS	possessive
DAT	dative		PREP	preposition
DEF	definite		PRES	present
DEM	demonstrative		PROG	progressive
DTR	ditransitive		PROX	proximate
DU	dual		PST	past
ERG	ergative		Q	interrogative marker
EXPL	expletive		R	realis
FEM	feminine		REFL	reflexive
FIN	finite		REL	relative agreement
FOC	focus		RM.PST	remote past
FRC	force marker		SCO	scope marker
FUT	future		SG	singular
FV	final vowel		SM	subject marker
GEN	genitive		SUBJ	subjunctive
IMPERF	imperfective		TOP	topic
INCOMPL	incompletive		TRANS	transitive
INDEF	indefinite		WH	wh-agreement

List of Figures

List of Tables

Chapter 1

On rule interactions

In this chapter, I set the stage for the discussion of opacity in this thesis. In section 1.1 I introduce the concept of opacity and related terminology. Section 1.2 gives a brief overview over the history of rule interaction in the generative paradigm of syntax from transformational grammar to Minimalism. The focus is on (i) the building blocks that can interact in the respective frameworks, (ii) the mode of rule application that the interactions require (sequential vs. simultaneous), (iii) the predictability of rule orderings (extrinsic vs. intrinsic), (iv) the modeling of different types of interactions in derivational and representational models of grammar. I will come back to these points in the following chapters when I introduce my own analysis and compare it to previous approaches of the phenomena under investigation. In particular, I will show in chapter 5 what new insights we can gain from the interaction of operations in Minimalism for the sequential vs. simultaneous and the extrinsic vs. intrinsic discussion.

1.1   Types of rule interactions

Traditionally, there are basically two components in a grammar: the base component and the transformational component. The former contains rules that generate the hierarchical structure of a linguistic expression, the underlying structure (US). The latter contains rules (R) that relate the underlying structure to the surface structure (SS) of a linguistic expression (cf. Chomsky 1965 on syntax and Chomsky and Halle 1968 on phonology). This is abstractly depicted in figure 1.1:

$$\boxed{\text{US}} \xrightarrow{R} \boxed{\text{SS}}$$
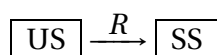
Figure 1.1: Transformational component

Rules apply to a structure if their structural description (SD), i.e., their context of application, is met. Often, more than one rule applies in the transformational component to derive the surface structure of a linguistic expression from its underlying structure. In this case, the rules may interact in intricate ways. For example, the application of a rule R_1 may block or facilitate the application of another rule R_2. Kiparsky (1965; 1968; 1971; 1973*a*) classifies rule interactions into *transparent* and *opaque* interactions and labels instances of these interaction types: Transparent interactions include *feeding* and *bleeding*. Opaque interactions comprise *counter-feeding* and *counter-bleeding*; McCarthy (1999) uses the terms *underapplication* and *overapplication* to characterize counter-feeding and

counter-bleeding, respectively.[1] These terms are defined in (1) and (2). Assume there are two rules R_1, R_2 and a structure S to which these rules apply:

(1)   *Transparent rule interactions*:
  a. Feeding:
    (i)   A rule R_1 creates the context for the application of a rule R_2.
    (ii)   If R_1 applies before R_2, R_1 feeds R_2.
    (iii) If R_1 does not apply to S, R_2 cannot apply because its SD has not been created by R_1 (assuming that the SD of R_2 is not met in S in the first place).
  b. Bleeding:
    (i)   A rule R_1 destroys the context for the application of a rule R_2.
    (ii)   If R_1 applies before R_2, R_1 bleeds R_2.
    (iii) If R_1 does not apply to S, R_2 can apply because its context in S has not been destroyed by R_1.

(2)   *Opaque rule interactions*:
  a. Counter-feeding (underapplication):
    (i)   A rule R_1 creates the context for the application of a rule R_2.
    (ii)   If R_1 applies before R_2, R_1 should feed the application of R_2.
    (iii) However, empirical evidence shows that R_2 has not applied although R_1 has, i.e., R_2 underapplies.
    (iv)   On the surface: A rule (R_2) has not applied although its context is given.
  b. Counter-bleeding (overapplication):
    (i)   A rule R_1 destroys the context for the application of a rule R_2.
    (ii)   If R_1 applies before R_2, R_1 should bleed the application of R_2.
    (iii) However, empirical evidence shows that R_2 has applied although R_1 has as well, i.e., R_2 overapplies.
    (iv)   On the surface: The rule R_2 has applied although its context is not given.

Counter-feeding and counter-bleeding are opaque because it cannot be read off of the surface structure of a linguistic expression (a) why rule R_2 has not applied although its context is given (counter-feeding), or why rule R_2 has applied although its context is not given (counter-bleeding).[2]

Opacity was first described in Chomsky (1951) (see also Chomsky 1975: 25-26): In Hebrew phonology, the two rules *Vowel Reduction* and *Spirantization* (in which a stop is weakened to a fricative) are in a counter-bleeding relation. Since *Spirantization* only applies after vowels, *Vowel Reduction* could potentially bleed *Spirantization* if it applied early and deleted a vowel that precedes a stop. However, on the surface there are instances of stops that have undergone *Spirantization* although the stop in question immediately

---

[1]Baković (2011: 65, n.1) attributes the prefix 'counter-' in counter-feeding and counter-bleeding to Newton (1971). On the origin of the terms *underapplication* and *overapplication* in the phonological literature see Baković (2011: 43).

[2]Kiparsky's original definition of opacity (in phonology) reads as follows (Kiparsky 1973*a*: 79):

(i)   A process P of the form $A \rightarrow B$ / C __ D is defined as being opaque to the extent that there are phonetic forms in the language having either
  a. A in environment C __ D, or
  b. B derived by the process P in environment other than C __ D
  c. B not derived by the process P (i.e. underlyingly or derived by another process) in environment C __ D.

follows a consonant instead of a vowel. Hence, on the surface it is not clear why *Spirantization* could apply given that its structural description is not met. If, however, the rules are ordered such that Spirantization applies before Vowel Reduction, the output can be produced: Spirantization could apply because the stop was preceded by a vowel at that point; the vowel is deleted only afterwards.

In what follows, I provide abstract examples of transparent and opaque interactions. Suppose there is a language L with a grammar G; G contains the following three rules, with no fix order between them intended (A, B and C are abstract symbols; '→' indicates that the symbol to the left of the arrow is replaced by the symbol to the right of the arrow in the context x, represented by '/x__'):

(3)   *Some rules of G in L:*
    a. R_1: B → C / A __
    b. R_2: A → D
    c. R_3: C → D

Assume that the rules apply sequentially and that each rule can apply only once. Given the input AB, the following transparent and opaque rule interactions can occur (among others):

(a) Feeding:
    If R_1 applies before R_3, R_1 feeds R_3 because the former creates the context for the latter by introducing the symbol C that R_3 manipulates. R_3 cannot apply to the input AB, but to the intermediate representation AC created by R_1.

    (4) $\text{AB} \xrightarrow{R_1} \text{AC} \xrightarrow{R_3} \text{AD}$

(b) Bleeding:
    Assume that R_2 applies before R_1. The SD of both R_1 and R_2 is met in the input. However, the earlier application of R_2 destroys the context for the application of R_1 because the former manipulates a symbol in the input that is part of the SD of the latter. The asterisk in the representation in (5) means that R_1 cannot be applied to the intermediate representation created by the application of R_2.

    (5) $\text{AB} \xrightarrow{R_2} \text{DB} \xrightarrow{*R_1}$

(c) Counter-feeding:
    If the output AC is derived from the input AB we have an instance of opacity. Looking at the output, it is unclear why rule R_3 has not applied to AC although its SD is given. This can be derived if R_3 applies before R_1, the reverse of the feeding order. First, R_3 tries to apply to the input AB but since its context is not met it does not evoke any changes. Afterwards, R_1 applies and replaces B by C. Now the input for R_3 is created, but it is too late for R_3 to apply. Thus R_1 counter-feeds R_3.

    (6) $\text{AB} \xrightarrow{(\,R_3\,)} \text{AB} \xrightarrow{R_1} \text{AC}$

   (d) Counter-bleeding:

When looking at the output DC derived from the input AB, it is unclear why rule R_1 could apply because its context is not given: C follows D, but not A. This is an instance of counter-bleeding. The pattern follows if R_1 applies before R_2, the reverse of the bleeding order. First, R_1 applies and turns A into C. A is changed to D only afterwards by R_2, obscuring the context for the application of R_1.

(7) $AB \xrightarrow{R_1} AC \xrightarrow{R_2} DC$

Opacity has also been described in terms of *surface truth,* most explicitly by McCarthy (1999: 332): "The generalization expressed by a phonological rule is not surface true if there are surface counterexamples to that generalization".[3] This view of opacity as non-surface truth emerges if rules are reinterpreted as generalizations over the output (filters). This can be demonstrated on the basis of the abstract examples above. The output generalization corresponding to (3-c) is that there can be no occurrences of C on the surface (because the rule changes every C into D). This generalization is not surface-true in (6). Likewise, the output generalization corresponding to (3-a) is that the symbol C can only occur in the output when adjacent to A (unless C is adjacent to another symbol already in the input, which is not the case in the abstract examples). In the counter-bleeding case in (7), this generalizations is not surface-true: There is an instance of C preceded by a symbol other than D.

Baković (2011) provides a detailed overview of opacity phenomena in phonology. He shows that counter-feeding and counter-bleeding are not the only instances of over-application and underapplication. Indeed, there are much more instances such as mutual bleeding, self-destructive feeding, Duke-of-York derivations, to name just a few phenomena he discusses. In addition, he argues that (a) there are cases of opacity that cannot simply be accounted for by rule ordering (in contrast to counter-feeding and counter-bleeding, see the abstract derivations above), and (b) that not all instances of counter-feeding and counter-bleeding result in opacity. A prominent example of underapplication that is not the outcome of a counter-feeding rule order is disjunctive blocking (cf. Baković 2011: 48): In this case, several rules compete for application in a given context. But unlike in classic counter-feeding cases, only one of the rules can apply. Assume the rule R_1 applies to the input. Afterwards, no other rule can even attempt to apply, even if the context for another rule R_2 is met after the application of R_1. Thus R_1 blocks the application of the other rules; R_2 thus underapplies.

Disjunctive blocking is a core ingredient in realizational models of morphology (see e.g. Anderson 1992; Halle and Marantz 1993; 1994; Corbett and Fraser 1993; Aronoff 1994; Stump 2001): Morphosyntactic features are realized postsyntactically by realization rules (of rather different form in the various frameworks). A realization rule applies if it has a subset of the morphosyntactic features of the input element. Due to underspecification of the features in the realization rules, it is possible that several rules can in principle apply to an input. However, only one rule can do so (it is chosen by a version of the Elsewhere Principle). This rule blocks all other rules, even if the context for them is met as well.

Examples of other types of underapplication and overapplication as listed in Baković (2011) can also be found with the interaction of syntactic rules.[4] In this thesis,

---

[3]See Grimshaw (1997: 375) for the observation that OT constraints are likely not to be surface-true because they are violable.

[4]For instance, syntactic rules can be in a *mutual bleeding* relation: The early application of a rule R_1

however, I will confine myself to the classic cases of opacity, viz. counter-feeding and counter-bleeding, that can be accounted for by rule ordering as abstractly depicted in (6) and (7).

Rule interactions have been a major topic in phonology and syntax since the earliest days of generative grammar. For an overview of interactions of phonological rules see e.g. Chomsky and Halle (1968), Anderson (1969), Anderson (1974), Koutsoudas et al. (1974), Kenstowicz and Kisseberth (1977; 1979), Baković (2011); on rule interaction in morphology see Embick (2000; 2010); Arregi and Nevins (2012). The most comprehensive treatment of this topic in syntax is Pullum (1979), providing a critical discussion of the literature on rule ordering from the beginnings of transformational grammar in the 1950s until the mid-1970s; see also Ross (1967); Williams (1974); Kayne (1975); Perlmutter and Soames (1979), McCawley (1984; 1988), Řezáč (2004), Lasnik (2001) and Brody (2002).

1.2 Rule interactions in syntax

In this section I will summarize how transparent and opaque rule interactions have been treated at different stages of generative grammar. Section 1.2.1 introduces how syntactic transformations interact and what these interactions tell us about whether and how rules are ordered. Section 1.2.2 shows what strategies have been proposed to capture transparent and opaque interactions in the Government and Binding framework. Finally, in section 1.2.3 I will introduce the Minimalist operations Merge and Agree and how they can interact.

1.2.1 Rule interactions in transformational grammar

1.2.1.1 Examples of interactions

In the "standard theory" of transformational grammar (Chomsky 1965) the grammar consists of a base component and a transformational component. The base component contains phrase structure rules (PSR) that generate hierarchical deep structures (DS) with elements from the lexicon; PSRs will not concern us here. The transformational component contains transformations (T) that map a structure onto another structure in order to produce the surface structure (SS) of a linguistic expression. A deep structure can undergo several transformations until the surface structure is reached. In this case, intermediate structures (IS) are created that neither correspond to the DS nor the SS. The most widespread assumption in the transformational grammar literature was that transformations apply sequentially, i.e. one after another (see section 1.2.1.2 for detailed discussion

---

destroys the context for the subsequent application of a rule R_2, but R_2 would also bleed R_1 if the order of rules was reversed. If R_1 applies before R_2, R_1 overapplies because it could potentially be bled by R_2. An example of mutual bleeding from the Minimalist literature is the interaction of case assignment and A-movement. Assume that an XP can only be A-moved if it has an unvalued case feature (*Activity Condition*, Chomsky 2001), and that there is a head H that triggers both A-movement of an XP to SpecHP and case assignment to XP if H c-commands XP. If XP first moves to SpecHP, subsequent case assignment to XP by H is bled because H no longer c-commands XP. But the reverse order also leads to bleeding: If H first assigns case to XP, XP cannot be A-moved to SpecHP anymore because it is has a case value and is thus inactive. The bleeding order movement before case assignment could be used to derive hyperraising (movement out of a finite clause e.g. in Brazilian Portuguese), and subject-object reversal (in Bantu): In both cases, an XP is A-moved out of a domain D where is should have been assigned case; hence, movement out of D should be bled, but it is not (counter-bleeding). This follows if the order of operations in languages with these constructions is such that the XP moves *before* it is deactivated by case assignment.

of the mode of rule application): The first transformation T_1 takes as its input the DS and creates an intermediate structure, IS_1; the second transformation T_2 takes as its input the output of T_1 and creates another representation IS_2 and so on until all obligatory transformations have applied and the SS is derived. This is schematically illustrated in figure 1.2:

a. Lexicon + PSR → $\boxed{\text{DS}}$

b. $\boxed{\text{DS}} \xrightarrow{T_1} \boxed{IS_1} \xrightarrow{T_2} \boxed{IS_2} \quad \ldots \quad \xrightarrow{T_n} \boxed{\text{SS}}$
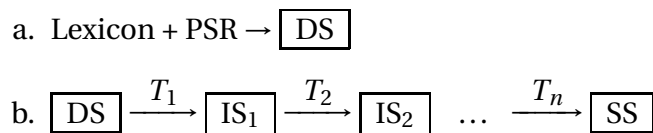
Figure 1.2: Model of the grammar ('standard theory')

In this model of grammar, the "rules" that can interact are transformations. In this subsection, I will present examples for all types of interaction in (1) and (2) from the transformational grammar literature.[5] We will come back to most of these examples in the course of this chapter when discussing different ways of deriving opacity effects. In what follows, the subscripts *n* and *k* are movement indices, and *i* and *j* are referential indices; the base position of a moved constituent is indicated by an underscore.

First, consider the following case of *feeding* that arises through the interaction of the two transformations Passive and There-insertion. These are defined in (8) and (9), respectively.

(8)   *There-insertion (Perlmutter and Soames 1979: 51-52)*:
        There replaces the indefinite subject NP of a verb of existence. The subject NP moves to the right of the verbal element.

(9)   *Passive (cf. McCawley 1988: 76)*:
        A passive clause is derived from an active clause by "replacing the subject by one of the NPs in the VP, inserting *be* before the VP, and either deleting the underlying subject or combining it with *by* and putting the resulting combination at the end of the VP".

There can only be combined with a restricted class of verbs, e.g. with verbs of existence such as *be*, cf. (10). In an active transitive sentence with a verb that is not from the class of verbs of existence, *there* cannot occur, cf. (11).

(10)   *'There' combined with 'be' (Perlmutter and Soames 1979: 42)*:
        a.   There is a mouse in the bathtub.
        b.   There is a mouse over there.

(11)   *'There' combined with a transitive verb (Perlmutter and Soames 1979: 50)*:
        a.   The policeman killed a demonstrator.
        b.   *There killed the policeman a demonstrator.

The Passive transformation introduces the auxiliary *be*, cf. (12-a) based on (11-a):

(12)   a.   A demonstrator was killed by the policeman.
        b.   There was a demonstrator killed by the policeman.

---

[5]The definitions of the transformations are formulated in a very surface-oriented way; nowadays, the corresponding rules / operations would be defined in a more abstract way. But the formulations from the transformational literature are sufficient to illustrate interaction types and the general question about the mode of rule application which the interactions pose.

If the subject of a passivized clause is an indefinite NP, There-insertion can apply since Passive introduced *be*, a verb of existence, cf. (12-b). Thus, if Passive applies before There-insertion, the former feeds the latter.

Next, consider the interaction of the transformations Passive and Equi-NP-deletion in English, which instantiates a *bleeding* configuration (example taken from McCawley 1988: 153f.). Passive was already defined in (9), Equi-NP-deletion is introduced in (13):

(13)   *Equi-NP-deletion (based on McCawley 1988: 120, 138)*:
       The subject of a non-finite clause is deleted if it is co-referent with an NP in the main clause.

Assume that the sentence in (14-a) has the deep structure in (14-b), in which the subjects of the matrix clause and the embedded clause are co-referential.

(14)   *Passive bleeds Equi-NP-deletion (McCawley 1988: 153)*:
       a.   SS: John$_i$ doesn't want anyone to be hurt by him$_i$.
       b.   DS: John$_i$ not want (that) he$_i$ hurt someone.

In principle, Equi-NP-deletion could apply in (14-b) with the result that the subject of the embedded clause is deleted. If passive applied afterwards, we would expect that the underlying subject of the embedded clause cannot be realized in a *by*-phrase because it has been deleted before by Equi-NP-deletion. However, since the underlying subject can appear in a *by*-phrase in the surface structure (14-a), Passive must have applied before Equi-NP-deletion. First, Passive maps (14-b) onto the structure in (15).

(15)   *Passive applied to (14-b)*:
       John$_i$ not want (that) someone be hurt by him$_i$.

Afterwards, Equi-NP-deletion can no longer apply because the NP in the embedded clause that is co-referent with the matrix subject is no longer the subject of the embedded clause. Hence, the underlying subject can be overtly realized in a PP. Passive in the embedded clause thus bleeds Equi-NP-deletion.[6]

I now turn to opaque interactions of transformations. A well-known example of *counter-feeding* is the interaction of the transformations Wh-movement and Wanna-contraction in English (cf. Lakoff 1970; Bresnan 1978), as defined in (16) and (17).

(16)   *Wh-movement (based on Question Movement in Perlmutter and Soames 1979: 251)*:
       The wh-phrase is moved to the sentence-initial position, thereby creating a gap.

(17)   *Wanna-contraction*:
       want and *to* can be contracted to *wanna* if they are adjacent.

(18-a) shows that a wh-phrase in English cannot remain in its theta-position (unless it is an echo-question). Rather, it must be moved to the initial position of the clause (accompanied by a rule of *do*-insertion), see (18-b). (18-c) illustrates Wanna-contraction: Since *want* and *to* are adjacent in the control construction in (18-b), they can be contracted, see (18-c).

---

[6]Further transformations need to apply to (15) to derive the SS in (14-a), but these are not relevant for the present discussion.

(18)   *Wh-movement and Wanna-contraction in control constructions*:
  a. *You want to meet who?
  b. Who do you want to meet?
  c. Who do you wanna meet?

However, *want* and *to* cannot be contracted in an ECM-construction like (19-a); this leads to ungrammaticality (cf. (19-b)), although *want* and *to* are adjacent on the surface. Hence, it is unclear why Wanna-contraction did not apply although its context is met on the surface, a case of underapplication.

(19)   *Blocking of wanna-contraction in ECM constructions*:
  a. Who do you want to meet Mary?
  b. *Who do you wanna meet Mary?
  c. [$_{TP}$ you want who to meet Mary]

The impossibility of Wanna-contraction in ECM constructions can be explained as follows (cf. Bresnan 1978): There is a stage in the derivation of (19-b) where the wh-phrase intervenes between *want* and *to*, as shown in (19-c); this is the case before Wh-movement to the matrix clause applies. Hence, if Wh-movement applies after Wanna-contraction, it follows that the latter is counter-fed by the former: At the point where Wanna-contraction applies, the wh-word still intervenes between *want* and *to*; it is moved 'out of the way' only afterwards (Wh-movement comes too late to feed Wanna-contraction). In the control construction in (18-b), however, the wh-word never intervenes between *want* and *to*, cf. the underlying structure in (20).

(20)   Who$_k$ do you want PRO to meet ___$_k$?

Under the standard assumption that the empty category PRO occupies the subject position of the embedded clause (Chomsky 1981), and that PRO does not count as an intervener, *want* and *to* can be contracted in control contexts.[7]

 An example of *counter-bleeding* that has been recurring in the literature of transformational grammar is the interaction of Reflexivization and Imperative Subject Deletion in English (see e.g. Postal 1964; Jacobs and Rosenbaum 1968; Lakoff 1968; Perlmutter 1971; Perlmutter and Soames 1979). If the direct object of a transitive verb in an imperative is co-referent with the 2nd person subject, the object must be realized as a reflexive pronoun:

(21)   *Reflexivization in imperatives*:
  a. (You$_i$) Defend yourself$_i$!
  b. *(You$_i$) Defend you$_i$!

The brackets in (21) indicate that the subject of an infinitive can optionally be deleted. The two relevant transformations are defined as follows:

(22)   *Reflexivization (McCawley 1988: 26)*:
  A deep structure NP must be expressed as a reflexive pronoun if a preceding NP
  of the same clause has the same reference.

---

[7]We will come back to the question whether empty elements that intervene between *want* and *to* always block contraction in section 1.2.2.

(23)   *Imperative subject deletion (ISD, optional, McCawley 1988: 25)*:
         A deep structure 2nd person subject (*you*) is deleted when it is the subject of an
         imperative sentence.

In principle, ISD could bleed Reflexivization: Since the transformation Reflexivization
makes reference to the subject of a transitive verb, it cannot apply if the subject is deleted
prior to Reflexivization by the application of ISD. However, there is no bleeding in (21-a).
On the surface, it is unclear why Reflexivization could apply because its structural de-
scription is not met in the output (the co-referent subject is missing), a case of counter-
bleeding. The opaque interaction can be derived if Reflexivization applies before ISD: At
the point where the former applies, the subject is still present in the structure, it is deleted
only afterwards.

   Another well-known class of examples of *counter-bleeding* that also involves reflex-
ives is called "Antecedent Removal" by Pullum (1979: 110ff.): On the surface, a reflex-
ive pronoun precedes its antecedent, but since Reflexivization only applies if the reflex-
ive pronoun is c-commanded by its antecedent in an A-position (Principle A, Chomsky
1981) opacity arises.[8] An instance of this pattern is shown in (24-b) (cf. the discus-
sion in Pullum 1979: 35ff.), where Reflexivization interacts with Topicalization. Simpli-
fying somewhat, Topicalization is a transformation that puts a (non-wh) constituent in
clause-initial position:

(24)   *Interaction of Reflexivization and Topicalization*:
         a. I_i never buy anything for myself$_i$.
         b.   [Myself$_i$]$_k$, I_i never buy anything for ___$_k$.

(24-a) shows the application of Reflexivization: The reflexive pronoun occurs because a
co-referent NP in the same clause precedes the complement of the preposition. Topi-
calization of the reflexive pronoun is possible, cf. (24-b), although on the surface it is
unclear how Reflexivization could apply given that its structural description is not met:
The potential antecedent, the subject pronoun *I*, follows the reflexive and hence does not
c-command it. Since Topicalization could bleed Reflexivization but does not, this is a case
of counter-bleeding. The opacity can be derived by ordering of transformations: Reflex-
ivization applies before the reflexive pronoun is topicalized.

   Another *counter-bleeding* interaction is discussed in Ross (1969: 262-264). He ar-
gues that Wh-movement (defined in (16)) must apply before Sluicing (cf. (25)) in order
to derive the fact that the wh-word in the embedded clause in (26-a) is not deleted by
Sluicing.

(25)   *Sluicing (cf. Ross 1969: 252)*:
         Delete everything but the preposed constituent of an embedded clause under the
         condition that the remainder of the question is identical to some other part of a
         preceding sentence.

(26)   *Interaction of Sluicing and Wh-movement Pullum (1979: 33)*:
         a.   [$_{CP}$ He is writing something [$_{CP}$ but you can't imagine [$_{CP}$ what]]].
         b.   [$_{CP}$ He is writing something [$_{CP}$ but you can't imagine [$_{CP}$ he is writing
              wh+something]]].

---

[8]In the definition of Reflexivization in (22), the crucial factor is linear precedence instead of c-command.
For the simple examples in this section it suffices to equate 'x precedes y' with 'x c-commands y'.

The D-structure representation of (26-a) is given in (26-b). If Sluicing applied before Wh-movement in the most deeply embedded clause, it would delete the wh-phrase contained in it, and Wh-movement would thus be bled. However, since the wh-phrase occurs in the embedded clause in (26-b), Wh-movement must have applied earlier than Sluicing. So Sluicing counter-bleeds Wh-movement.

A final prominent example of counter-bleeding is Anti-reconstruction. Look at the example in (27) where a Principle C violation is circumvented by movement:

(27)   *Principle C violation under reconstruction (Lebeaux 1988: 103)*:
         [Which pictures near John$_i$]$_k$ did he$_i$ destroy ___$_k$?

A Principle C violation obtains if an R-expression is c-commanded by a co-indexed (viz. co-referential) element (cf. Chomsky 1981):

(28)   Principle C:
         An R expression must not be c-commanded by a co-indexed element.

In (27), the R-expression *John* is contained in a wh-phrase that is base-generated as the internal argument of the verb *destroy*. In its base-position, it would violate Principle C because the R-expression is c-commanded by a co-referential pronoun. Wh-movement of the wh-phrase containing the R-expression should bleed a Principle C effect because on the surface, the R-expression is no longer in the c-command domain of the pronoun. Indeed, this is the case in (27), the sentence is grammatical. Interestingly, there is an argument/adjunct asymmetry with respect to Principle C effects under Wh-movement: Only R-expressions contained in adjuncts cancel a principle C violation; R-expressions contained in arguments, however, continue to cause a Principle C violation. This effect is known as the *Freidin-Lebeaux-Generalization* (van Riemsdijk and Williams 1981; Freidin 1986; Lebeaux 1988; 1990) Compare (29-a) (repreated from (27)) and (29-b).

(29)   *Argument/adjunct asymmetry for Principle C (Lebeaux 1988: 103)*:
         a.   [Which pictures near John$_i$]$_k$ did he$_i$ destroy ___$_k$?
         b.   *[Which pictures of John$_i$]$_k$ did he$_i$ destroy ___$_k$?

In (29-a), the R-expression is contained in an adjunct to the noun *picture* and no Principle C violation obtains under wh-movement, as expected; a Principle C violation is bled. In (29-b), however, the R-expression is contained in an argument to the noun *picture* and Principle C is violated, although the R-expression is not in the c-command domain of the co-indexed pronoun on the surface – a case of counter-bleeding. Lebeaux (1988: ch. 3.3) explains this effect through rule ordering: First, he assumes that Principle C holds at any level (DS, SS, IS). It follows that if an R-Expression is in the c-command domain of a co-indexed pronoun at DS, prior to Wh-movement, the structure will inevitably be illicit, even if subsequent transformations move the R-expression out of the offending domain. This explains the ungrammaticality of (29-b): The wh-phrase containing the R-expression is c-commanded by the pronoun in its D-structure position, indicated by underline. But the same state of affairs seems to hold in (29-a) where the R-expression is contained in an adjunct. In order to explain the difference, Lebeaux (1988: 105) proposes that adjuncts need not be present at D-structure, they can be added late at S-structure by the rule Adjoin-α, i.e. adjuncts can be inserted counter-cyclically.[9]

---

[9]I use the term *rule* here because Lebeaux's (1988) analysis is couched in the Government and Binding framework (see section 1.2.2). But for the present purposes we can think of Adjoin-α as a transformation

(30)   *Adjoin-α (cf. Lebeaux 1988: 105)*:
       Chomsky-adjoin an adjunct into another tree (counter-cyclically).

Adjoin-α can apply after Wh-Movement. If it does, the adjunct *near John$_i$* in (30-a), containing the R-expression, is not present in the base position of the wh-phrase and thus not c-commanded by the co-indexed pronoun. The adjunct is inserted into the wh-phrase only after Wh-Movement has taken place. In the S-structure representation, the adjunct is too high in the structure to be c-commanded by the pronoun. Hence, Principle C is not violated on any level if the R-expression is contained in an adjunct. Crucially, late insertion of XPs (by Adjoin-α) is restricted to adjuncts.[10] Thus, the argument DP in (29-b) is moved from its base position to its surface position; and in its base position, the R-expression contained in the argument DP inevitably violates Principle C. This explains the contrast between R-expressions in moved adjuncts vs. moved arguments. Late merger of adjuncts has become the standard analysis of the Anti-reconstruction effect and is also pursued by Freidin 1994; Chomsky 1995; Epstein et al. 1998; Safir 1999; Fox 1999; 2000 among many others, but see Chomsky 2001 for a different approach; for detailed discussion and evidence against the adjunct/argument dichotomy based approach see Fischer (2004).[11]

Further instances of opaque interactions of syntactic transformations can be found in Pullum (1979). In the next section I will discuss what transparent and opaque rule interactions tell us about the mode of rule application.

---

that maps an input structure onto another structure that contains an adjunct, i.e. the transformation inserts material.

[10]Late insertion is only available for adjuncts because they are not subject to the *Projection Principle*. This principle states that selectional requirements of lexical items must be preserved at every syntactic level (cf. Chomsky 1981; 1986*b*). Consequently a moved argument must be represented by an element (e.g. a trace) in its D-structure base position. Since adjuncts are not selected, they are not represented at D-structure. Therefore, a Principle C violation for R-expressions contained in adjuncts is avoided.

[11]One might wonder why the effect is known as "Anti-reconstruction" because Lebeaux's (1988) account does not literally involve reconstruction; the wh-phrase containing the R-expression is not lowered into its base position in order to derive the Principle C violation in (29-b). Lowering is not required because Lebeaux assumes that Principle C holds on all levels of representation. However, it has also been argued that the Binding Theory only holds at a single level, i.e. at LF (cf. e.g. Chomsky 1995). Given this assumption, another transformation, viz. Reconstruction, is needed in order to account for the ungrammaticality of examples like (29-b) (unless the copy theory of movement is pursued, see below). Reconstruction puts the wh-phrase back into its base position. If the reconstructed wh-phrase contains the R-expression (in case it is contained in an argument), a Principle C violation obtains at LF because the wh-phrase is in the c-command domain of the co-referential pronoun. The grammatical (29-a) instantiates an "Anti-reconstruction effect" because it seems that the wh-phrase is not reconstructed, otherwise a Principle C violation should occur. Importantly, in an account that makes use of Reconstruction (cf. e.g. Freidin 1986; 1994), regardless of how exactly this is formally implemented, one would have to assume that the insertion of the adjunct (by Adjoin-α) must be ordered after Reconstruction in order to explain the absence of a Principle C violation. It does not suffice to insert adjuncts after Wh-Movement, but before Reconstruction because the adjunct containing the R-expression would then be reconstructed as well and should induce a Principle C violation. But the order Reconstruction before Adjoin-α is impossible: Reconstruction is an LF-process and if adjoin-α applies after it, it must be an LF-process, too; but then the adjunct could not be phonologically realized: Processes at LF cannot feed PF in the Y-model. Furthermore, if Adjoin-α applies to a constituent X after Reconstruction of X, where is α adjoined to? This presupposes that reconstruction movement leaves a reflex in its structurally higher surface position where α can adjoin to. To conclude, Anti-reconstruction effects cannot be modeled in any obvious way if Reconstruction literally involves lowering of an element to its base position.

Lebeaux's (1988) account is similar to more recent minimalist accounts of Anti-reconstruction effects that assume that there is no Reconstruction operation (lowering of a moved phrase into its base position). Rather, movement involves copying of the element that is to be moved and remerging it (Chomsky 1995; Fox 1999; 2000). Late merge of an adjunct containing an R-expression is thus adjunction to the higher copy of the wh-phrase; the adjunct is not present in the lower copy which is c-commanded by the co-indexed pronoun and hence, no Principle C violation obtains.

1.2.1.2 Sequential vs. simultaneous application

Until now, I have presupposed that transparent and opaque rule interactions are resolved by rule ordering: A rule that applies early can feed or bleed the application of a subsequent rule; if the order of rules is reversed, we get counter-feeding and counter-bleeding, respectively. However, there are alternative ways to derive at least a subset of rule interactions. In this section, I will summarize the proposals made for the mode of rule application in transformational grammar.

A central question in the phonological as well as the syntactic literature in the 1960s and 1970s was the following: If more than one rule applies to a deep structure to obtain a surface structure as in figure 1.2, are these rules ordered with respect to each other and if yes, how are they ordered? Three major hypothesis are discussed in the literature (see e.g. Chomsky and Halle 1968):[12]

STRICT ORDERING HYPOTHESIS (see, among many others, Shamir 1961; Chomsky 1965; Chomsky and Halle 1968; Kisseberth 1972a; Williams 1974; Kenstowicz and Kisseberth 1979): Rules are ordered on a list for a given language (cf. Chomsky 1965: 67). Rules apply sequentially, one after another. A rule R is only applicable when its turn on the list has come, neither before nor after the point. If its structural description is met at that point, the rule can or must apply (depending on whether the rule is optional or obligatory); if its structural description is not met at that point, it does not apply and the next rule on the list is considered. The first rule on the list takes as its input the deep structure representation of a linguistic expression. All rules that apply later take as their input the output of the immediately preceding rule, cf. figure 1.3. Each rule can apply only once.

    a. List of ordered rules:

        (i) R_1
       (ii) R_2
      (iii) R_3
      (iv) ...

    b. Application of rules:

$$\boxed{\text{DS}} \xrightarrow{R_1} \boxed{\text{IS}_1} \xrightarrow{R_2} \boxed{\text{IS}_2} \ldots \xrightarrow{R_n} \boxed{\text{SS}}$$
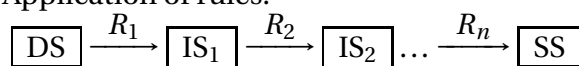
Figure 1.3: Abstract example of strictly ordered rules

UNORDERED SEQUENTIAL APPLICATION HYPOTHESIS (Koutsoudas 1972; Ringen 1972; Lakoff and Kisseberth 1972): Rules apply sequentially, one after another. However, there is no fixed order of rules. Either an arbitrary order is established among the rules for each derivation and they apply in that order, or no order is fixed at all and the rules apply in a random fashion, cf. figure 1.4. Each rule can apply only once.

---

[12]The hypotheses have been given a number of different names in the literature. In what follows, I use the most common and self-explanatory labels.

a. Derivation 1 – R$_1$ applies before R$_2$:

$$\boxed{\text{DS}} \xrightarrow{R_1} \boxed{\text{IS}_1} \xrightarrow{R_2} \boxed{\text{SS}}$$
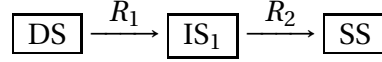
b. Derivation 2 – R$_2$ applies before R$_1$:

$$\boxed{\text{DS}} \xrightarrow{R_2} \boxed{\text{IS}_1} \xrightarrow{R_1} \boxed{\text{SS}}$$
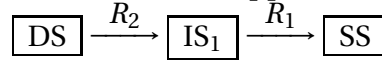
Figure 1.4: Abstract example of unordered sequential application

SIMULTANEOUS APPLICATION HYPOTHESIS (Pullum 1979): Rules are not ordered. A rule can or must apply (depending on whether it is optional or obligatory) as soon as its structural description is met. If more than one rule is applicable to a structure, all rules whose structural description is met apply simultaneously, cf. figure 1.5.

$$\boxed{\text{DS}} \xrightarrow[R_1]{R_2} \boxed{\text{SS}}$$
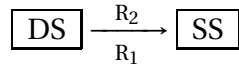
Figure 1.5: Abstract example of simultaneous application

A number of alternative hypotheses have been proposed in the literature. Often, these hypotheses mix ingredients of the three major hypotheses (cf. McCawley 1988: 157). For example, other hypotheses are the *Iterative Application Principle, Directional Iterative Application* (see Kenstowicz and Kisseberth 1979: 325ff., Johnson 1971 and Morin and Friedman 1971 for discussion). The properties of unordered application and strict ordering are combined in Anderson (1969; 1974) such that rules apply sequentially, but they are only partially ordered; simultaneous application and unordered application are combined in Koutsoudas et al. (1974). Koutsoudas (1972); Koutsoudas et al. (1974) and following them Pullum (1979) have adopted a mixture of simultaneous and sequential application to which I will come back in section 5.2.

For the present discussion, I concentrate on the three hypotheses introduced above. Transparent and opaque rule interactions have played a central role in choosing between these hypotheses. In sum, the *Strict Ordering Hypothesis* has gained the most support because the two alternative hypotheses make empirically wrong predictions when it comes to transparent interactions. I will briefly summarize the arguments in what follows.

Let me begin with the *Simultaneous Application Hypothesis* according to which all rules whose structural description is met in the input are applied simultaneously. Simultaneous application has been considered to be a serious alternative to sequential application because it can account for opaque rule interactions (counter-feeding and counter-bleeding, see e.g. Kenstowicz and Kisseberth (1979: 291-92)): Since all rules apply to deep structure simultaneously, no intermediate structures are created and hence no rule can feed or bleed the applicability of another rule. For instance, in the abstract bleeding example in (5) (repeated in (31)), the structural description of two rules R$_1$ and R$_2$ in (3) are met in the input structure AB.

(31)   AB $\xrightarrow{R_2}$ DB $\xrightarrow{*R_1}$

But early application of R$_2$ to the input AB creates the IS DB in which the SD of R$_1$ is no longer met because R$_1$ can only apply in the context of the symbol A, deleted by R$_2$. Such an interaction is not possible under simultaneous application: If the structural description of a rule is met at deep structure, it applies, regardless of what changes the other rules cause; hence R$_1$ and R$_2$ would both apply to DS. The avoidance of a potential bleeding

or feeding configuration by simultaneous application is exactly what is needed to produce counter-feeding or counter-bleeding: In the abstract example, the application of R_2 bleeds the application of R_1. But if they apply simultaneously, no IS is created by R_2 that serves as the input for R_1. Thus, there can be no bleeding of R_1. This account of opacity gets by without sequential application of rules in a fixed order. Kisseberth (1972b) and Kenstowicz and Kisseberth (1979: ch.8) argue against the *Simultaneous Application Hypothesis* (which they call *Direct Mapping Hypothesis*): Although it can handle opaque interactions, it cannot account for transparent interactions precisely due to the lack of intermediate structures. Let me explain this for feeding: A rule R_2 cannot apply to the deep structure, but application of R_1 creates the context for the application of R_2. This means that R_2 applies to the output structure that arises through application of R_1 to the deep structure. Under simultaneous rule application, however, the input for every rule is the deep structure. A feeding relation can thus not be described; the same is true for bleeding. Hence, in order to describe transparent rule interactions one must make reference to intermediate representations that neither correspond to deep nor surface structure. This is only possible if rules apply sequentially.

The existence of transparent interactions argues against simultaneous applications, but we still need to choose between the *Strict Ordering Hypothesis* with a fixed order of rules, and the *Unordered Sequential Application Hypothesis* with random ordering of rules. If a rule R_1 feeds a rule R_2 in language L, this follows under the former hypothesis if e R_1 precedes R_2 on the list and thus always applies before R_2. Under the latter hypothesis, R_1 and R_2 may apply in any order, the important point is that there is one possible order (R_1 is ordered before R_2) that produces a transparent interaction.

Kenstowicz and Kisseberth (1979: ch.8) also provide an argument against the *Unordered Sequential Application Hypothesis* (which they call Free Reapplication Hypothesis): Given this hypothesis, the two rules R_1 and R_2 may also apply in the reverse order which leads to counter-feeding. But in many languages, we get obligatory feeding if R_1 and R_2 interact, optional counter-feeding does not lead to a grammatical output; the same holds for bleeding relations which are usually not in free variation with counter-bleeding. At least in languages with an obligatory transparent interaction, rules cannot apply in any random order. Kenstowicz and Kisseberth (1979) make these arguments on the basis of interaction of phonological rules (see also Chomsky and Halle 1968: 342ff.); the same reasoning can be found for the interaction of syntactic rules in McCawley (1988).

To summarize, only the *Strict Ordering Hypothesis* can handle transparent interactions. Moreover, it can also capture opaque interactions by reversing the order of rules that produces a transparent interactions. The *Simultaneous Application Hypothesis* is too weak in that it cannot account for transparent interactions, whereas the *Unordered Sequential Application Hypothesis* is too strong in that it predicts that transparent and opaque interactions of two rules can freely alternate. Hence, the strict ordering of rules seems to be the only hypothesis that can handle transparent and opaque rule interaction in a uniform way. Indeed, this has also been the most widely adopted view in transformational grammar.

1.2.1.3   Extrinsic vs. intrinsic ordering

Given the conclusion that rules must be strictly ordered, another question arose in the discussion: If two rules interact, does the order follow from anything? Chomsky (1965) calls an order that must be stipulated *extrinsic*; if the order follows from independent factors, e.g. the way the rules are formulated, he calls it *intrinsic*. If, for example, R_2 manipulates

the symbol B and R_1 introduces B, the latter must necessarily apply before the former if B indeed undergoes the structural change described in R_2; hence, the order is intrinsic.[13] Pullum (1979: 12) criticizes this distinction because there may be a configuration in which it is unclear whether the order is extrinsic or intrinsic in Chomsky's sense. Suppose there is a principle that is part of the grammar of every language; this principle demands that if the grammar has a rule with the property P and another rule with the property Q, then P always precedes Q. Now, is the order P ≻ Q in a given language L intrinsic or extrinsic? We could call it intrinsic because the order does not have to be stipulated – there is a universal principle in the grammar from which it follows; on the other hand, the principle itself might not follow from anything, and so the order it determines is extrinsic. Pullum suggests that the crucial distinction is actually whether the postulated orderings are *language-specific*, i.e. the central question is whether two languages can differ solely in the order of rules. Pullum (1979: 13) thus replaces the extrinsic / intrinsic distinction with the parochial / universal distinction:

(32)   *Parochial ordering constraint*:
          The order of rules is determined in the grammar of an individual language, i.e. languages can differ solely in the order of rules.

(33)   *Universal ordering constraint*:[14]
          The order of rules is not a property of the grammar of an individual language; rather, it follows from universal principles.

According to Pullum, most researches dealing with the determination of rule ordering have used the terms extrinsic (= parochial) and intrinsic (= universal) in that sense of language-specific vs. universal anyway. Since Pullum's terminology did not prevail, I will continue to use the more wide-spread notions extrinsic and intrinsic, though with the crucial distinction introduced by Pullum, i.e. whether two languages can differ solely in the order of rules. This will be important for the discussion in section 5.2.

Starting from Pullum's (1979) definitions, the core question is what qualifies as a universal principle that predicts rule ordering. I will briefly discuss the three most prominent principles introduced in transformational grammar (more will follow in section 5.2).

(a) The *Obligatory Precedence Principle*

---

[13]Note that the way in which rules are formulated does not actually *predict* the order of rules. In the abstract example in the text, R_1 feeds the application of R_2 by introducing a symbol that R_2 manipulates. However, nothing prevents the reverse order; it would simply produce counter-feeding. Thus, there seems to be a background assumption in Chomsky's reasoning that the order of R_1 and R_2 is intrinsic: The number of rules that successfully apply to a given input should be maximal. If two rules are in a feeding or counter-bleeding relation, both of them manipulate their input; but if they are in a bleeding or counter-feeding relation, only one of them manipulates the input. The feeding order of R_1 before R_2 is superior to the reverse counter-feeding order with respect to the number of successfully applied rules. The assumption is thus the following: If two rules can in principle induce a structural change in the input, choose the order of rules that produces the maximal the number of successfully applied rules (= intrinsic ordering). If there is evidence for the reverse order, this order is extrinsic.

[14]The order of rules can also follow from something else than a universal principle. For example, Pullum (1979: ch.1) assumes with Chomsky that the order can follow from the formulation of the rules. Furthermore, it can follow from the architecture of the grammar: Postcyclic rules, which Pullum identifies as what would be called morphophonological rules nowadays, necessarily apply after cyclic rules (rules that apply in syntax proper); precyclic rules would always apply before cyclic rules (although Pullum argues that there are no precyclic rules). Note that the conception of postcyclic morphophonological rules is reminiscent of the postsyntactic PF component in GB and Minimalism under a realizational model of morphology such as Distributed Morphology (Halle and Marantz 1993; 1994): Operations at PF cannot feed or bleed syntactic operations, but they can be fed or bled by syntactic operations.

(b) The *Elsewhere Principle* (Specificity Condition, Elsewhere Principle, Blocking Principle, Panini's Principle, Proper Inclusion Principle)

(c) The *Cycle* (Cyclic Principle)

I begin with the *Obligatory Precedence Principle* originally formulated in Ringen (1972) and adopted by Perlmutter and Soames (1979: 133) under the label *Immediate Characterization*.

(34)   *Obligatory Precedence Principle*:
       Obligatory rules precede optional rules.

Perlmutter and Soames (1979: 182ff) show that this principle determines the order Reflexivization before Imperative Subject Deletion (ISD) that is necessary to derive the example in (21-a), an instance of counter-bleeding: ISD should bleed Reflexivization, but doesn't. Reflexivization is obligatory as we have seen, but ISD is optional – an imperative can contain a subject, see (35):

(35)   *Optional Imperative Subject Deletion (Perlmutter and Soames 1979: 23)*:
       a.   You go home at once!
       b.   Go home at once!

Given (34), it follows that Reflexivization precedes ISD; the latter comes too late to bleed the former. Note that the principle makes no predictions if two rules are both optional or both obligatory.

The *Elsewhere Principle* (also called Specificity Condition, Blocking Principle, Panini's Principle, Proper Inclusion Principle, cf. Anderson 1969; Kiparsky 1973*b*; Sanders 1974; Koutsoudas et al. 1974; DiSciullo and Williams 1987; Fanselow 1991; Anderson 1992; Lumsden 1992; Noyer 1992; Halle and Marantz 1993; 1994; Williams 1994; Halle 1997; Williams 1997; Wiese 1999; Stump 2001) demands that if two rules compete for application, the more specific rule applies and blocks the less specific rule. Hence, the principle is disjunctive: Only one of the rules can apply, even if the application of the more specific one does not destroy the context for the application of the less specific one. A version of this principle is given in (36):

(36)   *Elsewhere Condition (version proposed by Pullum 1979: 50)*:
       For any representation R, which meets the structural description of each of the two rules A and B, A takes applicational precedence over B with respect to R if and only if the structural description of A properly includes the structural description of B.

An example for its application in syntax is the interaction of Extraposition and It-deletion (Pullum 1979: 47f.). *it* and the clause S it is associated with are assumed to be underlyingly adjacent, but they must not be adjacent on the surface, cf. (37-a). Since the transformation Extraposition in (38-a) is more specific than It-deletion in (38-b) (the former properly contains the structural description of the latter), the conflict is resolved by extraposing S, cf. (37-b). Thus, Extraposition blocks It-deletion.

(37)   *Interaction of Extraposition and It-deletion*:
       a.   *It that Purvis has been made Professor of Verbs is great.
       b.   It is great that Purvis has been made Professor of Verbs.

(38)   *Structural descriptions of Extrapositon and It-deletion (Pullum 1979: 50)*:
    a.   Extraposition: X [$_{NP}$ it [$_S$ that/for X]] X
    b.   It-deletion: X [$_{NP}$ it [$_S$ X]] X

The most widely adopted principle that determines rule ordering is the *Cycle* or the *Cyclic Principle.* Since it will figure prominently in this thesis, I will discuss it in some detail. The *Cycle* was first formulated in Chomsky et al. (1956: 75) as the "transformational cycle" and used extensively for the interaction of phonological rules in Chomsky and Halle (1968); in syntax the Cycle was introduced to govern the application of transformations in Chomsky (1964; 1965):[15]

(39)   *Cyclic Principle (definition taken from McCawley 1988: 28)*:
    When one domain to which transformations can apply is contained in another, the applications of transformations to the smaller domain precede the applications of transformations to the larger domain.

Let me go through an example. In transformational grammar, a syntactic structure consists of domains. Standardly, each S and NP node (corresponding to CP and DP in contemporary terminology) constitutes a domain. The domains are distinguished by labeling them with a number index such that the topmost S node is labeled S_0 and lower S nodes are labeled with number indices increasing by one. A structure with three S domains would thus look as follows:

(40)   *Structure with three S domains*:
    [$_{S_0}$ NP [$_{VP}$ V [$_{S_1}$ NP [$_{VP}$ V [$_{S_2}$ NP [$_{VP}$ V NP]]]]]]

Transformations can apply to subdomains in a structure. The Cyclic Principle states that an operation to the cyclic domain D_x will precede any operation to the cyclic domain D_{x-1}. An example McCawley introduces is the interaction of the transformations Passive and Equi-NP-deletion in English (cf. section 1.2.1.1). In (41-a) with the DS in (41-b), Passive bleeds Equi-NP-deletion.

(41)   *Passive bleeds Equi-NP-deletion (McCawley 1988: 153)*:
    a.   SS: John$_i$ doesn't want anyone to be hurt by him$_i$
    b.   DS: [$_{S_0}$ John$_i$ not want [$_{S_1}$ (that) he$_i$ hurt someone]]

The DS has two domains, S_0 and S_1. Crucially, Passive applies only to the lower domain S_1 because only the embedded verb is passivized. Equi-NP-deletion, however, compares the reference of the subject of the matrix and the embedded clause; it's domain is thus S_0. The Cycle now correctly predicts that Passive applies before Equi-NP-deletion since the former applies to a domain that is contained in the domain of application of the latter. The result is that early application of Passive bleeds Equi-NP-deletion because after passivization, the structural subjects of the two clauses are no longer co-referent, see the intermediate structure after application of Passive to S_1 in (42).

(42)   *Passive applied to S_1*:
    [$_{S_0}$ John$_i$ not want [$_{S_1}$ (that) someone be hurt by him$_i$]]

---

[15]The definition of the Cycle in Chomsky and Halle (1968: 349) reads as follows: "Phonological rules apply in linear sequence to each phrase of the surface structure, beginning with the smallest, and proceeding to successively larger phrases until the maximal domain of phonological processes is reached."

Again, the question arises what happens if more than one rule applies to the same domain. In this case, the *Cycle* makes no predictions about the order of operations. Consider the interaction of the transformations Reflexivization and ISD again.

(43)   *ISD and Reflexivization in English*:
         a.   Defend yourself!
         b.   *Defend you!

A widely accepted analysis of (43) in transformational grammar is that it has the deep structure in (44-a), with the IMP(erative) morpheme being a daughter of S, just like the subject NP and the VP. First, Reflexivization applies and converts the object NP into a reflexive pronoun because it is co-referential with the subject, see (44-b). Afterwards, ISD applies and deletes the 2nd person subject in the presence of IMP, see (44-c).

(44)   *Counter-bleeding interaction of Reflexivization and ISD*:
         a.   S   *Reflexivization* →

                        IMP   NP   VP

                                   you   V   NP

                                          defend   you

         b.   S   *ISD* →

                        IMP   NP   VP

                                   you   V   NP

                                          defend   yourself

         c.   S

                 IMP   VP

                         V   NP

                       defend yourself
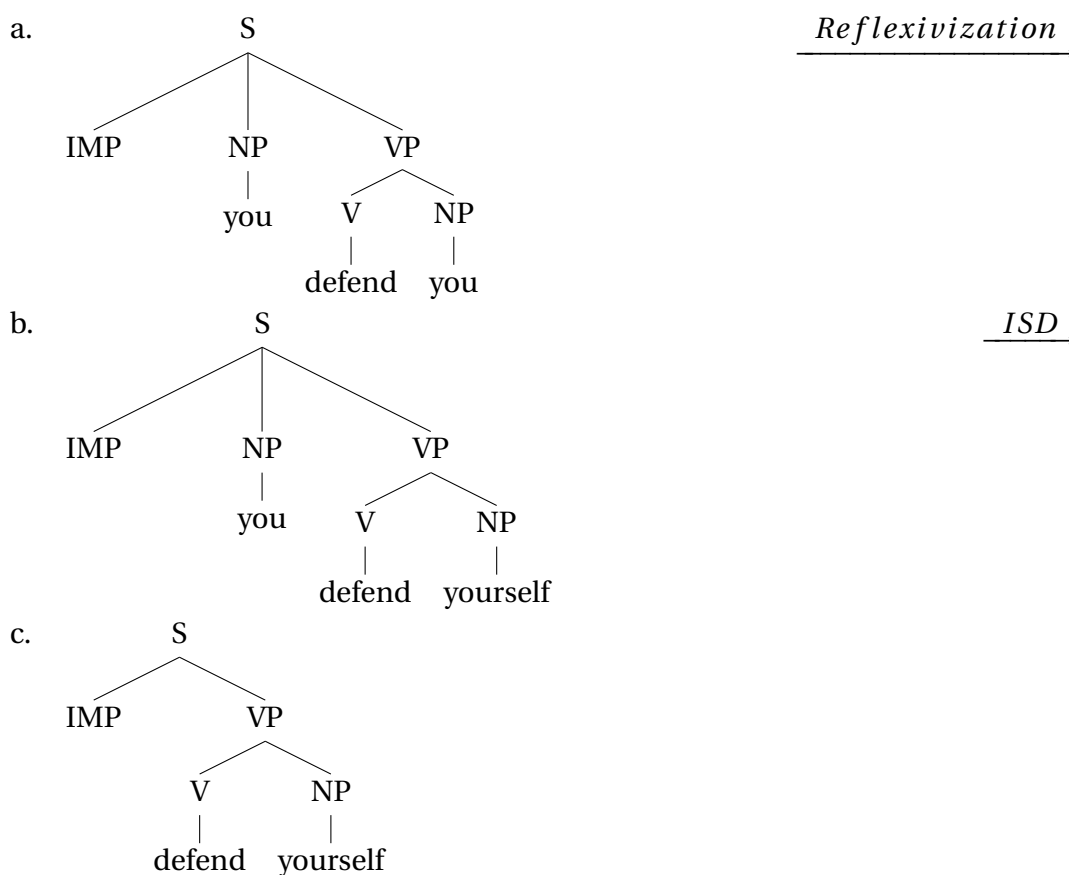
Recall that it is important that Reflexivization precedes ISD. If the transformations applied the other way around with the subject deleted before Reflexivization tries to apply, we would get the ungrammatical sentence in (43-b). The problem is that there is only a single S node in the structure, and hence only a single cycle to which both transformations apply. In this case, the Cyclic Principle does not predict the order of operations and it seems to be the case that the transformations must still be ordered extrinsically.

    However, it has been a desideratum to avoid extrinsic ordering whenever possible, and to make orderings follow from independent factors. Intrinsic ordering is conceptually preferred because it avoids the need to memorize a number of specific orderings. Two solutions have been proposed in the literature to avoid extrinsic ordering of rules that (seemingly) apply to a single cycle. Each is based on the idea that there are actually more

cycles in a structure than was previously thought. The solutions are (i) reanalysis of the deep structure and (ii) increasing the number of nodes that count as cyclic nodes.

Let us start with solution (i). A rather trivial example is again the interaction of Reflexivization and ISD presented by McCawley (1988). He argues that the deep structure of (43-a) is not (44-a) but rather (45) with IMP being the sister instead of a daughter of S. As a consequence, there are two cyclic domains in the clause (Lakoff 1972 for a similar proposal).

(45)   *Deep structure of (43-a) according to McCawley (1988: 159)*:
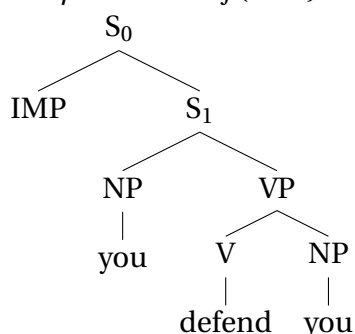
Given this structure, the Cyclic Principle correctly predicts that Reflexivization applies before ISD: Since ISD makes reference to the IMP morpheme, it must apply in the domain S_0; Reflexivization applies to S_1 since it only involves the two arguments of the verb that are included in the lower cycle. Since rules in lower domains precede rules in domains containing the former domains (cf. (39)), the correct ordering between the two rules follows independently. In this example, more S cycles are created by changes in the deep structure (which McCawley argues to be independently motivated by semantic considerations). However, such changes in deep structure may not be motivated for all examples in which rules interact. For this reason, solution (ii) has been introduced: McCawley (1988: 158) proposes to postulate more cyclic domains than S and NP, viz., it is not only S and NP that constitute a cyclic domain, but other nodes may do so as well. In particular, he suggests that every projection is a cyclic domain (McCawley 1988: 165, see also McCawley 1984). If transformations are relativized to apply to a particular cyclic domain, the increase in the number of cyclic domains has the advantage that the chance that two or more rules apply in the same domain is considerably decreased and the predictive power of the Cycle is maximized. For example, McCawley proposes that object-controlled Equi-NP-deletion as in (46) applies in the VP domain. If there were another transformation T_2 applying to S_0 in the same structure, then the Cyclic Principle would predict that Equi-NP-deletion applies before T_2 because the former applies in a lower domain, V_0. Under the previous definition of cyclic domains, the two transformations would have applied in the same cyclic domain S_0 and would have to be ordered extrinsically.

(46)   *Object-controlled Equi-NP-deletion*:
         $[_{S_0}$ John forced $[_{V_0}$ Mary$_i$ $[_{S_1}$ to $[_{V_1}$ (she$_1$ → Ø) empty the garage]]]]

Williams (1974) argues for a very similar concept. In order to explain the way in which rules interact, he proposes that the domain of application of a rule is large or small, depending on whether it applies early or late. Williams assumes the hierarchy of domains in (47-a), which refers to nodes in the clause structure in (47-b):

(47)   *Domains in Williams (1974)*:

    a.   Inclusion hierarchy: $\bar{S} \supset S \supset Pred \supset VP$

    b. $[_{\bar{S}}$ Comp $[_S$ NP $[_{Pred}$ $[_{VP}$ V NP$]]]]$

For him, each of the phrases (= $\bar{S}$, S, Pred Phrase, VP) constitutes a cyclic domain. Adopting the *Cycle*, he argues that a rule with the maximal domain D (i.e. the minimal XP in the inclusion hierarchy including the elements affected by the rule) applies before rules of any larger domain (Williams 1974: ch.1). In a simple clause, this has by and large the same effect as McCawley's (1988) proposal: Transformations in a phrase XP precede transformation in a dominating phrase YP.[16] To give an example, consider the interaction of the transformations Dative movement and Passive. Williams (1974: 17, 106ff) argues that Dative movement is a rule that applies to the domain Pred, whereas Passive is a rule in the domain S.

(48)   *Definition of transformations*:

    a.   Dative Movement$_{PRED}$: In a ditransitive context, an NP in a *to*-PP that bears the goal theta-role is moved in front of the NP with the theme theta-role. The preposition *to* is deleted.

    b.   Passive$_S$: See (9).

The empirical observation is that Dative movement feeds Passive of the goal NP:

(49)   John was given a book (by Mary).   *Williams (1974: 17)*

The underlying (simplified) structure of (49) is (50-a). Applying Dative movement to it results in (50-b). Finally, Passive applies and turns the goal NP into the subject of the clause.

(50)   a. $[_S$ Mary gave $[_{PRED}$ a book to John$]]$

    b. $[_S$ Mary gave $[_{PRED}$ John a book$]]$

    c. $[_S$ John was given a book by Mary$]$

Dative movement must apply before Passive in English because the reverse order results in an ungrammatical sentence.

(51)   *A book was given John.

Given the Cyclic Principle and relativization of the two transformations to certain domains, this order follows automatically, it is intrinsic: Dative Movement must apply before Passive because the former transformation applies in a domain that is included in the domain of the latter. The cyclic domains in Williams (1974) and McCawley (1988) are very small and the chance that two rules apply to the same small domain is therefore minimized. If such a situation occurred, however, the order of the rules must still be extrinsic, fixed language-specifically, as Williams (1974: 20) states. Do such cases exist? The answer is yes. The interactions I will be concerned with are exactly of that type: Two rules apply within the same domain, even under McCawley's strict version with every projection being a cyclic node. If there is no other principle that steps in in such a situation, extrinsic ordering is still needed. I will come back to the intrinsic / extrinsic dichotomy in section 5.2, when the data and the analysis will have been introduced.

---

[16]The *Cycle* with the domains proposed by McCawley (1988) and the *Inclusion hierarchy* introduced by Williams (1974) make different predictions when it comes to transformations that apply across clauses. William's theory is more restrictive than the Cycle in this case, see section 5.2.1 for discussion.

1.2.2 Rule interactions in representational frameworks: GB

We have seen that in a derivational framework where rules (e.g. transformations) apply sequentially, transparent and opaque rule interactions follows naturally from ordering. But how can such interactions be handled in a representational framework that does not incorporate rule ordering, but rather operates with constraints on the surface structure? The problem arises first and foremost for opaque interactions which are characterized by the fact that it cannot be read off of the surface representation why a certain rule has (has not) applied although its context is not (is) given in the output (generalizations are non-surface true in opaque interactions, see section 1.1 for discussion). This question came up for syntactic rule interactions in the mid 1970ies with the introduction of traces (Fiengo 1974) and culminated in the complete shift from a rule-based derivational to a representational model with the rise of the Government and Binding theory (GB) in Chomsky (1981) and subsequent work until the early 1990s. The general strategy in representational frameworks to derive opaque interactions is to enrich representations, i.e. to introduce abstract material that encodes (the equivalent of) earlier stages of the derivation. In addition, constraints can refer to such elements. For syntax, the following abstract elements have been proposed:[17]

(a) traces (see e.g. Wasow 1972; Fiengo 1974; 1977; May 1978; Chomsky 1981; 1986*a*; Lasnik and Saito 1984; Pollard and Sag 1994)

(b) pronominal elements (pro and PRO, cf. Chomsky 1981; 1982; Rizzi 1986*a*)

(c) copies (Grinder 1972; Chomsky 1995)

(d) indices (see the indexing theory in Chomsky 1981 and the various subtypes in Chomsky 1986*a*)

Let us look at some of the cases of opacity in syntax introduced in the last section and see how they are derived in a representational model of syntax like the Government and Binding framework. But let me first introduce the architecture of the grammar assumed in GB:

---

[17]The problem of how to derive opacity also came up in phonology with the shift from a rule-based, derivational system to Optimality Theory (OT, Prince and Smolensky 2004), where all constraints apply simultaneously to an underlying representation, and not sequentially. In standard OT, where optimization is global, it is not possible to derive opacity without further assumptions. This has been repeatedly pointed out in the phonological literature, see e.g. Idsardi (2000); Clements (2000); Kiparsky (2000); Odden (2000); Davis (2000); Bermúdez-Otero (2003); Embick (2010). In recent years, many solutions have been put forward in the literature. Representations are enriched with abstract material that represents earlier stages of a derivation, and constraints can refer to these elements; cf. Turbid phonology (Goldrick 2000), virtual phonology (Bye 2009), Coloured Containment (van Oostendorp 2006; Trommer 2011). Furthermore, new types of constraints have been postulated: McCarthy (2007) introduces PREC-constrains that demand a certain order of constraint violations (in a model with serial optimization); and McCarthy (1999) uses constraints that select a suboptimal candidate that corresponds to the representation of the optimal candidate at what would have been an earlier stage of the derivation in a derivational model. A special type of faithfulness constraints then demands that the optimal candidates share certain properties with the candidate chosen by the selector constraint. McCarthy (1999) calls this theory *Sympathy Theory*; for an application of Sympathy to OT syntax see Müller (2002). Another strategy is to to reintroduce the derivational character by abandoning global evaluation and by replacing it with local optimization where the generator and the evaluation process are interleaved: Only small subparts of the structure are evaluated; the optimal candidate of the first evaluation is the input for the next evaluation and so forth. Constraints of a later evaluation can lead to changes that mask effects of constraints of an earlier optimization, which leads to opacity. This is the case in Harmonic Serialism (McCarthy 2010) and in Stratal OT (see e.g. (Kiparsky 2000; Bermúdez-Otero 2003; to appear)). This strategy of serial optimization has also been applied to OT syntax, see Müller (2002) and references cited there.
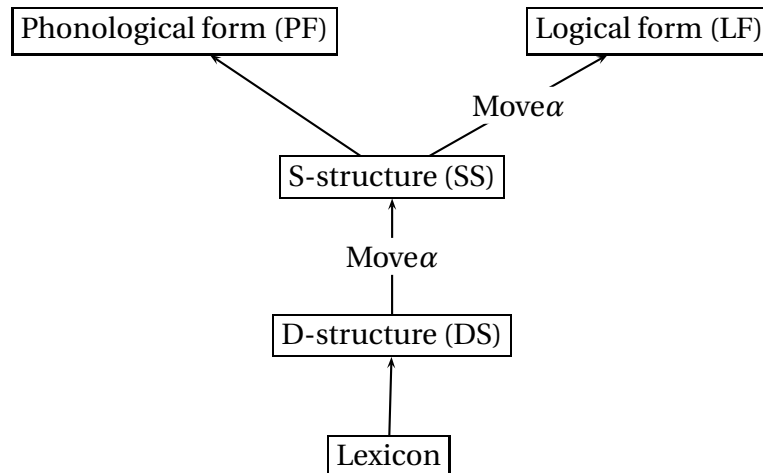
Figure 1.6: The Y-model in GB

In the Y-model, there are several levels of representation: D-structure contains the predicate-argument-structure; at this level, theta-relations are established. D-structure is mapped onto S-structure by the rule Moveα that displaces elements from their base position at D-structure to another position. Moveα also applies from S-structure to LF. Moveα is unrestricted, it can move anything anywhere. As a consequence, a large number of ungrammatical outputs is generated. Hence, constraints on the well-formedness of the surface structure, so-called 'filters', are imposed to filter out ungrammatical outputs. In this respect, there is no ordering of rules, only surface filters.[18]

We can now look at how opaque interactions are derived by enriched representations. I start with the counter-feeding interaction of Wanna-contraction and Wh-movement. Recall that *want* and *to* cannot be contracted in the ECM construction in (52), although they are adjacent on the surface:

(52)   *Interaction of Wanna-contraction and Wh-movement in ECM-constructions*:
   a.   Who do you want to meet Mary?
   b. *Who do you wanna meet Mary?

In a derivational framework, counter-feeding is derived if Wanna-contraction applies before Wh-movement. Since there is a stage in the derivation where *who* intervenes between *want* and *to*, Wanna-contraction cannot apply; movement of the wh-word to the sentence-initial position comes too late to feed contraction. This solution is re-encoded in a representational framework by assuming that non-head positions in a movement chain are occupied by traces. The enriched representation of (52-a) is shown in (53).

(53)   Who$_k$ do you want t$_k$ to meet Mary?

By assumption, the trace can act as an intervener although it is not phonetically realized (Chomsky 1976). Hence, *want* and *to* cannot be contracted because of the intervening trace. There is a problem with this account, however (see Pullum 1979: 147 who attributes the basic observation to Lightfoot 1976): Empty elements do not block Wanna-contraction per se. Recall that Wanna-contraction is possible in control infinitives, com-

---

[18]However, to a certain extent the architecture of the system implies ordering: D-structure precedes S-structure and S-structure precedes PF and LF. In principle, rules on a lower level can still feed or bleed rules on a subsequent level, but not vice-versa. See the discussion on the interaction of pre- and postcyclic rules with cyclic rules in transformational grammar in footnote 14.

pare (54-a) and (54-b). For θ-role reasons, there is an empty category PRO in the subject position of the embedded infinitive (receiving the agent role of *want*). To allow for Wanna-contraction in this example, this empty category must not block contraction, although it is an empty element just like a trace. To this end, Chomsky (1981: 21) stipulates that the abstract features of PRO are invisible to Wanna-contraction, but that the features of traces are visible. If this assumption is made, the counter-feeding interaction of Wanna-contraction and Wh-movement can be derived by the postulation of empty elements.[19]

(54)   *Interaction of Wanna-contraction and Wh-movement in control constructions*:
    a.   Who do you want to meet?
    b.   Who do you wanna meet?
    c.   Who$_k$ do you want PRO to meet t$_k$?

Another example that shows how empty elements can help to explain opaque surface forms is the counter-bleeding interaction of Reflexivization and ISD. In the imperative in (55-a), there is a reflexive pronoun although there is no co-referential c-commanding argument on the surface. Under rule ordering, the underlying subject is deleted only after Reflexivization has applied. In a representational framework this example is handled by postulating an empty pronominal element *pro* in subject position that can act as a binder (cf. Rizzi 1986*a*: 504f.). The structure of (55-a) is thus (55-b). Assuming that *pro* bears (inherent) ϕ-features, 2sg in this case, it follows that Reflexivization applies to the co-referential DP$_{int}$.

(55)   a.   Wash yourself!
    b.   *pro* wash yourself!

Let us now come to a more complex example where its is not sufficient to postulate empty elements, but where, in addition, constraints need to refer to those empty elements. Relevant examples come from "Antecedent Removal" contexts (Pullum 1979; 1992): An anaphor is not in the c-command domain of its antecedent at S-structure anymore because its has been dislocated from its base position at D-structure which was within this domain. Dislocation should bleed binding of the anaphor by its antecedent, but the sentence is grammatical, hence a case of counter-bleeding; see (56) where the anaphor is topicalized:

(56)   Myself, I never buy anything for.

Barss (1986) develops a formal account of such cases that does not involve rule ordering but relies on representational constraints. The basic idea is that movement leaves behind a trace and that a more complex version of Principle A can refer to traces. (57) and (58) are simplified versions of Barss' definitions, which are even more complex (cf. Barss'

---

[19]Note that the issue becomes even more problematic in Minimalism given (i) the copy theory of movement (Chomsky 1995) and (ii) the movement theory of control (Hornstein 2001; Boeckx et al. 2010). In the latter, the controller originates in the embedded clause and moves successive-cyclically to its surface position, leaving copies in intermediate positions, in particular also in the SpecTP position of the embedded infinitive. As a consequence, there would be a copy of *who* between *want* and *to* in (54-c), just as in the ECM construction in (53). So, under the copy theory of movement, there is no difference in the intermediate TP representation of the embedded clause between the ECM and the control case: A copy of *who* intervenes between *want* and *to*. Hence, Wanna contraction is predicted to be blocked or allowed in both cases, depending on whether copies count as intervenors or not, but it is not possible to distinguish the two contexts without further assumptions.

Chain accessibility sequences); these simpler definitions, taken from Müller (2013*b*), are sufficient for present purposes.

(57)   *Principle A (cf. Barss 1986: 104)*:
         At S-structure, an anaphor is chain-bound in its binding domain.

(58)   *Chain-Binding (cf. Barss 1986: 125-137)*:
         α chain-binds β if (a), (b), and (c) hold:
         a. α and β are co-indexed.
         b. α occupies an A-position.
         c.   (i) α c-commands β, or
              (ii) α c-commands a trace of γ, where $\gamma = \beta$ or γ dominates β.

What this version of Principle A essentially does is to say that an anaphor must be locally bound by its antecedent, but it must not be the anaphor itself that fulfills this condition; rather, it suffices if there is one member of the chain of which the anaphor is a part that does so. In the example in (56), it is the trace of the anaphor which is locally bound by the antecedent, see the S-structure representation in (59).

(59)   Myself$_k$ I never buy anything for t$_k$.

The trace encodes the earlier position of the anaphor at D-structure where Principle A was fulfilled. In this way, the trace enables us to understand why the sentence is grammatical although it violates (the standard version) of Principle A (see Fiengo 1977: 45 for an explicit statement that a trace "encodes part of derivational history [...] at the level of surface structure [...]").

        These examples may suffice to show the general direction of the reanalysis of opaque interactions in the GB framework. The question is whether all instances of opacity can be captured in a representational framework with enriched representations. We have seen in the discussion of the counter-feeding interaction of Wanna-contraction and Wh-movement that it is not trivial to derive the difference between ECM- and control constructions by simply making reference to empty elements; it requires a stipulation about which empty elements act as intervenors and which don't. In addition, there are more recent papers on opaque interaction that argue that the respective interaction cannot be reconstructed by simply looking at enriched output representations, see Assmann et al. 2012 on a counter-bleeding interaction and Lechner 2010 on a Duke-of-York derivation. I will not go into the details of the analyses here because we will see the same abstract configuration that Assmann et al. (2012) base their argumentation on in the counter-feeding cases that will be presented in chapter 2; what these data tell us about the derivational / representational dichotomy is discussed in detail in section 5.4.

        Is there a way to derive opacity cases where the postulation of empty elements and constraints on them is not sufficient in a representational framework? Maybe for some, but only at the price of an even more complex analysis. A suggestion along these lines is made in Müller (2013*b*) for the opacity that arises with V2-clauses in languages like German. The standard analysis of V2 is that the verb moves to the C head via T (and other intervening heads). On the surface, however, the structure is opaque (Chomsky 1986*a*, 1991): We would expect ungrammaticality because (at least) the trace of the T head intervenes between the verb in second position and its base position, which violates the *Head Movement Constraint* (HMC) (Travis 1984); but the output is grammatical, hence

this is a case of counter-bleeding.[20] Here the postulation of the trace does not eliminate the problem, rather on the contrary. See Baker (1988); Müller (2013*b*) for a suggestion on how this opacity effect could in principle be derived in a representational framework. The solution invokes (i) massive enrichment with indices in addition to traces, (ii) percolation of indices from a head to the head it moves to, and (iii) a non-trivial extension of the definition of intervention that makes reference to the indices.

To summarize, many instances of opacity can be reanalyzed by enriched representations, but in some cases the price to pay is a very complex analysis that not only makes reference to traces but needs to postulate additional types of empty elements (such as indices) to encode the derivational history; furthermore, basic concepts like Minimality may require more complicated definitions in order to refer to empty elements. The ordering approach to opacity provides a much simpler and more elegant way to capture the facts; it does not require any enrichment of the structure.

1.2.3 Rule interactions in Minimalism

In the beginnings of the 1990, there is a shift from the representational Government and Binding framework back to a derivational model of grammar: the Minimalist Program introduced in Chomsky (1995). In Minimalism, the elementary syntactic building blocks are operations instead of filters or transformations. The two basic operations are Merge and Agree (Chomsky 2000; 2001). Merge is a structure-building operation and Agree is an operation that triggers the exchange of feature values between two elements in the structure. Syntactic structure is built successively in a bottom-up fashion by sequential, alternating applications of Merge and Agree. Hence, there is no DS / SS distinction anymore. The basic operations generate the syntactic output structure, which is the input for PF and LF, step by step. The items that are used during the derivation, the lexical items, are taken from the numeration where they are assembled before the derivation starts. The Minimalist model of grammar is shown in figure 1.7:
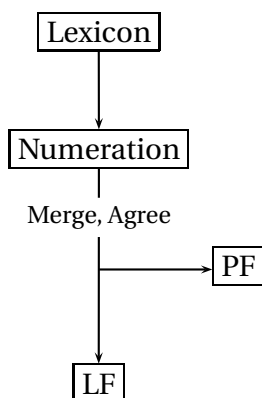
Figure 1.7: The Y-model in Minimalism

Since Minimalism employs a derivational model of grammar with sequential application of operations, the debate about the order of rules / operations has arisen anew. The present thesis contributes to this discussion by investigating transparent and opaque interactions in a particular configuration that has not been studied in much detail so far:

---

[20]The consequence drawn from these facts is that the HMC must not be evaluated at the output structure but rather at each movement step, i.e. it is a derivational rather than a representational or global constraint.

the interactions of Merge and Agree operations triggered by a single head. The central question is again whether we can learn something new from the interaction of elementary operations with respect to the mode of rule application (sequential vs. simultaneous, intrinsic vs. extrinsic). I will ultimately argue that the interactions provide evidence for the *Strict Ordering Hypothesis* (ordered sequential application) and the need for extrinsic ordering. The latter finding is important because the evidence for extrinsic ordering could not have been detected with the interaction of coarse transformations that induce large-scale changes; we can only see it when fine-grained operations interact. In subsection 1.2.3.1 I illustrate why and how Merge and Agree can interact. Subsection 1.2.3.2 shows the abstract configurations I will look at in chapters 2 to 4.

1.2.3.1   The interaction of elementary operations

In this thesis, I will be concerned with the elementary operations Merge and Agree.[21] Merge creates complex structures out of lexical items or previously created structures, and Agree establishes feature transfer between two elements in the structure. More formal definitions are given in (60):

(60)   *The elementary operations Merge and Agree (cf. Chomsky 2000; 2001)*:
    a.   Merge:
        (i)   Merge is a structure-building operation that takes two items α and β and creates a new item γ: $[_\gamma\ \alpha\ \beta]$.
        (ii) Merge is triggered by structure-building features on heads: [•F•].
    b.   Agree:
        (i)   Agree copies feature values from an element α (the goal / the controller) to an element β (the probe / the target): Unvalued features of a probe P are valued by features of the goal G if P and G are in a c-command relation.
        (ii) Agree is triggered by probe features that seek for a value: [∗F:□∗].

Structure-building features are depicted in bullets and probe features are depicted in asterisks in what follows; the empty box in the probe features symbolizes that the feature is lacking a value (the notation is taken from Sternefeld 2006; Heck and Müller 2007). Probes are located on functional heads and XPs provide goal features. Chomsky (2004) further subdivides Merge into *external* and *internal Merge*: In case of external Merge, α and β are both taken from the numeration or the workspace (α is distinct from / outside of β); in case of internal Merge α is not distinct from β, rather α is dominated by β. Internal Merge is thus movement reinterpreted as a special type of Merge. I will use the terms movement and internal Merge interchangeably throughout this thesis. The basic clause structure I will adopt is given in (61):

(61) $[_{CP}$ C $[_{TP}$ T $[_{vP}$ DP$_{ext}$ $[_{v'}$ v $[_{VP}$ DP$_{int}$]]]]]

---

[21]There may be more elementary operations, e.g. deletion and binding, that might interact in interesting ways with one another as well as with Agree and Merge. However, in this thesis I will exclusively focus on the two latter operations here. Note that many of the operations that were considered to constitute separate elementary operations beside Merge and Agree have been reanalysed as involving Agree in recent years: Agree is used to model binding (see among others Reuland 2001; Gallego 2010; Reuland 2011; Rooryck and Van den Wyngaerd 2011; Schäfer 2012), ellipsis (Aelbrecht 2010; Salzmann 2012), and c-selection (Wurmbrand 2013).

The internal argument (DP_{int}) of a transitive verb is merged as the sister of V. The external argument (DP_{ext}) is introduced in the specifier of the functional head v that takes the VP as its complement. Above vP, there are two more functional projections: TP and CP.

Why and how can Merge and Agree interact? The important point is that Agree is structure sensitive: It applies if the probe and the goal stand in a c-command relation. I will assume that either the probe must c-command the goal (downward Agree) or that the goal must c-command the probe (upward Agree); the direction of Agree is fixed language-specifically (see section 2.5.2.1 for discussion). C-command is defined in terms of sisterhood (cf. the derivational definition of c-command in Epstein et al. 1998: 10):

(62) *C-command*:
     An element α c-commands an element β if a. or b. holds:
     a. β is the sister of α.
     b. β is included in the sister of α.

The sisterhood relation is established and altered by Merge and thereby c-command relations are established or destroyed. External and internal Merge create new sisterhood relations between the merged XP and the head that triggers Merge of XP. In addition, internal Merge of XP destroys the sisterhood relation between XP and a head that existed before XP was moved to a higher position (assuming for the time being that movement of α does not leave behind any kind of representation in its base position). Hence, Merge can potentially feed or bleed Agree relations.

Let us look at an example for the creation of new c-command relations by Merge. Bruening (2005) is concerned with a feeding interaction of Merge and downward Agree in Passamaquoddy (Algonquian): In this language, a verbal prefix indicates ϕ-agreement with the most prominent argument (prominence is traditionally defined in terms of a participant/animacy hierarchy). If, in a transitive context, the subject (DP_{ext}) is more prominent than the object (DP_{int}), we have a direct context; if the object is more prominent than the subject, we have an inverse context. As Bruening argues, prominence is based solely on hierarchical positions in the syntax and not on the participant hierarchy. A DP is more prominent if it is the structurally highest DP in a given domain. First, he shows that the subject is always projected above the object; evidence comes from scope and variable binding data in the direct, which indicate that the subject c-commands the object, cf. (63-a).

(63) a. $[_{CP}\ C_{[*\phi:\square*]}\ [_{TP}\ [_{vP}\ DP_{ext}\ [_{v'}\ v\ [_{VP}\ V\ DP_{int}\ ]]]]]$ *direct*
     b. $[_{CP}\ C_{[*\phi:\square*]}\ [_{TP}\ [_{vP}\ DP_{int_k}\ [_{v'}\ DP_{ext}\ [_{v'}\ v\ [_{VP}\ V\ \underline{\ \ \ }_k\ ]]]]]]$ *inverse*

Interestingly, in the inverse where the object is more prominent, these tests suggest that the reverse holds, i.e. that the object c-commands the subject. Therefore, Bruening assumes that the object can be moved to the outer specifier of vP, thereby creating a new c-command relation, cf. (63-b). In its landing site, the moved object DP c-commands the subject DP; in addition, the subject ceases to c-command the object. The verbal prefix is the morphological realization of an Agree relation between a ϕ-probe located on C and the closest DP in its c-command domain. The fact that the prefix shows agreement with the most prominent argument is then simply a minimality effect: The probe agrees with the structurally highest DP and the more prominent an argument, the higher it is in the structure. In the default case, the closest and thus the most prominent argument is the subject DP because it is base-merged above the object (cf. (63-a)). If, however, object movement to the outer SpecvP has applied, the object is closer to the probe on C than the

subject (cf. (63-b)). Hence, internal Merge of the object to SpecvP feeds ϕ-Agree between C and the object because after movement, the object asymmetrically c-commands the subject. This analysis presupposes that vP-internal movement to SpecvP precedes Agree initiated by C. This order of operations follows from the Cycle if (at least) every XP is a cyclic domain (Williams 1974; McCawley 1984; 1988): DP-movement applies in the vP and Agree within CP. Since CP dominates vP, operations in the vP must precede those in CP.[22]

What I will be concerned with in this thesis are interactions of Merge and Agree in a smaller domain, viz. within a single cycle (even under the strictest definition of cyclic domain proposed by McCawley 1984; 1988, see above). These are cases where the two basic operations are triggered by one and the same head H. Indeed, it is assumed in the Minimalist literature that some heads trigger more than one operation. For example, a dual role is attributed to the head v: It triggers external Merge of the external argument DP and downward Agree with the internal argument, where Agree results e.g. in ϕ-feature transfer from DP_{int} to v or case assignment of v to DP_{int}. The same holds for the T head in English: It triggers internal Merge of a DP to its specifier (the EPP-property) and downward Agree, i.e. ϕ-feature transfer from a DP plus case assignment to a DP (see Chomsky 2000 et seq.). v and T thus have two operation-inducing features, as abstractly shown for the head H in (64):

(64)   H { [•F•], [∗F:□∗] }

As pointed out by Müller (2009), this leads to a conflict at the stage of the derivation where H is merged with its complement (triggered by another structure-building feature not listed in (64)) if the *Earliness Principle* (Pesetsky 1989) holds. This principle demands that an operation applies as soon as its context is met. But if the context for *both* operations is met as soon as H is merged into the structure, both need to apply immediately. This situation arises e.g. with T in English: Once it is merged with vP, it can trigger Merge of an XP or Agree with a DP in its c-command domain, viz. with DP_{ext} or DP_{int}. The only way to adhere to the *Earliness Principle* would be to apply Merge and Agree simultaneously. Indeed, Chomsky (2008) proposes that operations within a certain domain, the phase, apply simultaneously; irrespective of the exact definition of phases (see chapter 2), operations triggered by a single head will always be within the same domain. Recall that simultaneous application can only produce opaque interactions but not transparent interactions (see the discussion in section 1.2.1.2). It has been observed, however, that there are instances of feeding and bleeding between Merge and downward Agree when triggered by a single head. Hence, the two operations cannot (always) apply simultaneously; rather, sequential ordering of the two operations is required. Let us look at an example of bleeding and counter-bleeding interactions of downward Agree and internal Merge from the literature, as abstractly depicted in (65):

---

[22]See Müller (1998) for an example of an *opaque* interaction that also follows from the Cycle: In remnant movement configurations, an XP from which an element E is subextracted surfaces in a derived position, cf. (i):

(i) $[_{ZP} [_{XP} \ldots \underline{}_n \ldots]_k [_{Z'} Z \ldots [_{YP} E_n [_{Y'} Y [_{WP} W \underline{}_k]]]]]$

Usually, this evokes a violation of the *Freezing Principle* according to which moved phrases are islands. Movement of the XP should bleed subextraction of E from XP. However, the result is grammatical, an instance of counter-bleeding (*Anti-Freezing*). In contrast to classic freezing effects, the subextracted element E surfaces below the XP from which it is subextracted, more precisely, it is part of a cyclic domain (= YP) that is included in the domain of the XP (= ZP). The Cycle thus predicts that subextraction from XP applies before the XP is moved. This automatically explains the Anti-Freezing effect: At the point when subextraction applies, the XP is still in its base position; it is frozen only by a subsequent movement step.
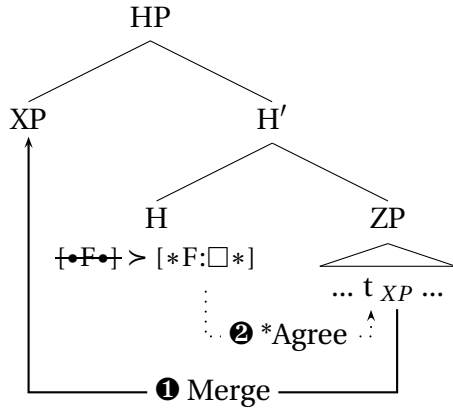
(65)   *Merge > Agree*   (66)   *Agree > Merge*
        (bleeding):   (counter-bleeding):
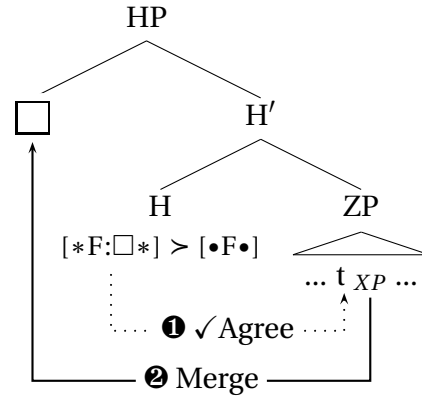
In each case we have a head H that triggers Agree and Merge. The phrase XP that is sup-posed to be merged with H′ is included in the complement of H (hence, a case of internal Merge). Under downward Agree, the probe must c-command the goal to be valued (but not vice versa). Assume that the goal for the probe on H is XP. If internal Merge of XP to SpecHP applies first (discharging the structure-building feature) as in (66-a), the probe on H that seeks for a value in a second step cannot find a goal since the potential goal XP is no longer in the c-command domain of H, and hence, the probe feature is not discharged by valuation.[23] Put differently, Merge before Agree (written as Merge > Agree) bleeds Agree; XP movement applies too early and thereby destroys the context for Agree.

If, however, Agree applies before Merge (written as Agree > Merge), as shown in (66-b), the XP is still in the c-command domain of H when the probe initiates Agree and the probe can be valued. XP moves out of the c-command domain of H only afterwards, i.e. movement applies too late to bleed Agree, a case of counter-bleeding: On the surface, it is not clear how Agree between H and XP could be established because the XP is not in the right structural position to be a goal for H. (or: the generalization that only XPs in the c-command domain of a probe can trigger Agree is not surface-true).

The counter-bleeding configuration in (66-b) corresponds to the situation that arises in English with subject-verb agreement (although it is virtually never explicitly char-acterized as a case of opacity in the literature). Under the standard Minimalist analysis of this phenomenon, the T head triggers internal Merge of the structurally highest DP to SpecTP (the EPP property) and ϕ-Agree with the subject. On the surface, the subject DP is in SpecTP and thus not in the c-command domain of T, nevertheless, Agree is success-fully established. Indeed, it is (mostly tacitly) assumed in the literature that Agree between the ϕ-probe on T and the structurally highest DP applies before movement of that DP to SpecTP in order to get the right result.[24]

Analyses where Agree and Merge are ordered on a single head have been proposed to an increasing degree in recent years: Bruening (2005); van Koppen (2005); Halpert (2012); Kalin and van Urk (2012); Richards (2013) and van Urk (in prep.) (as cited in

---

[23]This presupposes that the probe cannot Agree with the trace or copy in the base position of XP – if movement leaves behind an element at all. This will be further discussed in section 2.

[24]One might propose an alternative analysis of the English data that does not involve opacity: If Agree applies upwards between a head and its specifier, then it is not surprising that an Agree relation can be es-tablished between T and the EPP-moved subject DP in SpecTP. This can be read off of the surface represen-tation. However, there is evidence that Agree in English applies downwards, into the c-command domain of the probe: In constructions with an expletive in SpecTP, the verb agrees with the logical subject DP in the complement of T. This is illustrated in (i) for the 3rd person singular vs. plural forms of the verb *be*:

van Urk and Richards 2013: 6, fn.3) derive bleeding effects with the order Merge before Agree. Anand and Nevins (2005) and Asarina (2011) are concerned with feeding effects: Early internal Merge of an XP to SpecHP that intervenes for Agree between H and a lower YP feeds subsequent Agree H – YP. In all of these studies, Agree applies downwards, viz. the probe must c-command the goal. I will be concerned with such configurations in chapter 3. Feeding interactions that arise under the reverse direction of Agree, viz. if the goal c-commands the probe, are the subject of chapter 2.

All of the papers cited above are concerned with interactions of Merge and Agree in a single language. Therefore, the authors make no claim about whether the order of operations is universally fixed or variable (although they all assume that Merge applies before Agree to derive the bleeding / feeding effects). But since the standard analysis of English subject-verb agreement requires the reverse order Agree before Merge, we can conclude that the order of elementary operations on a single head differs between languages. That means that there seems to be a parameter for the order of Agree and Merge that is fixed language-specifically. Indeed, there are also a few studies that have explicitly argued for such a parameter based on cross-linguistic variation: Müller (2009) argues that the variable order of Merge and Agree on v derives the ergative (Merge > Agree) vs. accusative (Agree > Merge) argument encoding pattern (see Heck and Müller 2007 for further consequences of reordering Merge and Agree on other functional heads for argument encoding). Based on this work, Lahne (2008*a*) and Assmann et al. (2012) show that if the order of Merge and Agree on v is the same as on T, further properties that distinguish ergative from accusative languages can be derived (Lahne 2008*a* is concerned with word order restrictions in ergative languages that are not found in accusative languages, and Assmann et al. 2012 derive the ban on Ā-movement of the external argument of a transitive verb that holds in many ergative languages but not in accusative languages). To conclude, the state of the art in the literature is that if a head triggers more than one operation, (a) the operation-inducing features on that head must be ordered – at least if there are transparent interactions between the two, and (b) the order is either Merge before Agree or Agree before Merge and this is fixed in a language-specific manner.

1.2.3.2   Rule interactions in this thesis

In this thesis, I will be concerned with opacity effects that arise if Merge and Agree are triggered by a single head. However, the opacity is of a different type than the one in English, where it is sufficient to order internal Merge relative to Agree. Rather, I look at phenomena that suggest that both Merge > Agree and Agree > Merge holds on a single head:

(67)   *Abstract pattern of interactions studied in this thesis*:
         Merge > Agree > Merge

If Merge is a uniform operation, we cannot produce a symmetric order as in (67); either Merge applies before Agree or afterwards, but not both at the same time. Hence, there is opacity: Within a single language, Merge of an XP to SpecHP sometimes feeds or bleeds

---

(i)   *Downward Agree in English*:
         a.   There is a man in the garden.
         b.   There are two men in the garden.

In addition, there is evidence that the associate DP *two man* does not move to SpecTP at any level of representation (where it could enter into upward Agree with T): The postverbal DP cannot take scope over *seem* and cannot bind an anaphor or license a negative polarity item outside of VP; see Lasnik (1999*b*); Nevins (2004).

Agree initiated by H, but sometimes it has the opposite effect in exactly the same position. We can immediately draw the following conclusion: It is too simplistic to say that the order of Merge and Agree is fixed in a language-specific fashion because in some languages there is evidence that both orders are required. This conflict can be resolved if the Merge operation that applies before Agree and the Merge operation that applies after Agree are indeed two different types of Merge. Thus, Agree is not simply ordered with respect to Merge, but rather to *subtypes of Merge*:

(68)   *Abstract pattern of interaction studied in this thesis – revised version*:
        $Merge_1$ > Agree > $Merge_2$

I thus argue for a more fine-grained approach to elementary operations. I will implement this order of operations by ordering of operation-inducing features on a head. Consequently, $Merge_1$ and $Merge_2$ must have distinct triggers. Thus, depending on the type of Merge trigger, Merge of XP to SpecHP can apply either before or after Agree initiated by H, giving rise to transparent or opaque interactions of the two operations. By looking at configurations as in (68) where Agree "splits up" Merge operations, we can investigate which subtypes of Merge and thus which operation-inducing features need to be distinguished. Indeed, the split between types of Merge is not random; we find the same splits across languages over and over again. Chapters 2 and 3 are concerned with a split between intermediate and final movement steps in a movement chain; chapter 4 extends the general picture to other subtypes such as external vs. internal Merge and different types of final and intermediate internal Merge. As will be discussed in detail in chapter 5, the present analysis presupposes sequential application of rules in a fixed order; furthermore, it requires extrinsic ordering – neither the Cycle nor any other principle can predict all orders that are motivated by cross-linguistic variation in morphological reflexes of movement.

I will conclude with a historical remark: With the return to a derivational model in Minimalism, one would expect that rule interactions are at the center-stage of research again, especially opacity, which is problematic for theories with output-oriented constraints. Interactions of operations could provide evidence for a model where structure unfolds step by step through sequential application of operations. But surprisingly, this is not the case. There are quite a number of studies of interaction of operations across components where the order of operations follows from the architecture of the grammar.[25] There are, however, only a few studies on the interaction of elementary operations, especially of the type looked at here, where the order of operations does not immediately follow from the Cycle (see the references on page 30). The most comprehensive study of this type is presented in Řezáč (2004); however, he only looks at transparent interactions. It is even more astonishing that rule interactions are not prominent in Minimalist syntax when we look at the phonological literature: Within the framework of optimality-theory, opacity is currently an important topic because it provides a challenge for global optimization with simultaneous application of all constraints. In phonology, opacity phenomena have been causing a reintegration of derivations into an OT-system as e.g. in Stratal OT or Harmonic Serialism with local optimization (see footnote 17 for discussion and references, as well as Heck and Müller (2003) for the same conclusion based on in-

---

[25]See e.g. Embick 2000; Sauerland and Elbourne 2002; Lechner 2010; Watanabe 2012 for illustration that some opacity effects are a consequence of the fact that an operation O_1 that could in principle feed or bleed another operation O_2 comes too late to do so because O_1 is a postsyntactic operation (at LF or PF) and O_2 is a syntactic operation. Frampton (2009); Arregi and Nevins (2012) and Salzmann (2013) show that opacity effects at the morphosyntax / phonology interface provide evidence for a certain order of PF operations.

teractions of OT-constraints in syntax). The present thesis aims at closing the gap in the syntactic literature by investigating in particular opaque interactions and their relevance for the derivational / representational distinction. I will argue for a strictly derivational model of syntax that requires both extrinsic and intrinsic ordering of operations.

Chapter 2

Interactions of internal Merge and upward Agree

In this chapter I investigate the interaction of internal Merge (movement) and upward Agree triggered by a single head. I argue that opaque interactions of these operations provide evidence that Merge is not a uniform operation: Rather, we need to distinguish between Merge operations that trigger final and those that trigger intermediate movement steps because they can apply at different points of the derivation relative to upward Agree. The empirical basis for this claim are reflexes of successive cyclic movement that involve Agree between a head H and an XP moved to SpecHP. Section 2.1 introduces the concept and motivation for successive-cyclic movement. In section 2.2 I present the abstract patterns of reflexes of long-distance movement found cross-linguistically; examples for each pattern are provided in section 2.3. In section 2.4, I discuss previous analyses of the patterns and argue that two of them involve opacity: Some movement steps feed agreement while others counter-feed it. Section 2.5 develops an account of the interaction of agreement and movement in terms of ordering of elementary operations. I show that the ordering analysis is superior to previous approaches for both empirical and conceptual reasons.

2.1   On the notion of successive-cyclicity

In this section, I briefly summarize the motivation for the concept of successive-cyclicity as well as the question which positions serve as intermediate landing site. I continue with an overview of the empirical evidence for the cyclic nature of movement, viz. reflexes of movement in positions between the base position and the landing site of a moved XP.

2.1.1   Evidence for successive-cyclic movement

Ā-movement can in principle span an unlimited number of clauses, it is an unbounded dependency (cf. Ross 1967), cf. the example in (1) with movement across two clause-boundaries (the base position of the Ā-moved phrase is indicated by underline with the index k):

(1)   [$_{CP}$ What$_k$ [$_{IP}$ do you think [$_{CP}$ that [$_{IP}$ Mary believes [$_{CP}$ that [$_{IP}$ Bill bought ___$_k$]]]]]] ?

Chomsky (1973) proposes that such a long-distance dependency should be broken up into a sequence of more local steps, i.e. that movement applies successive-cyclically. He

assumes that movement makes a stop-over in the Comp position (corresponding to current SpecCP) of a clause before it continues moving to the next higher clause (Comp-to-Comp movement). This is motivated on the basis of the *Subjacency Condition* and wh-island effects:

(2)   *Subjacency Condition (Chomsky 1977)*:
        Movement must not cross two bounding nodes.

In English, the bounding nodes are NP and IP (DP and TP in current terminology). The *Wh-Island Condition* formulates the observation that wh-movement from a clause with a wh-element in SpecCP or C is ungrammatical in many languages. (3) provides an example of a *Wh-Island Condition* violation from Boeckx and Lasnik (2006: 150).[1]

(3)   *?[CP What_k [IP did Sue wonder [CP where [IP Bill bought ___k]]]] ?

Chomsky analyses this effect as a Subjacency violation: The wh-phrase moves from its base position directly to its scope position in the matrix clause and thereby crosses two bounding nodes (IP in the embedded and the matrix clause). But then why is the example in (1) grammatical where even more IP bounding nodes intervene between the base position and the landing site of the wh-phrase? The answer is that the wh-phrase does not move in one fell swoop but rather moves successive-cyclically through every Comp (SpecCP) position. (4) illustrates this for the example in (2). Each movement step crosses only one bounding node, viz. the IP in the respective CP, which is in accordance with *Subjacency*.

(4)   *Comp-to-Comp movement*:
        [CP What_k do [TP you think [CP ___k that [TP Mary believes [CP ___k that [TP Bill bought ___k]]]]]] ?

I will call the position (here Comp) where the wh-phrase surfaces its *final landing* site; other positions where it only makes a stop-over on its way to the final landing site are called *intermediate landing sites*.

     However, successive-cyclic movement is not available for the wh-phrase in (3): It cannot stop in Comp of the embedded clause because, by assumption, there is only one such position and this is already occupied by a wh-phrase. So *what* is forced to skip the embedded SpecCP and to move in one fell swoop, causing a violation of Subjacency. This derives the *Wh-Island Condition* violation in (3).

     Since the introduction of successive-cyclic movement, one of the central questions is what positions serve as intermediate landing sites. Whereas Chomsky (1973; 1981) postulates that only SpecCP is an intermediate landing site, Chomsky (1986a) assumes that Ā-movement also makes a stop-over in SpecVP in addition to SpecCP. For our initial example in (1) this would look as follows:

(5)   *Stop-overs in SpecVP and SpecCP*:
        [CP What_k do you [VP ___k think [CP ___k that Mary [VP ___k believes [CP ___k that Bill [VP ___k bought ___k]]]]]] ?

Sportiche (1989); Manzini (1994) and Takahashi (1994) propose even more local movement steps: Successive-cyclic movement makes a stop-over in the specifier of every

---

[1]In English, extraction of wh-arguments from a wh-island is generally considered to be less degraded than extraction of a wh-adjunct, therefore the question mark in (3).

phrase along the path of movement. If we assume that there are three core phrases CP, TP, and VP, we get the result in (6):

(6)   *Stop-overs in every SpecXP*:
        [$_{CP}$ What$_k$ [$_{TP}$ do you [$_{VP}$ ___$_k$ think [$_{CP}$ ___$_k$ that [$_{TP}$ Mary [$_{VP}$ ___$_k$ believes [$_{CP}$ ___$_k$ that [$_{TP}$ Bill [$_{VP}$ ___$_k$ bought ___$_1$]]]]]]]]] ?

In recent Minimalism, successive-cyclic movement is enforced by the *Phase Impenetrability Condition* (PIC):

(7)   *PIC (Chomsky 2000: 108)*:
        In a phase α with the head H, the domain of H is not accessible to operations outside α; only H and its edge are accessible to such operations.

The domain of a phase HP is the complement of a phase head H; the edge of HP consists of H and elements in the specifier of HP or adjuncts to HP. Chomsky (2000; 2001) assumes that vP and CP are phases. So, if a phrase XP is base-generated in the vP and has to move out of vP, e.g. to SpecCP, it needs to move through the specifier of vP. If XP did not move to this position, it would not be accessible for the C head because it would be trapped in the complement of v. The number of stop-overs of course depends on what XPs constitute phases. If vP and CP are phases, the outcome is similar to (5): Successive-cyclic movement goes through every SpecvP and SpecCP along the path of movement (see also Heck and Müller 2000). There is an ongoing debate whether some other (but not all) XPs beside vP and CP are phases; see Heck and Zimmermann (2004); Svenonius (2004); Kramer (2010) for arguments that DP is a phase, Abels (2003) on the phase status of PP, and Richards (2004; 2011); Chomsky (2005) on the phase status of TP (at least as an option in some language). It has also been proposed that every phrase is a phase, and hence that successive-cyclic movement must go through the specifier of every phrase along the path of movement. This amounts to the strategy exemplified example in (6), as suggested by Sportiche (1989); Manzini (1994) and Takahashi (1994); see, among others, Epstein and Seely 2002; Bošković 2002; 2007*b*; Boeckx 2003; Müller 2004; 2011; Chomsky 2008 for more recent proposals along these lines making reference to the PIC.[2]

An even more local approach would be to have the information that an element is extracted available on every projection, not only on the XP level. This is e.g. the case with *Slash Feature Percolation* in GPSG and HPSG (Gazdar 1981; Gazdar et al. 1985; Pollard and Sag 1994).

Abels (2003) classifies analyses of successive-cyclic movement with respect to the question whether all categories crossed by movement are treated alike or not. If all XPs or projections are targeted by intermediate movement steps (or if none of them is as in unbounded movement of the type in (1)), the path is uniform. If only designated categories provide stop-over positions for intermediate movement steps (e.g. vP and CP, but not TP), the path is punctuated.

The question which positions serve as intermediate landing site is an empirical one. Indeed, a great deal of empirical evidence has been accumulated since the 1970s for the assumption that movement applies successive-cyclically. The evidence comes from morphophonological, syntactic and semantic reflexes of movement that appear either in intermediate positions along the path of movement (syntactic and semantic reflexes), or

---

[2]In some analyses, the phase status of an XP is not fixed. Rather, it can be lost or acquired in the course of the derivation (see Gallego 2005; den Dikken 2007); or the phase status depends on the syntactic context (see Svenonius 2001; Bobaljik and Wurmbrand 2005).

on heads whose specifier the moving element uses as an intermediate landing site (morphophonological reflexes). In this thesis, I will exclusively be concerned with morphological as well as morphosyntactic reflexes of the latter type, i.e. where the reflex does not occur in the intermediate position itself, but rather on a head in the vicinity of this position. The reason for this choice is that – following standard assumptions – I take reflexes of movement to be ultimately the result of an agreement relation between a head and an XP moved to its specifier (cf. among many others Torrego 1984; Henry 1995 on inversion in Spanish and Belfast English, respectively, Rizzi 1990 on the *que/qui*-alternation in French, Collins 1993 on pronoun choice in Ewe, Chung and McCloskey 1987; Kinyalolo 1991; Schneider-Zioga 1995; Torrence 2012 on complementizer agreement in Bantu languages, Watanabe 1996 on wh-agreement, as well as Chung and McCloskey 1987; McCloskey 1990; Chung 1998 on complementizer agreement in Irish and Chamorro). This agreement relation involves feature transmission from XP to H (or alternatively, checking of features on H), cf. (8).

(8)
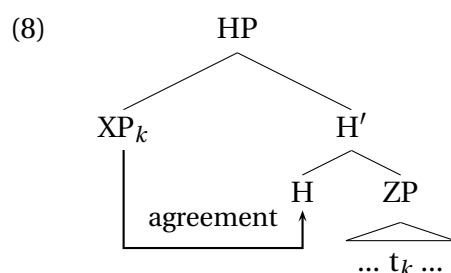
The feature transmitted to H (or checked on H) is then realized on H; the exponent is the reflex of movement.[3,4] In my analysis of reflexes of movement, agreement will be modeled by the operation Agree and movement by the operation Merge. In this way I can investigate the interaction of Merge and Agree on the basis of reflexes of movement. Note that the agreement relation between XP and H in (8) will involve upward Agree because the controller (XP) c-commands the target (H) of feature transmission. For this reason, I will focus on the interactions of Merge and upward Agree in this chapter; interactions with downward Agree will be the topic of chapter 3.

As for the empirical evidence for the positions that serve as intermediate landing sites, it is uncontroversial that SpecCP and SpecvP constitute such positions. We will see a number of examples for morphological reflexes on C and v in the data. For other heads, the issue is debated. In the data I look at, we will find evidence from reflexes of movement for the hypothesis that SpecTP and Spec*n*P (*n* being a functional head in the extended projection of the noun) serve as intermediate landing sites, too. For the sake of concreteness, I will assume that all phrases are phases and that movement must thus pass through the specifier of every phrase along the path of movement.

---

[3]Often, the nature of the relation between the head and its specifier is not made explicit in the literature. If it is, there are many different implementations, transfer of morphosyntactic features (agreement) being just one of them. Schneider-Zioga (1995) discusses the mechanism of Comp-indexing: The head gets the same index as the phrase moved to its specifier (similarly to what Finer (1985) proposes for the modeling of switch-reference) and thereby inherits the relevant features of the XP that are realized morphologically. There are also approaches where the relation is not one of transfer but rather involves binding, cf. the Ā-binding approaches to reflexes of movement that will be discussed in section 2.4.5.1 (Epée 1975; Clements 1984*a*; Chung and McCloskey 1987; Haïk 1990; Schneider-Zioga 1995).

[4]See Müller (2011) for a completely different view on the phenomenon: In his system, reflexes of XP-movement are not reflexes, viz. not the result of movement; rather, the movement-related morphology enables movement of XP (by opening up a barrier in which the XP would otherwise be trapped).

2.1.2 Reflexes that are not treated in this thesis

As pointed out above, I will only look at reflexes of movement that manifest themselves on a head H adjacent to the (intermediate or final landing) site of an XP, and I will ignore reflexes that solely occur in the intermediate position SpecHP itself. Since this thesis is concerned with the operation Agree that copies features from one element to another, viz. from the moving XP to the head H to whose specifier XP moves, we can only learn something about the interaction of Agree and Merge when we see the realization of Agree on the head. When a moving XP has a reflex in SpecHP but no visible effect on the morphological form of H, we do not know whether H enters into Agree with the XP in its specifier. Nevertheless, reflexes of the latter type (occurring in SpecHP) also provide strong evidence for the successive-cyclic nature of movement. For the sake of completeness, I briefly summarize them in sections 2.1.2.1 and 2.1.2.2 (see also Boeckx 2008*b*; Lahne 2008*b* for an overview).

2.1.2.1 Morphosyntactic reflexes in specifiers

In some languages, a wh-phrase cannot only be pronounced in its scope position, but in addition it can occur in positions below it. This phenomenon is known as *wh-copying* (see du Plessis 1977; McDaniel 1989; Fanselow and Mahajan 2000; Höhle 2000; Felser 2004; Nunes 2004 for data). Examples are given in (9) from Afrikaans and German. Under long wh-movement from an embedded clause, the wh-phrase is not only realized in its scope position, viz. SpecCP of the matrix clause, but also in the specifier of the embedded clause(s). Here and in what follows, I will set the reflex of movement, i.e. the element in the intermediate position, as well as the element in the final position in bold.

(9)   *Wh-copying in Afrikaans and German*:

    a. [$_{CP}$ **waaroor**$_k$ dink jy   [$_{CP}$ **waaroor**$_k$ dink die bure   [$_{CP}$ **waaroor**$_k$
            whereabout think you   whereabout think the neighbors   whereabout
            stry   ons ___$_k$ die meeste]]]
            argue we   the most
            'What do you think the neighbors think we are arguing about the most?'
                                            Afrikaans, du Plessis (1977: 725)

    b. [$_{CP}$ **Wie**$_k$ glaubst du, [$_{CP}$ **wie**$_k$ sie das ___$_k$ gelöst hat]]
            how believe you   how she that   solved has
            'How do you believe that she has solved that?'
                             German, Fanselow and Mahajan (2000: 220)

Assuming that the wh-phrases in the embedded clauses are the spell-out of copies of the wh-phrase in the matrix clause left by movement, these data provide evidence for successive-cyclic movement through every SpecCP along the path of movement.

    The same conclusion can be drawn from *partial wh-movement* (see van Riemsdijk 1982; Cole 1982; McDaniel 1989; Horvath 1997; Müller 1999; Sabel 2000; Bruening 2006 for discussion): The wh-phrase does not (overtly) move to its scope position, but only to the specifier of an intermediate CP, cf. the example from Indonesian in (10-a) We will see numerous examples of this type in section 2.3.

(10)   *Partial wh-movement (+ scope marking)*:

    a. [$_{CP}$ Tom meng-harap [$_{CP}$ **apa**$_k$ yang Mary beli   ___$_k$]]
        Tom TRANS-expect   what FOC Mary bought
        'What does Tom expect Mary bought?'   *Indonesian, Saddy (1991: 192)*

    b. [$_{CP}$ Was meinst du, [$_{CP}$ **wem**$_k$ Peter   die Leute   ___$_k$ vorgestellt
        SCO think   you   who.DAT Peter.NOM the people.ACC   introduced
        hat]]
        has
        'To whom do you think Peter introduced the people?'

                                              German, Sabel (2000: 411)

In some languages, the scope position of the wh-phrase is marked by a scope marker (glossed as 'SCO'), an expletive wh-element that is, however, not phonologically identical to the "real" wh-phrase in the intermediate SpecCP position, see the German example in (10-b) with the scope marker *was*.

    Another effect of successive-cyclic movement in intermediate positions is that a part of the moved phrase is stranded there.[5] Here I will exemplify this for quantifier float (see du Plessis 1977; Barbiers 2002 for stranding of material other than quantifiers). Under Sportiche's (1988) seminal analysis, the quantifier Q and the NP it is associated with initially form a constituent: [$_{NP}$ Q NP]. On the surface, Q and its associate NP can be separated, however. Quantifier floating comes about when the NP moves alone and leaves behind the quantifier, i.e. quantifier floating is in fact quantifier stranding. Crucially, there are languages where the quantifier strands in intermediate positions along the path of movement. McCloskey (2000) describes this pattern for a variety of English. West Ulster English generally allows the stranding of quantifiers associated with a wh-phrase, cf. (11):.

(11)   *Quantifier Float in West Ulster English (McCloskey 2000: 58)*:

    a. [What all]$_k$ did you get ___$_k$ for Christmas
    b. [What]$_k$ did you get ___$_k$ **all** for Christmas

Interestingly, under long wh-movement, the quantifier cannot only be moved along with the wh-phrase to its scope position (cf. (12-a)), or be stranded in the base position of the wh-phrase in the embedded clause (cf. (12-b)), but it can also be stranded in an intermediate position, viz. SpecCP of the embedded clause (cf. (12-c)).

(12)   *Quantifier Float under long wh-movement (McCloskey 2000: 61)*:

    a. [$_{CP}$ [What all]$_k$ did he say [$_{CP}$ (that) he wanted ___$_k$]]
    b. [$_{CP}$ What$_k$ did he say [$_{CP}$ (that) he wanted ___$_k$ **all**]]
    c. [$_{CP}$ What$_k$ did he say [$_{CP}$ **all** (that) he wanted ___$_k$]]

Assuming with Sportiche (1988) that quantifiers occupy the positions the associated NP can occupy, McCloskey (2000) takes these facts as evidence for successive-cyclic movement of the wh-phrase through the intermediate SpecCP.[6]

    In all of the cases summarized above, there is a morphosyntactic reflex of movement in a specifier position, but no reflex on the head that project this position. Hence, there is no evidence for an Agree relation between the moving phrase and a head along the path of movement. The same is true for semantic reflexes of movement, to which I will turn now.

---

[5]See Barbiers et al. 2009 for a stranding analysis of partial wh-movement.

[6]In addition, McCloskey (2000) assumes that the wh-word undergoes DP-internal movement from the complement position of the D-element *all* to SpecDP. For ease of exposition, I ignored this movement step in the examples.

2.1.2.2 Semantic reflexes

Semantic evidence for successive-cyclic movement comes from reconstruction to intermediate landing site. I will mention a few illustrative examples. For more data and discussion see among others Lebeaux (1990); Fox (2000); Abels (2003; 2012); Boeckx and Grohmann (2007); Agüero-Bautista (2007).

Castillo et al. (2009) present the following argument for successive-cyclic A-movement through SpecTP: A reflexive pronoun can be bound by an NP that on the surface does not fulfill the requirements to qualify as an antecedent. In (13-a) the antecedent *John* is not the closest binder for the reflexive; the complement of the preposition, *Mary*, is able to c-command out of the PP and thus would be a closer binder. Nevertheless, the sentence is grammatical with the binding indicated (the subscripts *i* and *j* are referential indices).

(13)   *Evidence for successive-cyclic A-movement (Castillo et al. 2009: 93)*:
      a. John_i seems to Mary_j to appear to himself_i to be happy.
      b. $[_{TP_3}$ [John_i]$_k$ seems to Mary_j $[_{TP_2}$ ___$_k$ to appear to himself_i $[_{TP_1}$ ___$_k$ to be ___$_k$ happy]]]

This can be explained if intermediate landing sites are postulated: If *John* moves successive-cyclically (at least) through every SpecTP, as indicated in (13-b), it is in the right structural position to bind the anaphor in the specifier of TP_2. Technically, binding into intermediate positions (or the base position) can be established by reconstruction of the moved phrase XP into those positions.[7]

The mirror image obtains when it is the anaphor instead of the antecedent that, on the surface, is not in the right structural position to be bound. In (14-a), the anaphor is not c-commanded by a coreferent NP and hence, the sentence should be ungrammatical. However, it is possible for the NP *John* to bind the anaphor.[8] This can be explained if the NP containing the anaphor moves successive-cyclically (at least) through every SpecCP on the way to its final landing site, cf. (14-b) (therefore these anaphors are called pit-stop reflexives, see also Boeckx 2008*b*; Abels and Bentzen 2011). The reflexive pronoun is locally bound by the coreferent NP *John* in the intermediate landing site $SpecCP_1$.

(14)   *Evidence for successive-cyclic Ā-movement: pit-stop reflexives (Barss 1986: 25)*:
      a. $[_{CP_1}$ $[_{NP}$ Which pictures of himself_i] did John_i think Fred liked]
      b. $[_{CP_2}$ $[_{NP}$ Which pictures of $\text{himself}_i]_k$ did John_i think $[_{CP_1}$ ___$_k$ Fred liked ___$_k$]]

In the examples discussed so far, we have seen evidence from reflexes in specifier positions that SpecCP and SpecTP serve as intermediate landing sites. I will now turn to the topic of this chapter: reflexes on heads in the vicinity of final and intermediate landing sites of a moved XP.

---

[7]Several implementations of reconstruction have been proposed. Either, the XP is literally moved downward into an intermediate position or an intermediate representation of the XP such as a trace or a copy is taken into account when computing the binding possibilities.

[8]The anaphor can also be bound by *Fred*. The reason is that *Fred* locally binds the anaphor if the NP containing the anaphor is in its base position, prior to Ā-movement. What is crucial for the argument for successive-cyclic movement, however, is that it is possible that *John* binds the reflexive pronoun. That this is an option cannot be read off of the surface structure because *John* does not c-command the anaphor there. It can also not be explained by taking into account the base position of the moved phrase because *Fred* is a closer binder than *John* for the anaphor in base position.

2.2 Patterns of reflexes of successive-cyclic movement

In this section I will introduce three abstract patterns of morphosyntactic reflexes of cross-clausal Ā-movement that surface on a heads in the vicinity of a final or intermediate landing site of movement. The attested variation in the distribution of reflexes will provide evidence for a split between final and intermediate landing sites. To facilitate matters, I will refer to the clause that contains the final landing site of a moved XP as the *final clause* and to clauses that exclusively contain intermediate landing sites of XP movement as *non-final clauses*. In the example in (15) where movement from CP_3 to CP_1 is supposed to stop in every SpecCP, CP_2 and CP_3 are non-final clauses, and CP_1, excluding CP_2 and CP_3, is the final clause.

(15) $[_{CP_1}$ What$_k$ do you think $[_{CP_2}$ __$_k$ that Mary believes $[_{CP_3}$ __$_k$ that Bill bought __$_k$]]]
            $\underbrace{\qquad\qquad\qquad}_{final}$ $\underbrace{\qquad\qquad\qquad\qquad}_{non-final}$ $\underbrace{\qquad\qquad\qquad\qquad}_{non-final}$

2.2.1 Pattern I: reflexes in final and non-final clauses

The most prominent pattern found with morphological reflexes of cross-clausal Ā-movement is the following: The reflex occurs once in *every* clause that is crossed by overt movement of an XP, i.e. in the clause in which the moving XP originates, in the clause in which it has its overt landing site and in all clauses in between. Put differently, it occurs in the final and non-final clauses. This is abstractly depicted in (16) for wh-movement of an XP from CP_3 to $SpecCP_1$, across two clause-boundaries. For ease of exposition, I assume that movement goes through every SpecCP, although there are probably more intermediate landing sites. Assume that wh-movement leaves a reflex 'R' on a C head crossed by movement (how R is realized, i.e. if it is a tonal, a morphological or a syntactic reflex, is irrelevant for the discussion). In a pattern I language, R occurs on the C head of each of the three CPs.

(16)   *Pattern I*:
        $[_{CP1}$ wh-XP$_k$ $[_{C'}$ **C$_1$-R** ... $[_{CP2}$ ___$_k$ $[_{C'}$ **C$_2$-R** ... $[_{CP3}$ ___$_k$ $[_{C'}$ **C$_3$-R** ... ___$_k$]]]]]]

There is massive evidence that this reflex is indeed related to movement (the abstract configurations of these scenarios are illustrated in (17)): First of all, the reflex is triggered in dependencies that show the classic diagnostics for displacement: The dependencies are island-sensitive and induces strong or weak cross-over effects (see Ross 1967, Postal 1971 on island-sensitivity and cross-over effects of movement). This is shown in (17-a) for XP movement from a complex NP island, and in (17-b) and (17-c) for weak and strong cross-over. Weak cross-over involves movement of XP across a coreferent pronoun that is included in a phrase YP; strong cross-over involves movement of XP across a coreferent pronoun. Note that (17-a-c) are also ungrammatical without the reflex of movement; it is not the distribution of the reflex that leads to ungrammaticality but the presence of an island / a coreferent pronoun. Second, there can be no reflex on any C head if XP does not undergo Ā-movement and stays in its base position. This may be so either because XP does not bear a feature (like the [WH]-feature) that forces its movement to the left periphery, cf. (17-d), or because the language allows for wh-in situ, cf. (17-e). Third, if the respective language allows for partial movement of XP to a position below its scope position, the reflex only shows up on C heads that are crossed by *overt* Ā-movement. If, for example, the XP takes scope in $SpecCP_1$ but moves only up to $SpecCP_2$, the reflex will be found on C_3 and C_2, but not on C_1, cf. (17-f). Even if there is subsequent covert movement

of XP to its scope position, it does not trigger a reflex in any of the languages I have looked at.[9] Finally, if the moving XP originates in the clause that contains its scope position and Ā-moves to this position, there can be no reflex on the C head of any lower clause that is not crossed by movement. In the example in (17-g) where XP originates in CP_1 and moves to $SpecCP_1$, we find the reflex only on C_1, but neither on C_2 nor on C_3.

(17)   *Abstract configurations of XP movement in a pattern I language*:

     a.   XP-movement out of a Complex NP island:

       *$[_{CP_1}$ wh-XP$_k$ $[_{C'}$ **C$_1$-R** ... $[_{CP2}$ ___$_k$ $[_{C'}$ **C$_2$-R** ... $[_{NP}$ N $[_{CP3}$ ___$_k$ $[_{C'}$ **C$_3$-R** ... ___$_k$]]]]]]]]

     b.   Weak cross-over induced by XP-movement across YP containing ZP:

       *$[_{CP_1}$ [wh-XP$_i$]$_k$ $[_{C'}$ **C$_1$-R** ... $[_{CP2}$ ___$_k$ $[_{C'}$ **C$_2$-R** $[_{YP}$ Y ... ZP$_i$] ... $[_{CP3}$ $[_{C'}$ ___$_k$ **C$_3$-R** ... ___$_k$]]]]]]]

     c.   Strong cross-over induced by XP-movement across ZP:

       *$[_{CP_1}$ [wh-XP$_i$]$_k$ $[_{C'}$ **C$_1$-R** ... $[_{CP2}$ ___$_k$ $[_{C'}$ **C$_2$-R** ... ZP$_i$ ... $[_{CP3}$ $[_{C'}$ ___$_k$ **C$_3$-R** ... ___$_k$]]]]]]]

     d.   XP in base position (no wh-feature):

       $[_{CP_1}$ **C$_1$(*-R)** ... $[_{CP2}$ **C$_2$(*-R)** ... $[_{CP3}$ **C$_3$(*-R)** ... XP]]]

     e.   Wh-in-situ:

       $[_{CP_1}$ **C$_1$(*-R)** ... $[_{CP2}$ **C$_2$(*-R)** ... $[_{CP3}$ **C$_3$(*-R)** ... wh-XP$_k$]]]

     f.   Partial movement of XP:

       $[_{CP_1}$ **C$_1$** ... $[_{CP2}$ wh-XP$_k$ $[_{C'}$ **C$_2$-R** ... $[_{CP3}$ ___$_k$ $[_{C'}$ **C$_3$-R** ... ___$_k$]]]]]

     g.   Extraction of XP from CP_1:

       $[_{CP_1}$ wh-XP$_k$ $[_{C'}$ **C$_1$-R** ... ___$_k$ $[_{CP2}$ **C$_2$(*-R)** ... $[_{CP3}$ **C$_3$(*-R)** ...]]]]

To summarize, the exponent R is a reflex of *overt* movement; covert movement does not trigger it, as becomes clear in wh-in-situ and partial movement constructions.

2.2.2   Pattern II: reflexes solely in the final clause

In languages with pattern II, cross-clausal Ā-movement of an XP triggers a morphological reflex only in the clause in which XP surfaces, but neither in the clause in which the XP originates nor in any non-final clause between the base position and the final landing site of XP. This is schematically shown in (18) (for a reflex R on the head C triggered by wh-movement).

(18)   *Pattern II*:

     $[_{CP1}$ wh-XP$_k$ $[_{C'}$ **C$_1$-R** ... $[_{CP2}$ ___$_k$ $[_{C'}$ **C$_2$** ... $[_{CP3}$ ___$_k$ $[_{C'}$ **C$_3$** ... ___$_k$]]]]]]

The wh-phrase is base-generated in CP_3 and moves to $SpecCP_1$, crossing two CPs on its way. But it only leaves a reflex on the C-head of the final clause C_1. If the reflex additionally showed up on the C head of either of the two non-final clauses or on both, ungrammaticality would result.

     The dependency has the same properties as in pattern I languages: The reflex does not occur if XP stays in its base position, cf. (19-a) ad (19-b); in case of partial movement, the reflex occurs only in the clause where the XP lands, but not in clauses above this position, cf. (19-c); if the XP originates in the clause that contains its final landing site, only this clause exhibits the reflex, cf. (19-d).

---

[9]An alternative approach is that an in-situ wh-phrase / a partially moved wh-phrase does not undergo covert movement to it scope position but is rather bound by an operator in this position (unselective binding, cf. Heim 1982). Since there is no covert movement in the first place under this analysis, there can also be no reflex of covert movement.

(19)   *Abstract configurations of XP movement in a pattern II language*:
      a.   XP in base position (no wh-feature):
           $[_{CP_1}$ **C$_1$(*-R)** ... $[_{CP2}$ C$_2$ ... $[_{CP3}$ C$_3$ ... XP $]]]$
      b.   Wh-in-situ:
           $[_{CP_1}$ **C$_1$(*-R)** ... $[_{CP2}$ C$_2$ ... $[_{CP3}$ C$_3$ ... wh-XP$_k$ $]]]$
      c.   Partial movement of XP:
           $[_{CP_1}$ C$_1$ ... $[_{CP2}$ wh-XP$_k$ $[_{C'}$ **C$_2$-R** ... $[_{CP3}$ ___$_k$ $[_{C'}$ C$_3$ ... ___$_k$ $]]]]]$
      d.   Extraction of XP from CP$_1$:
           $[_{CP_1}$ wh-XP$_k$ $[_{C'}$ **C$_1$-R** ... ___$_k$ $[_{CP2}$ C$_2$ ... $[_{CP3}$ C$_3$ ... $]]]]$

Hence, the occurrence of R in pattern II languages is also related to an *overt* dependency (in syntax proper, not at LF) between the C head and the XP. Whether the dependency actually involves movement or base-generation is a question that I will address in section 2.4.1. But let me anticipate that at least in some languages with pattern II there is evidence for movement (island-sensitivity, cross-over effects).

2.2.3   Pattern III: reflexes solely in non-final clauses

Pattern III is the mirror image of pattern II: Under cross-clausal extraction of an XP, a morphological reflex only occurs in the clauses that are crossed by movement but in which the XP does not land, viz. the reflex only occurs in non-final clauses. The pattern is illustrated in (20) (as before, with a reflex R on the C head triggered by wh-movement).

(20)   *Pattern III*:
         $[_{CP1}$ wh-XP$_k$ $[_{C'}$ C$_1$ ... $[_{CP2}$ ___$_k$ $[_{C'}$ **C$_2$-R** ... $[_{CP3}$ ___$_k$ $[_{C'}$ **C$_3$-R** ... ___$_k$ $]]]]]]$

In (20), the wh-phrase is base-generated in the most deeply embedded CP$_3$ and moves overtly to SpecCP$_1$. We find a morphological reflex of this movement in CP$_3$, the clause in which the moving XP originates, and in CP$_2$, a clause crossed by movement. But there must not be a reflex on C$_1$ because the wh-phrase lands in the specifier of that head. Just as pattern I and pattern II reflexes, pattern III reflexes are tied to *overt* movement: The dependency is island-sensitive and exhibits weak cross-over effects. Furthermore, there is no reflex on any C head if XP is not moved at all, cf. (21-a) and (21-b). In case of partial wh-movement to SpecCP$_2$, the reflex cannot occur in a clause above the overt landing site of the moved XP (and also not on C$_2$ because is projects the final landing site of overt movement), cf. (21-c). If XP originates in C$_1$, there can be no reflex in the lower clauses CP$_2$ and CP$_3$ because they are not crossed by XP-movement; there is actually no reflex at all then because CP$_1$ is the clause where XP surfaces.

(21)   *Abstract configurations of XP movement in a pattern III language*:
      a.   XP in base position (no wh-feature):
           $[_{CP_1}$ C$_1$ ... $[_{CP2}$ **C$_2$(*-R)** ... $[_{CP3}$ **C$_3$(*-R)** ... XP $]]]$
      b.   Wh-in-situ:
           $[_{CP_1}$ C$_1$ ... $[_{CP2}$ **C$_2$(*-R)** ... $[_{CP3}$ **C$_3$(*-R)** ... wh-XP$_k$ $]]]$
      c.   Partial movement of XP:
           $[_{CP_1}$ C$_1$ ... $[_{CP2}$ wh-XP$_k$ $[_{C'}$ C$_2$ ... $[_{CP3}$ ___$_k$ $[_{C'}$ **C$_3$-R** ... ___$_k$ $]]]]]$
      d.   Extraction of XP from CP$_1$:
           $[_{CP_1}$ wh-XP$_k$ $[_{C'}$ C$_1$ ... ___$_k$ $[_{CP2}$ C$_2$ ... $[_{CP3}$ C$_3$ ... $]]]]$

2.2.4   Pattern IV: no reflex in any clause

Patterns II and III are the mirror image of one another: One has a reflex in clauses (final or intermediate) where the other does not, and vice versa. By this logic, we expect a fourth pattern, the opposite of pattern I: There is no reflex of movement in any clause. Agreement between a head H and an XP moving to SpecHP fails in both final and non-final clauses. I will henceforth refer to this as pattern IV. It might seem to be misleading to talk about a pattern of a reflex if there is no reflex at all. Indeed, pattern IV is ambiguous: It can emerge if H initiates agreement with an XP but for some reason H does not find an XP that can function as an agreement controller; alternatively, pattern IV could also arise simply because H does not initiate agreement in the first place (it does not have a feature that triggers this agreement relation). In the latter case, we should not speak about an *agreement* pattern because there has not even been an attempt to Agree. But we will encounter languages where pattern IV freely alternates with one of the other patterns. Since patterns I – III show reflexes of movement, we know that the relevant heads in these languages do trigger agreement; the optional absence of the reflex then cannot simply be attributed to the general absence of agreement-triggering features on H. In such languages, it is actually meaningful to talk about a pattern IV reflex of movement even in the absence of a reflex: There is an agreement relation, but it can fail to result in feature transfer so that there is no reflex.

2.2.5   The relevance of the patterns in the literature

The distribution of reflexes of movement is of interest to syntacticians because it helps us to determine where and how often Ā-movement makes a stop-over on the way to the surface position of the moved element. In the theoretical literature on the locality of movement, pattern I reflexes figures prominently. This is not surprising because pattern I languages exhibit the largest number of reflexes: There is (at least) one in every clause along the path of movement; in pattern II and pattern III languages, we see reflexes only in a subset of these clauses. As a consequence, patterns II and III are much less discussed. Especially in recent formal work on reflexes of movement, the existence of pattern II is at most mentioned in passing (see e.g. Boeckx 2008*b*: 36, fn.7); alternatively, reflexes that do not occur in every clause along the path of movement are said to be no "real" reflexes of successive-cyclic movement (see Lahne 2008*b*); and even if these patterns are described, the distribution of the reflexes is not derived (see e.g. Torrence 2012 who analyses pattern I but none of the other patterns of reflexes in Wolof). Pattern III is virtually absent from the comparative formal literature; this may be due to the fact that it is barely described and seems to occur less frequently than patterns I and II. There are only two formal approaches (Müller 1999; Abels and Muriungi 2008), but they are not comparative: The analyses are developed to capture solely pattern III but it is not shown how the other patterns could be derived. There are a few analyses of the variation between patterns I and II from the 1980s and early 1990s (see sections 2.4.4 and 2.4.5); the only more recent work on the pattern I/II alternation comes from Bošković (2007*a*) and Reintges et al. (2006) (see section 2.5.3.1). The latter state explicitly that "no other patterns are attested" (p.167), contrary to fact: pattern III is attested, see section 2.3.3 for data. Thus, there is not a single analysis that tries to account for the whole range of patterns of reflexes of long-distance movement. In section 2.5 I will present a uniform approach to reflexes of movement that relies on the order of elementary operations and that predicts the attested variation.

2.3 Data

In this section, I will illustrate the abstract patterns with language data. For each language, I will show that the reflex tracks overt movement by investigating its distribution in the aforementioned contexts (extraction from islands, cross-over configurations, wh-in-situ, partial movement). Most of the data, especially those that exemplify patterns I and II, are well-known and have repeatedly been discussed in the literature. The reader who is familiar with the phenomena can go directly to the data summary on page 83.

The present collection of phenomena is not meant to be exhaustive. There are a lot more examples especially for patterns I and II; as for pattern III, the languages listed in section 2.3.3 are the only ones with this pattern that I know of. The selection presented here is made on the basis of two criteria: First, most of the languages are chosen because their reflexes of movement have properties that are of theoretical relevance for the discussion in the present and upcoming chapters (e.g. optionality between patterns, co-occurrence of different patterns). Second, other languages are included in order to give an impression of the cross-linguistic variation found in the exponence of reflexes of movement (apart from the different patterns). For one, we find **phonological**, **morphological** and **syntactic reflexes**. Furthermore, following the terminology introduced by Zentz (2013), phonological and morphological reflexes can be classified as involving **addition**, **deletion**, or **replacement** of exponents that occur without movement. We will see examples for each of these subclasses. Note that not all reflexes will occur on the C head, as in the abstract examples; reflexes will also be found on other heads but the abstract distribution pattern remains the same.

Before we begin, a few comments on the presentation of the data are in order: As before, referential indices are represented by the subscripts i and j; movement dependencies are identified by subscripts k and n. Underlining marks an extraction site; this is to remain neutral about the nature of the element in this position (trace, copy, etc.). I will not indicate all movement steps that apply in the examples; I only mark the base position of a moved element, but not its intermediate landing sites. First and foremost, this is for the sake of readability. Moreover, it is debated which positions serve as intermediate landing sites (see the previous section), and I will not make any claims about it at this point. The location of the reflexes will provide empirical evidence which positions are targeted by movement.

2.3.1 Pattern I reflexes

In this section, I present languages with a pattern I reflex. We will encounter the entire range of variation: phonological, morphological and syntactic reflexes that either result in addition, replacement or deletion.

2.3.1.1 WH-agreement in Chamorro

Chamorro (Austronesian, Guam and Mariana Islands) exhibits a pattern I reflex of movement on the verb that involves replacement of regular morphology: Simplifying somewhat at this point, verbs along the path of movement indicate the grammatical function of the extracted element and can no longer bear the regular subject-verb-agreement morpheme. Let me begin with some basic facts: Chamorro is head-initial language with otherwise relatively free word order. The unmarked order is VSO; adverbs follow the arguments, cf. (22):

(22)   *Declarative sentence (Chung 1998: 21, 25)*:

Ha-ätan   i   taotao mansu i   guagä'-ña
AGR-watch the person tame   the fish.basket-AGR
'The tame man looked (in) his basket.'

The verb itself inflects for transitivity of the predicate, mood (realis vs. irrealis) as well as person and number of the subject, indicated by a prefix or infix:

(23)   *Agreement markers,*
        active transitive (Chung 1998: 26):

	Realis	Irrealis
1sg	hu-	(bai) u-
2sg	un-	un-
3sg	ha-	u-
1incl du/pl	ta-	(u)ta-
1excl du/pl	in-	(bai) in-
2du/pl	in-	in-
3du/pl	ma-	uma-

(24)   *Agreement markers,*
        intransitive realis (Chung 1998: 27):

	Realis
sg/du	-um-
pl	man-

(25)   *Agreement markers,*
        intransitive   irrealis   (Chung
        1998: 27):

	Irrealis
1sg	(bai) u-
2sg	un-
3sg	u-
1incl du	(u)ta-
1excl du	(bai) in-
2du	in-
3du	u-
1incl pl	(u)tafan-
1excl pl	(bai) infan-
2pl	infan-
3pl	u-fan

Following Chung (1998), I gloss the subject-verb-agreement morpheme simply as 'AGR' in what follows; the actual person/number combinations are not relevant for the discussion. In constituent questions, the wh-word is in sentence-initial position, preceding the complementizer (if overt) and the verb, cf. (26). Wh-movement is obligatory in Chamorro, the wh-phrase cannot stay in-situ (Chung 1998: 209f.).

(26)   *Question formation in Chamorro (Chung 1998: 209-210)*:

a. Hayi$_k$ ma'añao-mu   ___$_k$
    who   WH.OBL.afraid-AGR
    'Who are you afraid of?'

b. [Ginin hayi]$_k$ na un-konni' i   neni ___$_k$
     from who   C AGR-take the baby
    'From whom did you take away the baby?'

Crucially, the morphological form of the verb changes when Ā-movement (wh-movement, relativization, focus movement) has taken place. I illustrate this with wh-movement. The verb varies according to the grammatical function of the extracted phrase. This phenomenon is known as *wh-agreement* in the literature on Austronesian languages. In the examples that follow, wh-agreement forms are glossed as 'WH' followed by an abbreviation that specifies the case of the extracted element (nominative, objective, oblique). The wh-agreement exponent is set in bold in case it is segmental. Consider the examples in (27). (27-a) provides the baseline declarative sentence; the other sentences are interrogative sentences based on (27-a). The verb changes its form when wh-movement takes place. It has an infix in (27-b) and (27-c) (subject and object question,

respectively); in (27-d) where an oblique XP is questioned there is nominalization. Nominalization can be identified by the possessive affix on the verb that agrees in person and number with the subject, and by the fact that the internal argument of the verb is in the oblique case instead of the unmarked case (Chung 1998: 242). Note that object extraction, signaled by an infix, can additionally be accompanied by nominalization, cf. (27-c).[10]

(27)   *Wh-agreement with clause-bound extraction (Chung 1998: 236)*:

    a.   Ha-fa'gasi si   Juan i   kareta
          AGR-wash CASE Juan the car
          'Juan washed the car.'   *declarative*

    b.   Hayi$_k$ f**um**a'gasi   ___$_k$ i   kareta
          who   WH.NOM.wash   the car
          'Who washed the car?'   *SU extraction, V: [+real], [+trans]*

    c.   Hasa$_k$ f**in**a'gasése-nña   si   Henry ___$_k$ pära hagu
          what WH.OBJ.wash.PROG-3SG.POSS CASE Henry   for   you
          'What is Henry washing for you?   *DO extraction, V: [+real], [+trans]*

    d.   Hafa$_k$ pära fa'gase-mmu   ni   kareta ___$_k$
          what FUT WH.OBL.wash-2SG.POSS CASE kareta
          'What are you going to wash the car with?'
          (Lit: 'What will be your washing of the car with?')
                          OBL extraction, V: [–real], [+trans]

The following table gives an overview of the morphological changes. The left hand column contains the case (grammatical function) of the extracted XP and the right hand column shows the corresponding exponent(s) + additional conditions on the realization of these exponents. Oblique XPs include oblique complements of intransitive predicates, instruments and comitatives.

(28)   *Wh-agreement exponents (Chung 1998: 236)*:

case of extracted XP	exponent
nominative	-*um*- infix
	(if the predicate is realis and transitive)
objective	optional nominalization + -*in*-infix
	(infix occurs if the predicate is transitive)
oblique	nominalization + optional -*in*- infix
	(infix occurs if the predicate is unaccusative)

To summarize, when the extracted element is the subject (nominative) or the direct object (objective) of the verb, the change only occurs on transitive verbs, but not on intransitive verbs; in any case, the verb has an infix (depending on mood in case of subject extraction). If, however, an oblique is extracted, both transitive and intransitive verbs are affected by the special morphology; the verb is nominalized. As we can see in (27), wh-

---

[10]Case is encoded by a particle that precedes the noun. Chamorro distinguishes three cases (unmarked, oblique, local) that have different realizations for common nouns, proper names and pronouns:

(i)   *Case markers in Chamorro (Chung 1998: 50)*:

	Unmarked	Oblique	Local
Common noun	–	ni	gi
Proper name	si	as	gias
Pronoun	–	nu	giya

agreement replaces the usual subject-verb-agreement morphology. In (27-b) – (27-d), the subject agreement prefix *ha-*, present in the declarative sentence (27-a), is absent. Only if there is no special wh-agreement exponent in a given context does the usual subject-verb-agreement form occur even under extraction. This is the case, for example, if the subject of a transitive irrealis verb is extracted (infixation only occurs with transitive realis verbs), cf. (29):

(29)   *No wh-agreement under extraction (Chung 1998: 238)*:

Hayi$_k$ pära u-sangani   yu' ___$_k$ ni   ansa
who FUT 3SG.IRR-say.to me   CASE answer
'Who is going to tell me the answer?'   *SU extraction, V: [–real], [+trans]*

There is evidence that wh-agreement is indeed a reflex of movement: First of all, the movement types that trigger wh-agreement are island-sensitive. This is illustrated in (30-a) for wh-movement out of a complex NP (*Complex NP constraint*, cf. Ross 1967). Furthermore, wh-dependencies show strong cross-over effects when the gap is c-commanded by a coreferential pronoun, cf. (30-b) vs. (30-c); note that Chamorro allows for pro-drop of subjects, direct objects, and possessors.

(30)   *Island-sensitivity and strong cross-over (Chung 1998: 211-212)*:[11]

   a. *[Hayi siha na famagu'un]$_k$ un-rispeta   [$_{NP}$ ädyu i   palao'an [$_{CP}$ ni
       who PL L children   AGR-respect   DEM the woman   C
       f*um*a'na'gui   ___$_k$]]
       WH.NOM.teach
       'Which children do you respect the woman who taught?'   *CNP island*
   b. [Hayi$_i$]$_k$ malago'-ña   pro$_j$ na un-na'facho'chu'   ___$_k$
       who   WH.OBL.want-AGR   C WH.OBJ-make.work
       'Who$_i$ does he$_j$ want you to hire t$_i$?'
   c. *[Hayi$_i$]$_k$ malago'ña   pro$_i$ na un-na'facho'chu'   ___$_k$
       who   WH.OBL.want-AGR   C WH.OBJ-make.work
       'Who$_i$ does he$_i$ want you to hire t$_i$?'   *SCO*

Crucially, under long extraction every verb along the path of movement exhibits wh-agreement (if there is a wh-agreement form at all in the given context). In (31), the verb in the embedded clause indicates that an object has been extracted. The main verb also shows a wh-agreement form, thus wh-agreement follows pattern I.[12]

(31)   *Long wh-movement in Chamorro (Chung 1998: 247)*:

[$_{CP}$ Hafa$_k$ **sin**angani   hao ni chi'lu-mu [$_{CP}$ malago'-ña   ___$_k$]]
what WH.OBJ.PASS.say.to you CASE sibling-AGR   WH.OBL.want-AGR
'What did your sister tell you that she wants?'   *OBL extraction*

---

[11]The element glossed L in (30) is a linker. It shows up if a noun has a modifier. If the modifier is to the right of the noun, the linker is the suffix -*n*; if the modifier is to the left of the noun, it is realized as the particle *na* (Chung 1998: 231).

[12]In fact, the situation with long-distance extraction is a bit more complicated. As demonstrated in Chung (1994), the verbs in the clauses above the clause from which the extracted element originates obligatorily exhibit wh-agreement only if the moved operator is non-referential (i.e., "when it is novel and lacks descriptive content", Chung 1998: 248); if it is referential (i.e., "when it is familiar or has descriptive content", Chung 1998: 248) wh-agreement on the higher verbs is optional.

Although wh-agreement occurs throughout the movement path, it is not the case that every verb agrees with the moving XP, as (31) might suggest. In general, only the verb in the clause from which the Ā-moved phrase originates indicates the grammatical function of this phrase. All higher verbs do not indicate the grammatical function of the extracted phrase but rather of the clause from which the moving phrase is extracted, as becomes obvious in the examples in (32):[13]

(32)   *Long wh-movement in Chamorro (Chung 1998: 249-250)*:

   a.   Hafa$_k$ malago'-ña   si   Magdalena [pära ta-chuli'   ___$_k$]
        what   WH.OBL.want-3SG.POSS   CASE   Magdalena   FUT   WH.OBJ-bring
        'What does Magdalena want us to bring?'

   b.   Hayi$_k$ si   Manuel h*in*assóso-nña   [ch*um*uli'   ___$_k$ i
        who   CASE   Manuel   WH.OBJ.think.PROG-3SG.POSS   WH.NOM.take   the
        salappi']
        money
        'Who does Manuel think has taken the money?'

In (32-a) the direct object of the embedded verb is extracted. Therefore, this verb shows the objective wh-agreement form. The embedded clause is an oblique complement of the matrix verb. It therefore shows the oblique wh-agreement form (nominalization), and not agreement with the grammatical function of the wh-phrase, which would result in objective wh-agreement. In (32-b), the subject of the embedded verb is extracted out of an object clause. Hence, the embedded verb shows nominative wh-agreement and the matrix verb has the objective wh-agreement form. In (31) where an object is extracted from a complement clause we get objective wh-agreement on both the embedded and the matrix verb, creating the illusion that both verbs agree with the wh-phrase. Despite this particular pattern, wh-agreement instantiates pattern I: It occurs in every clause crossed by movement.

   Note that Chamorro does not allow for partial wh-movement or wh-in-situ, so there is no evidence from these constructions for the link of the reflex to overt or covert movement. But I will present data from a closely related language with wh-agreement that does have these constructions in section 2.6.3.3; we will see that in this language wh-agreement is tied to overt movement.

2.3.1.2   Complementizer selection in Irish

Modern Irish (Indo-European, Ireland) exhibits a pattern I reflex that involves replacement: The default form of the complementizer is replaced by a special form under Ā-extraction. Modern Irish is a verb-inital language. The unmarked word order is VSO; oblique argumens and adverbials follow the core argument, cf. (33).[14]

(33)   *Declarative sentence (McCloskey 1990: 67)*:

        Bhéarfaidh mé an t-aírgead do Chaaimhín i   nDoire inniu
        give.FUT   I   the money   to Kevin   in Derry today
        'I'll give Kevin the money in Derry today.'

---

[13]A similar pattern can be found in Tagalog (Austronesian), cf. Rackowski (2005); Rackowski and Richards (2005).

[14]The glosses and the representation of the complementizer varies to some extent in the literature cited in this section. I have adjusted the glosses to make them coherent.

Subject pronouns must be null if the verb is inflected for person and number of the subject (McCloskey 1990: 74). In case of Ā-movement the extracted element precedes the complementizer and the verb. Crucially, the form of the complementizer varies depending on whether Ā-movement has taken place, as first described in McCloskey (1979). An embedded finite declarative sentence is introduced by the complementizer *go*.[15] In the examples that follow, this marker is combined with the past tense marker -*r* and changes its surface form to *gu-* in this context. If an element of the clause undergoes Ā-extraction, leaving a gap in its base position, the form of the complementizer obligatorily changes to a form glossed as *aL*.[16] This basic pattern is illustrated in (34) for relativization and wh-movement, but it can also be found in comparatives, clefts and certain adverbial clauses (see McCloskey 1979: 153ff. for examples). (34-a) shoes a declarative sentence without Ā-extraction and the default complementizer *go*; (34-b) and (34-c) provide examples with Ā-movement.

(34)   *go / aL-alternation with clause-bound Ā-movement (McCloskey 2001: 67, 70)*:

      a.   Deir said **gu**-r   ghoid na síogaí í
          say   they *go*-PST stole   the fairies her
          'They say that the fairies stole her away.'   *declarative*

      b.   an ghirseach$_i$ [$_{CP}$ [OP$_i$]$_k$ **a**   ghoid na   síogaí ___$_k$]
          the girl   *aL* stole   the fairies
          'the girl that the fairies stole away'   *DO relativization*

      c.   [Cá fhad   ]$_k$ **a**   bhí   siad fá   Bhaile Átha Cliath ___$_k$
           Q   length   *aL* be.PST they around Dublin
          'How long were they in Dublin?'   *ADJ wh-movement*

There is evidence that the the *go / aL*-alternation is indeed a reflex of movement. First of all, the dependencies that require *aL*-marking are subject to the standard island constraints. This is shown in (35) for movement from a complex NP, and in (36) for movement from a wh-island CP introduced by the interrogative particle *ar ~ an* (*Wh-Island Condition*, cf. Chomsky 1973). In the a.-examples, the operator OP$_i$ undergoes the fatal movement step. Note that *nachN* is the negative allomorph of the complementizer.

(35)   *Complex NP Constraint in Modern Irish (McCloskey 1979: 30)*:

    a.   *an fear$_i$ [$_{CP}$ [OP$_i$]$_k$ **a**   phóg   mé an [$_{NP}$ bhean$_j$ [$_{CP}$ [OP$_j$] **a**   phós   ___$_k$]]]
       the man   *aL* kissed I   the   woman   *aL* married
       'the man who I kissed the woman who married'   *relativization*

    b.   *[$_{CP}$ [Cén   fear]$_k$ **a**   phóg tú   an [$_{NP}$ bhean$_i$ [$_{CP}$ OP$_i$ **a**   phós   ___$_k$]]]
         which man   *aL* kissed you the   woman   *aL* married
       'Which man did you kiss the woman who married?'   *wh-movement*

(36)   *Wh-islands in Modern Irish (McCloskey 1979: 32)*:

    a.   *bean$_i$ [$_{CP}$ [OP$_i$]$_k$ nachN bhfuil fhios agam [$_{CP}$ an   bpósfadhc   duine
       woman   C.NEG I know   PART.Q would.marry person
       ar   bith ___$_k$]]
       any
       'a woman who I don't know if anyone would marry'   *relativization*

---

[15]See McCloskey (1979: 10ff.) and McCloskey (1990; 2001) for arguments that *go* and its alternants are complementizers; see Sells (1984) and Noonan (1992; 2002) for a different view.

[16]*aL* causes a mutation of the following segment, viz. lenition. The symbol *L* is is an abbreviation for lenition (McCloskey 2001: 78).

b. *[$_{CP}$ [Cén   sagart]$_k$ **a** d'fhiafraigh Seán díot   [$_{CP}$ ar   bhuail tú ___$_k$]]
      which priest   *aL* asked   John of.you   PART.Q hit   you
   'Which priest did John ask you if you hit?'   *wh-movement*

In addition, an Ā-dependency that triggers the occurrence of *aL* induces a weak cross-over effects when a phrase containing a coreferent pronoun is crossed, cf. (37):

(37)   *Weak cross-over effect in Modern Irish aL-chains (McCloskey 1990: 110)*:

*[fear$_i$] [$_{CP}$ [OP$_i$]$_k$ **a**   d'fhág a$_i$ bhean ___$_k$]
   man   *aL* left   his wife
'a man that his wife left'

Furthermore, if an element is Ā-moved within a clause, the special form *aL* can only appear in this clause, but not in clauses embedded in the former clause, cf. (38-a); the embedded clause in introduced by the default complementizer *go*. In (38-b) extraction takes place from CP$_2$; the lower clause CP$_3$ contains *go* because it is not crossed by movement.

(38)   *Extraction from the matrix clause (McCloskey 1979: 151)*:

a.   [an fear]$_i$ [$_{CP_1}$ [OP$_i$]$_k$ **a** shíl   ___$_k$ [$_{CP_2}$ **go** mbeash sé$_i$ ann]]
      the man   *aL* thought   *go* would.be he there
   'the man that thought he would be there'
b.   [an fear]$_i$ [$_{CP_1}$ [OP$_i$]$_k$ **a**   dúirt sé [$_{CP_2}$ **a** shíl   ___$_k$ [$_{CP_3}$ **go** mbeadh sé$_i$
      the man   *aL* said he   *aL* thought   *go* would.be he
   ann   ]]]
   there
   'the man that he said thought he would be there'

Irish does not allow for wh-in-situ nor for partial wh-movement, so we cannot test whether the reflex is tied to overt movement or also to covert movement. Under long extraction, the change from *go* to *aL* obligatorily applies in every clause along the path of movement (see also McCloskey 1990: 74-75, McCloskey 2002: 185):

(39)   *go / aL-alternation with cross-clausal Ā-movement (McCloskey 1979: 54)* :

a.   [an t-úrscéal]$_i$ [$_{CP}$ [OP$_i$]$_k$ **a** mheas mé [$_{CP}$ **a** dúirt sé [$_{CP}$ **a** thuig
      the novel   *aL* thought I   *aL* said   he   *aL* understood
   sé ___$_k$]]]
   he
   'the novel that I thought he said he understood'   *long relativization*
b.   [$_{CP}$ [cén   t-úrscéal]$_k$ **a** mheas mé [$_{CP}$ **a** dúirt sé [$_{CP}$ **a** thuig   sé
      which novel   *aL* thought I   *aL* said he   *aL* understood he
   ___$_k$]]]
   'Which novel did I think he said he understood?'   *long wh-movement*

Thus, complementizer choice in Irish exhibits pattern I.

2.3.1.3   Complementizer agreement in Wolof *u*-chains

Wolof (Niger-Congo, Senegal) has a pattern I reflex of movement that involves addition of material: Complementizers agree in noun class with an operator in their specifier. Wolof has about 15 noun classes that are also expressed on determiner elements (following the noun) and adjectives (Torrence 2012: 1149f.). The basic word order is SVO; adjuncts follow the objects, cf. (40):

(40)   *Declarative sentence in Wolof (Torrence 2012: 1149):*

Xaj y-i   lekk-na-ñu ceeb b-i   ci   kër   g-i
dog CL.PL-DEF.PROX eat-FIN-3PL rice CL-DEF.PROX PREP house CL-DEF.PROX
'The dogs ate the rice at the house.'

Complementizer agreement will be illustrated for wh-movement in what follows, but it also occurs in relativization (see Torrence 2012: 1153ff.). Examples are given in (41) for the so-called *u*-construction:

(41)   *u-construction in Wolof (Torrence 2012: 1151):*

a. Ø$_k$ **k**-u ___$_k$ togg ceeb ak jën
    Q   CL-u   cook rice and fish
    'Who cooked rice and fish?'   *SU question*

b. Ø$_k$ **y**-u jigéén j-i   togg ___$_k$
    Q   CL-u woman CL-DEF.PROX cook
    'What(pl) did the woman cook?'   *DO question*

c. Ø$_k$ **f**-u jigéén j-i   togg-e   ceeb ak jën ___$_k$
    Q   CL-u woman CL-DEF.PROX cook-LOC rice and fish
    'Where did the woman cook fish and rice?'   *ADJ question*

In this construction, the complementizer surfaces as *u* and is preceded by a class prefix that indicates class agreement with the wh-phrase.[17] As Torrence (2012) convincingly argues, the wh-phrase in these examples is located in the left periphery, but it is zero, a null operator, indicated by Ø in the examples. The interpretation of the sentence, i.e. whether a person, an object or a place is questioned, can be inferred from the noun class marker on the complementizer that agrees with the zero operator: In (41) you can see that the class marker varies with the interpretation of the operator.[18]

Question formation with the *u*-construction clearly involves movement. First of all, the dependency is sensitive to islands. This is illustrated in (42) for wh-movement from an adjunct island (cf. Huang 1982; Chomsky 1986*a*) and a coordination island (*Coordinate Structure Constraint*, cf. Ross 1967); for more island tests see Torrence 2013*a*.

(42)   *Wh-movement from islands in Wolof (Torrence 2012: 1157, Torrence 2013a: 182f.):*

a. *[$_{CP}$ Ø$_k$ **y**-u xale b-i   dem [$_{CP}$ laata Bintë togg-al   Móódu ___$_k$]]
        Q   CL-u child CL-DEF.PROX leave   before Binte cook-BEN Moodu
    'What(pl) did the child leave before Binte cooked for Moodu?' *adjunct island*

---

[17]In relative clauses, the complementizer can surface as *u*, *i*, or *a*. The choice has an interpretative difference: It indicates the definiteness of the head noun of the relative clause. The *u*-form occurs with indefinite nouns, the *i*-form indicates that the head noun is definite and spatially or temporally proximal, the *a*-form indicates that the head noun is definite and spatially or temporally distal (Torrence 2012: 1155). Note that the same distinction is made for determiners, see e.g. the definite proximate determiners in (40). Hence, the complementizer not only agrees in noun class with the operator in its specifier, but also in definiteness. It is thus no surprise that we find *u* in *u*-chains because wh-operators are indefinite.

[18]Let me reproduce some of Torrence's arguments that the *u*-element is in fact a complementizer and not itself the wh-word (see Torrence (2012: 1157ff.)). For one, it can surface on the left edge of embedded clauses that do not contain a gap, i.e. from which no extraction has taken place. This would be mysterious if the *u*-element was an operator. Second, it is in complementary distribution with other complementizers in non-interrogative clausal complements. Furthermore, there can only be a single *u*-form in a clause, but in the *an*-construction, where the wh-word is overt, more than one wh-phrase can be moved to the left periphery (with an echo question reading); if the *u*-element was a wh-operator, we would expect that more than one u-element can be int he left periphery, but this is ungrammatical. Finally, in echo questions, overt wh-phrases can remain in their base position. *u*-elements, however, can never appear inside TP.

  b. *$\emptyset_k$ **l**-u ñu jend a-y   [$_{\&P}$ nen ak ___$_k$]
     Q CL-u 3PL buy INDEF-CL.PL   eggs and
     'What did they buy eggs and?'   *coordination island*

Furthermore, the silent wh-expression can be interpreted in the scope of a quantifier:

(43)   *Reconstruction in u-chains (Torrence 2012: 1176)*:

       Ø **l**-u [xale b-u nekk] bëgg
       Q CL-u child CL-u exist love
       'Who does every child love?'   *Wh > ∀, ∀ > Wh*

This can be accounted for if the wh-expression originates in the c-command domain of
the quantifier and is Ā-moved to SpecCP; it can be reconstructed into its base position
for interpretation, giving rise to the Wh > ∀ reading. The same facts hold for long wh-
movement (see Torrence 2012: 1176).

     Zero wh-expressions cannot remain in situ (see Torrence 2012: 1167 for an argument
from multiple wh-questions), so we cannot test whether class agreement on the comple-
mentizer is strictly tied to overt movement. Crucially, under cross-clausal wh-movement,
class agreement with the silent operator occurs on every C along the path of movement:

(44)   *Pattern I under long-distance wh-movement, u-chains (Torrence 2012: 1171)*:[19]

       a. [$_{CP}$ $\emptyset_k$ **f**-u a defe [$_{CP}$ **f**-u Maryam wax [$_{CP}$ **f**-u ñu teg tééré
              Q CL-u 2SG think   CL-u Maryam say   CL-u 3PL put book
           b-i   ___$_k$]]]
           CL-DEF.PROX
           'Where do you think Maryam said they put the book?'
       b. [$_{CP}$ $\emptyset_k$ **k**-u Kumba wax [$_{CP}$ ne **k**-u Isaa defe [$_{CP}$ ne **k**-u Maryam dóór
              Q CL-u Kumba say   FRC CL-u Isaa think   FRC CL-u Maryam hit
           ___$_k$]]]
           'Who did Kumba say that Isaa thought that Maryam hit?'

Hence, complementizer agreement in *u*-chains follows pattern I. Torrence (2012: 1172)
takes this as evidence for successive-cyclic movement of the (zero) wh-operator: It moves
(at least) through every SpecCP on the way to its final landing site and triggers agreement
on the complementizer in each case. Note that complementizer agreement throughout
the path of movement is only optional. Another grammatical variant is to have agreement
only on the complementizer in whose specifier the silent wh-phrase lands. All lower com-
plementizers bear the class marker *l-*, which Torrence (2013*b*) identifies as a default class
marker of the non-human class; its default character is evident because it is invariant, i.e.
it does not alternate with the class of the extracted wh-phrase; in addition, the *l-* class
marker is used when the class of an object is unknown, e.g. in the Wolof equivalent of the
question 'What is this?'.

(45)   *Pattern II under long-distance wh-movement, u-chains (Torrence 2012: 1171)*:

       [$_{CP}$ $\emptyset_k$ k-u Kumba wax [$_{CP}$ ne l-a   Isaa defe [$_{CP}$ ne **l**-a   Maryam dóór
          Q CL-u Kumba say   FRC EXPL-a Isaa think   FRC EXPL-a Maryam hit
       ___$_k$]]]
       'Who did Kumba say that Isaa thought that Maryam hit?'

---

[19]The morpheme *ne*, glossed as FRC, is a force marker.

Thus, complementizer agreement can exhibit pattern II in addition to pattern I: Only the complementizer adjacent to the silent wh-expression must agree in class; other complementizers only optionally do so. Note that the Ā-dependency in *u*-chains that triggers pattern II shows the same reconstructions effects as the Ā-dependency that causes pattern I (cf. (46)), indicating that both involve movement.

(46) *Reconstruction in Wolof u-chains under long wh-movement (Torrence 2013a: 163)*:

    a. [$_{CP}$ Ø$_k$ **k**-u a   foog [$_{CP}$ **k**-u [xale b-u nekk] bëgg ___$_k$]]
          Q CL-u 2SG think   CL-u child CL-u exist love
        'Who do you think that every child loves?'   *PI: Wh > ∀, ∀ > Wh*

    b. [$_{CP}$ Ø$_k$ **k**-u a   foog [$_{CP}$ l-a   [xale b-u nekk] bëgg ___$_k$]]
          Q CL-u 2SG think   EXPL-a child CL-u exist love
        'Who do you think that every child loves?'   *PII: Wh > ∀, ∀ > Wh*

Indeed, there is even more variation in the distribution of agreeing complementizers in Wolof. I will discuss the other patterns in section 2.5.3.2. Complementizer agreement as a pattern I reflex of movement occurs in many other languages as well. It is, for example, prominent in Bantu languages (see e.g. Schneider-Zioga (1995); Carstens (2003); Baker (2008*b*)).

2.3.1.4 *men*-deletion in Indonesian

Indonesian (Austronesian, Indonesia) exhibits a pattern I reflex of Ā-movement that manifests itself as deletion. The basic word order of this language is SVO; there is no agreement between the verb and its arguments, the strict word order is the primary cue of argument encoding. However, verbs bear prefixes indicating voices. Of central importance for the following discussion is the verbal prefix *men* (and its alternants that result from phonological processes) which encodes active voice. Since the following data are from Saddy (1991), who identifies *men* as a marker of transitivity, *men* is glossed as TRANS.[20] This marker is optional in declarative clauses:

(47) *Declarative sentence with an active transitive verb (Macdonald and Dardjowidjojo 1967: 238)*:

    Dia (mem)-beli   buku itu
    he TRANS-bought book DEM
    'He bought that book.'

In constituent questions, the wh-phrase moves to the clause-initial position and must be immediately followed by the marker *yang*, which Saddy (1991) identifies as a focus marker (see section 2.3.2.4 for further discussion of this marker). Crucially, the prefix *men* is no longer optional when an argument originating from within the VP is extracted; instead, it must be deleted, compare the object questions in (48-a) and (48-b) with and without *men*-deletion, respectively.

(48) *Obligatory men-deletion under wh-movement (Saddy 1991: 186)*:

    a. Siapa yang Sally cintai ___$_k$
        who$_k$ FOC Sally love
        'Who does Sally love?'

---

[20]However, it is clear that is is not a marker of transitivity because it also occurs with intransitive verbs; see section 2.6.2.2 for evidence.

    b. *Siapa$_k$ yang Sally men-cintai ___$_k$
       who   FOC Sally TRANS-love
       'Who does Sally love?'

If the subject of a transitive verb or an adjunct is extracted, however, *men*-deletion does not have to apply. As for local subject extraction, this is illustrated in (49). Note that since subject extraction is string-vacuous, overt movement is only detectable from the occurrence of the focus marker *yang*:[21] Saddy (1991) does not provide examples with adjunct extraction, but explicitly states that *men*-deletion is not obligatory in this context (p.214, n.3).[22]

(49)   *Subject questions in Indonesian (Saddy 1991: 186)*:

    Siapa$_k$ yang ___$_k$ (men)-cintai Sally
    who   FOC   TRANS-love Sally
    'Who loves Sally?'   *SU extraction*

The following facts suggest that *men*-deletion is a reflex of movement: First of all, wh-movement which triggers deletion, is island-sensitive as illustrated in (50-a) for a wh-island and in (50-b) for a complex NP island.

(50)   *Island-sensitivity of overt wh-movement in Indonesian (Saddy 1991: 190)*:

    a. *[$_{CP}$ Apa$_k$ yang kamu katakan [$_{CP}$ dimana kita beli   ___$_k$]]
       what FOC you mention   where we bought
       'What did you mention where we bought?'   *wh-island, wh-ex-situ*
    b. *[$_{CP}$ Siapa$_k$ yang kamu sukai [$_{NP}$ cerita [$_{CP}$ yang mengeritik ___$_k$ itu]]]
       who   FOC you like   stories FOC criticize   the
       'Who do you like stories that criticize?'   *CNPC, wh-ex-situ*

Furthermore, wh-movement induces weak cross-over effects:[23]

---

[21]This focus marker is absent in the corresponding wh-in-situ sentence.

[22]The dialect of Malay described in Cole and Hermon (1998; 2000) also has a process of *men*-deletion which, by and large, works as in Indonesian. Cole et al. (2006) provide examples which show that adjunct extraction does not lead to deletion of the verbal prefix:

(i)   *Adjunct extraction in Malay (Cole et al. 2006)*:

    a. [Di mana]$_k$ Ali mem-ukul Ahmad ___$_k$
       LOC which   Ali TRANS-hit Ahmad
       'Where did Ali hit Ahmad?'
    b. Bilamana$_k$ Ali mem-ukul Ahmad ___$_k$
       when   Ali TRANS-hit Ahmad
       'When did Ali hit Ahmad?'

[23]Saddy (1991) shows that wh-in-situ does not trigger strong or weak cross-over effects in Indonesian, which further indicates that it does not involve overt movement (with spell-out of the lowest position of the movement chain). This is consistent with the observation that *men*-deletion is not obligatory in wh-in-situ constructions.

(i)   *Strong and weak cross-over effects in questions in Indonesian (Saddy 1991: 207-208)*:

    a. dia$_i$ men-harp   Jon men-cintai siapa$_i$
       he   TRANS-expect Jon TRANS-love who
       'Who does he expect Jon to love?'   *SCO*
    b. Prof   dia$_i$ men-gira   saya men-centai siapa$_i$
       professor he TRANS-think I   TRANS-love who
       'Who does his professor think I love?'   *WCO*

(51)   *Weak cross-over effects (Cheng 1991: 67)*:

    *[siapa$_i$]$_k$ yang dosen-nya$_i$ suka ___$_k$
       who   FOC professor-his like
    'Who does his professor like?'

There is also evidence that *men*-deletion is tied to overt movement (but not to LF-movement): In addition to full wh-movement to the scope position, Indonesian also allows for partial wh-movement and wh-in situ.[24] If the wh-phrase, taking matrix scope, stays in situ, *men*-deletion is not obligatory; the prefix can be optionally realized, cf. (52-a) and (52-b). In case of partial wh-movement, *men* obligatorily deletes only in the clause(s) that the wh-phrase has overtly crossed, but not in the clause(s) above the partially moved wh-phrase, cf. (52-c): *men* is optional in the matrix clause that is not affected by movement but obligatorily absent in the embedded clause including the partially moved wh-phrase.

(52)   *Wh-in-situ and partial wh-movement (Saddy 1991: 186, 192, 194)*:

    a. Sally (men)-cintai siapa
        Sally TRANS-love who
        'Who does Sally love?'   *wh-in situ*

    b. [$_{CP}$ Tom tidak (meng)-harap [$_{CP}$ Mary (mem)-beli   apa]]
           Tom NEG TRANS-expect   Mary TRANS-bought what
        'What doesn't Tom expect Mary bought?'   *wh-in situ*

    c. [$_{CP}$ Tom (meng)-harap [$_{CP}$ apa$_k$ yang Mary (*mem-)beli   ___$_k$]]
           Tom TRANS-expect   what FOC Mary TRANS-bought
        'What does Tom expect Mary bought?'   *partial wh-movement*

In case of cross-clausal Ā-movement of a direct object, *men*-deletion must apply throughout the movement path, i.e. on all verbs between the base position and the final landing site of the wh-phrase that can bear *men* in the corresponding declarative sentence. (53-a) shows a wh-in situ question (in which there is no *men*-deletion); (53-b) illustrates obligatory prefix deletion under overt movement.

(53)   *Long-distance movement and men-deletion (Saddy 1991: 186)*:

    a. [$_{CP}$ Bill (men)-gira   [$_{CP}$ Tom (men)-harap [$_{CP}$ Fred (men)-cintai siapa]]]
           Bill TRANS-think   Tom TRANS-expect   Fred TRANS-love who
        ''Who did Bill think Tom expects Fred loves?   *wh-in situ*

    b. [$_{CP}$ Siapa$_k$ yang Bill (*men-)kira [$_{CP}$ Tom (*men-)harap [$_{CP}$ Fred
           who   FOC Bill TRANS-think   Tom TRANS-expect   Fred
        (*men-)cintai ___$_k$]]]
        TRANS-love
        ''Who did Bill think Tom expects Fred loves?   *wh-ex situ, deletion obligatory*

Recall that extraction of a transitive subject does not trigger obligatory *men*-deletion. Extraction of the transitive subject from an embedded clause, however, must lead to *men*-deletion on all verbs crossed by overt movement above the clause from which the subject originates.

---

[24]The three strategies are available for verbal arguments and arguments of prepositions; adjuncts are more restricted in the movements they can undergo, cf. Saddy (1991: 214, fn.3 and 6).

(54) *Extraction of an embedded subject (Saddy 1991: 187)*:

    a. [$_{CP}$ Bill (men)-beri [$_{CP}$ Tom (men)-harap [$_{CP}$ siapa (men)-cintai Fred]]]
        Bill TRANS-think   Tom TRANS-expect   who TRANS-love Fred
        'Who does Bill think Tom expects loves Fred?'   *wh-in situ*

    b. [$_{CP}$ Siapa$_k$ yang Bill (*mem-)beri [$_{CP}$ Tom (*mem-)harap [$_{CP}$ ___$_k$
        who   FOC Bill TRANS-think   Tom TRANS-expect
        (men)-cintai Fred]]]
        TRANS-love   Fred
        'Who does Bill think Tom expects loves Fred?' *wh-ex situ, deletion obligatory*

Hence, obligatory *men*-deletion in Indonesian is a reflex of overt movement that follows
pattern I, even if it does not apply under local subject extraction (which will be explained
in section 2.6.2.2).

    Finally, note that virtually the same pattern I reflex exists in Singaporean Malay
(Austronesian, Malaysia), described in Cole and Hermon (1998; 2000): The prefix *meng*
optionally precedes transitive verbs. It is obligatorily deleted under overt wh-movement
(as well as under relativization and focus movement) of VP-internal arguments on all
verbs between the base position and the overt landing site of the wh-phrase. As in In-
donesian, this is a reflex of overt movement: The movement types that trigger deletion
are island-sensitive; furthermore, deletion does not occur with wh-in-situ and it also does
not apply to verbs above the overt landing site of a partially moved wh-phrase.

2.3.1.5   Downstep deletion in Kikuyu

Kikuyu (Bantu, Kenya), a tone language, has a phonological pattern I reflex of movement
that involves deletion: Under Ā-movement, a downstep is deleted, which has several con-
sequences for tone assignment. In what follows, high tones are represented by an acute,
low tones are indicated by the absence of a symbol on the vowel, and extra low tones are
represented by a grave; the haček symbolizes a rising tone and the circumflex a falling
tone (Clements 1984*b*: 283).

    The basic word order in Kikuyu is SVO; adverbs follow the arguments, cf. (55-a). A
verbal prefix, glossed as *subject marker* 'SM', indicates class agreement with the subject.

(55) *Downstep shift in a declarative sentence (Clements et al. 1983: 7-8)*:

    a. mondo$^!$ a-hɛ-írɛ́$^!$ moanákɛ́ kaɣokó roːcíːně$^!$
        person SM-gave-PST.COMPL boy   chicken this.morning
        'Someone gave the boy a chicken in the morning.'   *underlying tones*

    b. mondo$^!$ ahɛírɛ́ moanákɛ́ káɣó$^!$kó roːcíːně$^!$ *tones after !-shift*

What is important for the present discussion is that in an affirmative declarative sentence,
the verb bears a downstep suffix, represented by the exclamation mark superscript.[25] Ac-
cording to Clements and Ford (1979), the downstep is a free floating tone that shifts to
other words and influences the tones of the following elements. Phonetically, it triggers
pitch lowering, i.e. the following tones are realized in a lower register. The downstep usu-
ally shifts to the immediately following word, unless this word ends in a high tone; in this
case, the downstep shifts even further to the right. If there is no element to the right, the
downstep is realized on the verb itself. The shifted downstep causes preceding underlying

---

[25]See Clements (1984*a*: 40, 53) and Clements (1984*b*: 312ff.) for details on the distribution of the down-
step, in particular on its absence under negation. For an overview of the morphology and phonology of
verbs in Kikuyu see Clements (1984*b*).

low tones to change into high tones, the process is blocked by underlying high tones. In (55-a), with underlying tones, there are several downsteps. We are interested in the one associated with the verb, the others are irrelevant for the purposes of this discussion. The verbal downstep shifts to the direct object 'chicken' (after the sequence of low tones in that element) because the closer indirect object ends in a high tone. As a consequence, the preceding underlying low tones of the direct object become high. This is shown in (55-b). In sentence-final position, the downstep has no effect on following words because the shift is clause-bound. Nevertheless the downstep is detectable because it blocks a general rule that otherwise applies in final position, namely the change of high to low tone, cf. Clements et al. (1983: 8).[26]

In an Ā-dependency, the operator occurs in clause-initial position and gets a focus prefix, compare the declarative sentence in (56-a) with the corresponding subject question in (56-b).

(56)   *Question formation in Kikuyu (Clements 1984a: 39)*:
    a. Kariokĭ á-ˈtɛ́m-írɛ́   mo-tě**ꜜ**
         Karioki SM-cut-PST.COMPL CL-tree
         'Karioki cut a tree.'   *declarative*
    b. nó-o$_k$ ___$_k$ o-tɛm-irɛ́   mo-te
         FOC-who   SM-cut-PST.COMPL CL-tree
         'Who cut a tree?'   *SU question*

Crucially, in case of wh-movement, relativization, and focus movement, the postverbal downstep is deleted and hence, it does not affect the following tones (Clements 1984*a*: 38). Look at the examples in (57). Note that a rule of question intonation forces a final rising tone to become a falling tone.

(57)   *Downstep deletion under wh-movement Clements et al. (1983: 9)*:
    a. nó-o$_k$ ___$_k$ o-hɛ-irɛ́   moanákɛ́ káɤokó ro:cîːnê
         FOC-who   SM-give-PST.COMPL boy   chicken this.morning
         'Who gave the boy a chicken this morning?'   *SU extraction*
    b. [né káɤokó karékó]$_k$ móndoˈ a-hɛ-íˈrɛ́   moànàkɛ ___$_k$
          FOC chicken which   person SM-gave-PST.COMPL boy
         'Which chicken did someone give to the boy?'   *DO extraction*
    c. [né ré   ]$_k$ móndoˈ a-hɛ-íˈrɛ́   moanákɛ́ káɤòkò ___$_k$
          FOC when   person SM-gave-PST.COMPL boy   chicken
         'When did someone give the boy a chicken?'   *ADJ extraction*

It the subject question in (57-a) the downstep is absent. This becomes clear when we compare it with the corresponding declarative sentence in (55-a): Without wh-movement, the downstep shifts to the direct object and turns the underlying low tones of its first two syllables into high syllables. But this is not the case in (57-a). Here, the second syllable of the word 'chicken' has low tone. The same is true for (57-c) where the adjunct is questioned. In (57-b), the direct object is questioned. Without Ā-movement, the downstep would have shifted to the sentence-final position and there it would have blocked the change of underlying high to low tones – a rule that regularly applies in sentence-final position. However, *moanake*, which has a sequence of underlying high tones (cf. (55-a)), now has low tones, showing that the rule applying in final position has applied and is not blocked by the downstep.

---

[26]Kikuyu has a number of other phonological processes in addition that will be ignored in what follows; for details, see Clements (1984*a*: 40-41) and references cited there.

Downstep deletion is indeed a reflex of movement. First of all, the dependencies that trigger this deletion are subject to the classic island constraints. I illustrate this for wh-movement, but the same holds for relativization and focus movement.[27] Wh-movement is in principle unbounded (cf. (58-a)), but is subject to the complex NP constraint (cf. (58-b)) and the wh-island constraint (cf. (58-c)):

(58)   *Island-sensitivity of wh-movement (Clements 1984a: 41-42):*[28]

    a.   [$_{CP}$ nó-o$_k$ ó-ɣw-eciíri-a   [$_{CP}$ Ngoɣe aúɣ-írɛ   [$_{CP}$ áte
           FOC-who SM-RM.PST-think-RM.PST   Ngoɣe SM-say-PST.COMPL   that
           Kama.ú a-ón-ꞌirɛ́   ___$_k$]]]
           Kama.ú SM-see-pst.COMPL
           'Who do you think Ngūgī said that Kamau saw?'

    b. *[$_{CP}$ nó-o$_k$ Kama.ú a-ón-ꞌírɛ́   [$_{NP}$ mó-ndo [$_{CP}$ o-reǎ
           FOC-who Kama.ú SM-see-PST.COMPL   CL-person   SM-DEM
           ó-ríng-ꞌírɛ́   ___$_k$]]]
           SM-hit-PST.COMPL
           'Who did Kamau see the person (that) hit?'   *CNP island*

    c. *[$_{CP}$ nó-o$_k$ Ngoɣe a-écirí-riɛ́   hiíhi [$_{CP}$ nó-o
           FOC-who Ngūgī SM-wonder-PST.COMPL   FOC-who
           o-ɔn-iꞌrɛ́   ___$_k$]]
           SM-see-PST.COMPL
           'Who did Ngūgī wonder who saw?'   *wh-island*

Furthermore, Kikuyu also allows for wh-in situ questions.[29] In these, none of the morphological changes associated with wh-ex situ questions occurs: The wh-word cannot be preceded by the focus particle and the downstep is not deleted (Clements 1984*a*: 42). Look at (59) which is the in-situ counterpart of the adjunct question in (57-c). It is clear that the downstep is not deleted because it causes the same changes as in the corresponding declarative sentence (55-a), i.e. it shifts to the direct object and changes the low tones of the object into high tones.

(59)   *Wh-in-situ in Kikuyu (Clements et al. 1983: 9):*

    mondoꞌ a-hɛ-írɛ́   moanákɛ́ káɣóꞌkò rè
    person SM-gave-PST.COMPL boy   chicken when
    'When did someone give the boy a chicken?'   *wh-ADJ*

Finally, Kikuyu also allows for partial wh-movement. Under partial wh-movement, only the clauses crossed by overt movement show deletion of the downstep, but structurally higher clauses do not. Compare the examples in (60) (the wh-phrase is set in bold). In (60-a) we have full wh-movement to the scope position; in (60-b) and (60-c) the wh-word undergoes partial movement (but it takes matrix scope):

---

[27]As in many other languages, the island violations can be circumvented by using a resumptive pronoun instead of a gap in the base position of the element in SpecCP. For examples, see Clements (1984*a*: 42).

[28]Clements does not provide a gloss for the word *hiíhi* in (58-c).

[29]There is one exception: Subjects are obligatorily extracted, i.e. there are no subject in-situ questions. Besides wh-in-situ, Kikuyu also has a focus-in-situ construction. As with wh-in-situ, it does not exhibit any of the morphological reflexes of overt focus-/wh-movement, cf. Clements (1984*a*: 43f.).

(60) *Partial wh-movement and downstep deletion (Clements 1984a: 47)*:

a. [$_{CP_1}$ **nó-o**$_k$ ó-ɣw-eciíri-a   [$_{CP_2}$ Ngoɣe aúɣ-írɛ   [$_{CP_3}$
    FOC-who SM-RM.PST-think-RM.PST   Ngūgī SM-say-PST.COMPL

    áte ___$_k$ o-ɔn-írɛ́   Kaanakɛ]]]
    that   SM-see-PST.COMPL Kanake   *full movement*

b. [$_{CP_1}$ ó-ɣw-ˈeciri-á   [$_{CP_2}$ **nó-o**$_k$ Ngóɣe aúɣ-írɛ   [$_{CP_3}$
    SM-RM.PST-think-RM.PST   FOC-who Ngūgī SM-say-PST.COMPL

    áte ___$_k$ o-ɔn-írɛ́   Kaanakɛ]]]
    that   SM-see-PST.COMPL Kanake   *partial movement*

c. [$_{CP_1}$ ó-ɣw-ˈeciri-á   [$_{CP_2}$ Ngóɣe aúɣ-írɛ́   [$_{CP_3}$ ate
    SM-RM.PST-think-RM.PST   Ngūgī SM-say-PST.COMPL   that

    nó-o$_k$ ___$_k$ o-ɔn-írɛ́   Kaanakɛ]]]
    FOC-who   SM-see-PST.COMPL Kanake
    'Who do you think Ngūgī said saw Kanake?'   *partial movement*

Let us begin with (60-c): We see that wh-movement has taken place in the lowest clause because the downstep is deleted: The rule that lowers high tones in sentence-final position has applied to *Kanake* (underlyingly, it ends in high tones: Kanákɛ́); if the downstep was present, it would have shifted from the verb to *Kanake* and would have blocked the sentence-final lowering rule. We see that the same holds in (60-a) and (60-b) where the lowest clause is also crossed by movement. Comparing (60-a) on the one hand and (60-b) and (60-c) on the other hand, we observe a tonal difference on the matrix verb: In (60-a) the downstep on the verb is deleted because wh-movement targets matrix SpecCP, but this downstep is present in (60-b) and (60-c) because the matrix clause is not affected by movement in the latter examples. With respect to CP$_2$, we see that movement has affected this clause in (60-a) and (60-b) when we compare it with (60-c): The downstep from the verb of CP$_2$ spreads into CP$_3$, behind the complementizer *ate*. One effect of this downstep shift is that the complementizer changes its low tone into a high tone. This process is absent in (60-c) where CP$_2$ is not the target of wh-movement.

To summarize, these facts suggest that downstep deletion is a reflex of movement that follows pattern I. Although tonal reflexes of movement are barely described, Kikuyu is not the only language with such a reflex: In Akɔɔse (Niger-Congo, Cameroon) non-subject extraction triggers the occurrence of a floating high tone prefix on the verb (Zentz 2013).

2.3.1.6 Inversion in Spanish

In Spanish (Indo-European, Spain), we find a syntactic reflex of Ā-movement: subject-verb-inversion. The basic word order is SVO. Under inversion, the verb raises past the subject. Inversion is obligatory if an operator moves to SpecCP (e.g. in questions, exclamatory sentences and comparatives). In addition, there is an argument-adjunct-asymmetry: Inversion only applies if an argument is Ā-moved, but not if an adjunct is extracted (Torrego 1984: 106). (61-a) vs. (61-b) and (61-c) vs. (61-d) illustrate inversion in matrix and embedded questions, respectively.

(61) *Invershion in Spanish (Torrego 1984: 103-104)*:

a. Qué$_k$ **querían**   esos dos ___$_k$
    what want.3PL.PST DEM two
    'What did those two want?'   *matrix question, inversion*

  b. *Qué$_k$ esos dos querían ___$_k$
    what DEM two want.PST
    'What did those two want?'   *matrix question, no inversion*

  c. No sabía   [$_{CP}$ qué$_k$ **querían** esos dos ___$_k$]
    NEG know.1SG.PST what want.3PL.PST DEM two
    'I didn't know what those two wanted.' *embedded question, inversion*

  d. *No sabía   [$_{CP}$ qué$_k$ esos dos querían ___$_k$]
    NEG know.1SG.PST what DEM two want.3PL.PST
    'I didn't know what those two wanted.' *embedded question, no inversion*

There is massive evidence that this is a reflex of movement: The dependencies that trigger inversion are island-sensitive. This is illustrated in (62) for wh-movement from a wh-island (cf. (62-a) and (62-b) (see also Rivero 1980)) and from a complex NP island (cf. (62-c)):[30]

(62) *Island effects in Spanish (Torrego 1984: 115, Rivero 1980: 380, Rivero 1978: 514)*:

  a. *[$_{CP}$ Cuánto$_k$ no sabes [$_{CP}$ quién pesa ___$_k$]]
    how.much NEG know who weighs
    'How much don't you know who weighs?'   *wh-island*

  b. *[$_{CP}$ Qué$_k$ preguntan [$_{CP}$ (que) quién tiene ___$_k$]]
    what ask.3PL   (that) who has
    'What do they ask who has?'   *wh-island*

  c. *[$_{CP}$ Qué$_k$ aceptas [$_{NP}$ la pretensión de [$_{CP}$ que no tienen ___$_k$]]]
    what accept.2SG the pretension of that NEG have.3PL
    'What do you accept the pretension that they do not have?' *CNP island*

Furthermore, they induce weak cross-over effects:

(63) *Weak cross-over effects in Spanish (García Mayo 1995: 6, Sũner 1988: 421)*:

  a. *[A quién$_i$]$_k$ visitó su$_i$ madre ___$_k$
    to whom visited his mother
    'Who did his mother visit?'

  b. *[A   cuales de ellos$_i$]$_k$ no aguanta ni su$_i$ madre ___$_k$
    which ones of them not stand even their mother
    'Which one of them can't even their mother stand?'

Hence, the dependency involves movement. Furthermore, the reflex is clearly tied to overt movement, as a number of facts suggest: First, inversion is neither required if the wh-word stays in-situ nor in yes-no questions, cf. (64):

(64) *No inversion without wh-movement (Torrego 1984: 104)*:

  a. Marta quiere qué
    Marta wants what
    'Marta wants what?'   *wh-in-situ*

  b. Marta quiere café
    Marta wants coffee
    'Does Marta want coffee?'   *yes-no-question*

---

[30]For some reason, wh-movement from an embedded interrogative clause is possible in Spanish if the subject of that clause is extracted. For discussion of this violation of the wh-island constraint see Torrego (1984: sec.3).

Furthermore, if extraction takes place from the matrix clause, inversion only applies in this clause, but not in the embedded clause(s) that are not affected by movement:

(65)   *Inversion in Spanish under wh-movement from the matrix clause (Torrego 1984: 110)*:

[$_{CP}$ [A quién]$_k$ **prometió** Juan ___$_k$ [$_{CP}$ que Pedro se   encargaría   de [$_{CP}$ que
     to whom   promised Juan   that Pedro REFL be.in.charge of   that
la gente sacara las entradas a tiempo]]]
the people buy   the tickets   on time
'To whom did Juan promise that Pedro would be in charge of the people buying their tickets on time?'

Having established that inversion in Spanish is a reflex of overt movement, let us now look at long wh-extraction. In this context, inversion obligatorily applies in every clause between the base position of the wh-phrase and its landing site. (66-a) provides the base line declarative sentence and (66-b) and (66-c) show that inversion must apply along the whole movement path (note that the auxiliary and the participle are treated as a complex head: both undergo inversion).

(66)   *Long extraction and inversion in Spanish (Torrego 1984: 108, 109)*:

a. [$_{CP}$ Juan pensaba   [$_{CP}$ que Pedro le   había   dicho [$_{CP}$ que la
        Juan think.3SG.PST   that Pedro him have.3SG.PST told   that the
   revista había   publicado ya   el artículo]]]
   journal have.3SG.PST published already the article
   'John thought that Peter had told him that the journal had published the article already.'   *declarative*

b. [$_{CP}$ Qué$_k$ **pensaba**   Juan [$_{CP}$ que le   **había**   **dicho** Pedro [$_{CP}$ que
        what think.3SG.PST Juan   that him have.3SG.PST told   Pedro   that
   había   **publicado** la revista ___$_k$]]]
   have.3SG.PST published the journal
   'What did John think that Peter had told him that the journal had published?'
                                                            long wh-movement, inversion

c. *[$_{CP}$ Qué$_k$ pensaba   Juan [$_{CP}$ que Pedro le   había   dicho [$_{CP}$ que
        what think.3SG.PST Juan   that Pedro him have.3SG.PST told   that
   la revista había   publicado ___$_k$]]]
   the journal have.3SG.PST published
   'What did John think that Peter had told him that the journal had published?'
                                                         long wh-movement, no inversion

Thus, inversion is a reflex of movement that follows pattern I (for variation on the distribution see Baković 1998). Torrego (1984: 108-109) takes this as evidence that wh-movement applies successive-cyclically through every embedded SpecCP on its way to the matrix SpecCP.

Subject-verb-inversion is also attested as a reflex of movement in other Romance languages such as French (Kayne 1972; Kayne and Pollock 1978), and in Belfast English (Henry 1995). As in Spanish, inversion in French is tied to overt movement: It is impossible without extraction, i.e. in wh-in-situ contexts and in clauses that are not affected by movement. Under cross-clausal Ā-movement, inversion applies to all verbs along the path of movement. Hence, inversion in French is also a pattern I reflex. We will see examples of inversion from Belfast English in section 4.2.2.

2.3.2   Pattern II reflexes

We have already seen an instance of a pattern II reflex: Complementizer agreement in Wolof alternates between patterns I and II. In this section, I present further examples of pattern II from complementizer agreement in Chamorro, *no*-marking in Duala, pronoun choice in Ewe, and focus marking in Indonesian. The latter involves replacement, the others involve addition. Other languages with pattern II that are not discussed here include complementizer agreement in Bùlì (Niger-Congo, Ghana, Hiraiwa 2005), relative tense in Hausa (Afro-Asiatic, Nigeria, Tuller 1986) and Coptic (Afro-Asiatic, Egypt, Reintges and Green 2004; Reintges et al. 2006), Moore (Haïk et al. 1985; Haïk 1990), and Haitian Creole (French based Creole, Takahashi and Gracanin Yuksek 2008).

2.3.2.1   C-agreement in Chamorro

We have seen in section 2.3.1.1 that Chamorro (Austronesian, Guam and Mariana Islands) has a pattern I reflex that involves replacement of the regular subject-verb-agreement morphology: wh-agreement on the verb. Furthermore, Chamorro exhibits another reflex of movement on the complementizer. In contrast to wh-agreement, complementizer agreement follows pattern II and results in addition of morphology to the C-head.

In sentences without operator movement, the choice of the complementizer depends on (i) the finiteness of the sentence (finite [+fin] vs. non-finite [–fin]), (ii) the clause type (interrogative [+Q] vs. non-interrogative [–Q]) and (iii) the embedding status (root [+root] vs. embedded [–root]). Infinitives and finite root declarative clauses do not have an (overt) complementizer whereas finite embedded declarative sentences are introduced by *na*; finite interrogative clauses (yes-no-questions) are introduced by *kao*, regardless of their embedding status, although this morpheme is optional in root clauses.

(67)   *Chamorro complementizers in clauses without operator movement (Chung 1998: 223)*:

syntactic context	form of C
[–fin] / [+fin, +root, –Q]	Ø
[+fin, –root]	*na*
[+fin, +Q]	*kao*

Under Ā-extraction, however, the form of the complementizer is determined by different factors: Its form does not depend on the properties of the clause but rather on those of the phrase in SpecCP. The relevant properties of the Ā-moved phrase are (i) its category (NP [+N] vs. PP [–N]), (ii) whether it denotes a location in time or space ([±locat]), (iii) whether it is the null relative operator ([+Rel]) or not ([–Rel]) (Chung 1998):

(68)   *Chamorro complementizers in clauses with operator movement (Chung 1998: 224)*:

syntactic context	form of C
[+N, –locat]	Ø
[+N, +locat, +Rel]	*änai*
[+N, +locat, –Rel]	*na* (Guam) or **nai/ni** (Saipan)
[–N]	*na* (Guam) or *nai/ ni* (Saipan)

Since the form of the complementizer co-varies with the properties of the Ā-moved XP in SpecCP, we can conclude that there is agreement between C and XP (e.g. in category), and we can exclude a potential alternative analysis of the alternation according to which the complementizer in Ā-dependencies is simply an interrogative C-head that has a different realization than a non-interrogative C-head. I illustrate some of the forms in (69):

(69) *Complementizer choice in sentences without operator movement (Chung 1998: 224-226)*:

    a.   Hafa$_k$ **Ø** malago'-mu   ___$_k$
          what C WH.OBL.want-AGR
          'What do you want?'   *[+N, –locat]*

    b.   Na'tungu' yu' [$_{CP}$ hafa$_k$ **Ø** malago'-mu   ___$_k$]
          make.know me   what C WH.OBL.want-AGR
          'Let me know what you want?'   *[+N, –locat]*

    c.   [Ginin hayi]$_k$ **na** un-risibi   katta ___$_k$
           from who   C AGR-receive letter
          'From whom did you receive a letter?'   *[–N]*

In the matrix question in (69-a) the complementizer is zero because the wh-word is a noun that denotes the patient and not a location. The same holds for the embedded question in (69-b). In (69-c) the fronted phrase is a PP and the complementizer is *na* (in the Guam dialect). As these data show, complementizer agreement co-occurs with wh-agreement; both reflexes are triggered by a the same Ā-dependency. I have already shown in section 2.3.1.1 that the respective Ā-dependencies involve movement (they are island-sensitive and show cross-over effects). Consequently, I will not repeat the data here.

    Interestingly, the complementizer alternation in Chamorro only affects the highest complementizer in the dependency: In the context of long-distance extraction only the C head in whose specifier the Ā-moved phrase lands alternates. The C heads in all other clauses crossed by movement are realized according to the constraints that regulate complementizer choice in sentences without Ā-movement, i.e. depending on the finiteness, clause type and the embedding status of the clause, cf. (70).

(70) *Complementizer choice in sentences with operator movement Chung (1998: 229f.)*:

    a.   Ngai'an$_k$ **nai** un-hassúsuyi   [$_{CP}$ **Ø** d*um*ingu Sa'ipan ___$_k$]
          when   C WH.OBJ.AGR-think.of.PROG   C INF.leave Saipan
          'When are you thinking of leaving Saipan?'

    b.   [Manu na lepblu]$_k$ **Ø** malagu'ñiha   [$_{CP}$ **na** u-taitai   ___$_k$]
          which L book   C WH.OBL.want-AGR   C WH.obj.AGR-read
          'Which book do they want to read?'
          Lit: 'Which book do they want that they should read?'

In (70-a) the wh-phrase originates in the embedded clause and triggers the operator extraction form *nai* on the complementizer in the matrix clause (because the fronted phrase is [–N]); but in the embedded clause we find the zero complementizer used with non-finite declarative clauses. If the embedded complementizer would also Agree with the moving wh-phrase, we would expect the same form *nai* as in the matrix clause. In (70-b), the wh-phrase is an NP that does not denote location or time in space, hence the complementizer chosen is the zero complementizer, cf. (68). But in the embedded clause we find the complementizer *na* used for embedded finite declarative clauses. Thus, Chamorro complementizer agreement exhibits pattern II.

2.3.2.2 *no*-marking in Duala

Duala (Niger-Congo, Cameroon) also shows a pattern II reflex of movement under Ā-movement. The basic word order in Duala is subject – verb – indirect object – direct object – adjuncts. If an XP is Ā-moved (focussed, questioned, relativized), it is moved

to the clause-initial position SpecCP, preceding the subject (Epée 1975: 211). In case a non-subject XP is displaced, the invariant marker *no* must be inserted after the first verbal element of the clause, which is an auxiliary or the main verb in the absence of auxiliaries. The marker does not show up if the subject of a clause is extracted (cf. Epée 1975; 1976*a*,*b*). This is illustrated in (71) for focus movement; the focussed constituent is followed by the focus marker *nde*.[31] (71-a) provides a declarative sentence with the basic word order. (71-b) – (71-e) illustrate focussing of various constituents of this clause. Note that the fronted constituent must be followed by the marker *nde*. In (71-b), the subject is topicalized and hence, the *no* particle does not occur in the clause. In (71-c) – (71-e), non-subjects are topicalized, which obligatorily leads to insertion of *no* after the verb.

(71)   *Focus movement in Duala (Epée 1976*b*: 194)*:

    a.   Kuo a   bodi nu moto kalati kiele
           Kuo 3SG give that man book yesterday
           'Kuo gave a book to that man yesterday.'   *declarative*

    b.   Kuo$_k$ nde ___$_k$ a   bodi nu moto kalati kiele
           Kuo FOC   3SG give that man book yesterday
           'It's Kuo who gave a book to that man yesterday.'   *SU extraction*

    c.   [nu   moto]$_k$ nde Kuo a   bodi **no** ___$_k$ kalati kiele
             that man   FOC Kuo 3SG give NO   book yesterday
           'It's that man Kuo gave a book to yesterday.'   *IO extraction*

    d.   kalati$_k$ nde Kuo a   bodi **no** nu moto ___$_k$ kiele
           book   FOC Kuo 3SG give NO that man   yesterday
           'It's a book Kuo gave to that man yesterday.'   *DO extraction*

    e.   kiele$_k$ nde Kuo a   bodi **no** nu moto kalati ___$_k$
           yesterday FOC Kuo 3SG give NO that man book
           'It's yesterday that Kuo gave a book to that man.'   *ADJ extraction*

The presence of *no* under subject extraction and its absence under non-subject extraction leads to ungrammaticality. This is shown in (72) for wh-movement:

(72)   *Wh-movement in Duala (Epée 1976*b*: 194, Biloa 1993: 69)*:

    a.   *nja$_k$ ___$_k$ a   bodi **no** Kuo moni
           who   3SG give NO Kuo money
           'Who gave Kuo the money?'   *SU question*

    b.   *nja$_k$ o   bodi moni   ___$_k$
           who you give money
           'Who did you give the money to?'   *IO question*

no-marking is a reflex of overt movement: First of all, the relevant Ā-dependencies are island-sensitive. This is illustrated in (73) for relativization from of a complex NP island and a wh-island. So far, we have only looked at wh-movement and focus movement, therefore (73-a) shows that relativization also triggers no-marking and is in principle unbounded.

(73)   *Island-sensitivity in Duala (Biloa 1993: 70f.)*:

    a.   moto$_i$ [$_{CP}$ nyena$_i$ na mongole **no** [$_{CP}$ na o   kwadi [$_{CP}$ na o   wen ___$_k$]]]
           man   who I think   NO   that you say   that you see
           'the man who I think that you said that you saw'

---

[31]In Duala, the 3rd person subject always co-occurs with a preverbal pronoun indicating the noun class of the subject, glossed as '3SG' in what follows.

b. *moto$_i$ [$_{CP}$ [nyena$_i$]$_k$ na neimbi **no** [$_{NP}$ mbo$_j$ [$_{CP}$ [nyenat$_j$]$_n$ ___$_n$ e kuko
man   who   I recognize NO dog   which   SM bit
___$_k$]]]
'the man who I recognized the dog which bit (him)'   *CNP island*

c. *moto$_i$ [$_{CP}$ [nyena$_i$]$_k$ na nyaka   **no** [$_{CP}$ [nja mmuto]$_n$ ___$_n$ bai
man   who   I am.astonished NO   who woman   married
___$_k$]]
'the man who I wonder which woman married (him)'   *wh-island*

Furthermore, Duala allows for wh-in-situ questions. In this context, the marker *no* can never show up after the first verbal element, even if a non-subject is extracted (cf. Epée 1975: 215-216). Compare the wh-ex-situ questions in (74-a) and (74-c) with their wh-in-situ counterparts in (74-b) and (74-d).[32]

(74)   *Wh-movement in Duala (Epée 1976b: 194-195)*:

a. [njika ponda]$_k$ Kuo a po   **no** ___$_k$
which time   Kuo 3SG come NO
'At what time will Kuo arrive?'   *wh-ADJ ex-situ*

b. Kuo a po njika ponda
Kou 3SG come which time
'At what time will Kuo arrive?'   *wh-ADJ in-situ*

c. [njika lambo]$_k$ nu muna a   boli **no** ___$_k$
which thing   that child 3SG do NO
'What did that child do?'   *wh-DO ex-situ*

d. nu muna a   boli njika lambo
that child 3SG do   which thing
'What did that child do?'   *wh-DO in-situ*

In an embedded question where wh-movement of a non-subject targets the embedded SpecCP, *no* can only show up in the clause affected by movement, viz. in the embedded clause. This is illustrated in (75-a) for focus movement and in (75-b) for wh-movement.

(75)   *Extraction in embedded clauses (Epée 1976b: 197)*:

a. [$_{CP}$ na si bi   [$_{CP}$ nga wenge$_k$ nde Kuo a   ben **no** kekise ___$_k$]]
I   not know   if   today FOC Kuo 3SG have NO exam
'I don't know if it's today that Kuo has an exam.'   *focus movement*

b. [$_{CP}$ baise Kuo [$_{CP}$ nje$_k$ a   pula **no** ___$_k$]]
ask   Kuo   what 3SG want NO
'Ask Kuo what he wants.'   *wh-movement*

To conclude, no-marking only occurs if overt movement has applied. One might wonder whether the absence of *no*-marking under local subject extraction is due to the fact that the subject is not extracted at all but stays in SpecTP (there is evidence from several languages that this is the case for wh-subjects). In Duala, however, this is not a possible

---

[32]Duala also has a focus-in-situ strategy, see Epée (1975: 212) for data. As with wh-in-situ, the particle *no* cannot appear with focus-in-situ, further indicating that it is a pure reflex of overt movement. The focus marker *nde*, however, is not a reflex of movement; it also occurs with focus-in-situ after the verb, so it just indicates that the sentence contains a focussed element. In this respect, Duala contrasts with Indonesian and Malay. In these languages, the focus marker only occurs under overt extraction, but not with in-situ phrases of the relevant type, i.e. wh-words.

explanation. Epée (1976*a*: 150) provides evidence that the subject can move, at least op-
tionally: If the subject is an operator of a certain class, it can trigger a special form of the
preverbal pronoun that agrees in noun class with the subject; instead of the regular form
a that also occurs in a declarative sentence with extraction, we can also find the special
form *nu* that obligatorily occurs under subject relativization. Epée takes the latter to be a
reflex of movement.[33] Note that in both cases, there is no *no*-marking; hence, the absence
of *no* in local subject extraction does not correlate with the absence of movement – it can
be absent even if movement (of the subject) has taken place.

(76)   *Local subject extraction (Epée 1976*a*: 150-151)*:

   a.   nja   a   pgi
        who 3SG come

   b.   nja$_k$ ___$_k$ nu   pqi
        who   3SG.REL come
        'Who came?'

Let us now look at long-distance extraction. If a non-subject is extracted from an em-
bedded clause, the particle *no* only surfaces in the clause where the Ā-moved XP lands;
it can neither show up after the first verbal element of the clause in which the moved XP
originates nor in lower clauses crossed by movement. (77-a) provides an example of long
object topicalization, (77-b) of long object wh-movement and (77-c) of long relativization.

(77)   *Long-distance extraction in Duala (Epée 1976*b*: 196, Biloa 1993: 69-70)*:

   a.   [$_{CP}$ ni   kalati$_k$ nde na ta   **no** na kwalane Kuo [$_{CP}$ na   a-angamente wana
        that book   FOC I   PST NO I   tell   Kuo   that 3SG-must   bring
        ___$_k$]]
        'That's the book I told Kuo that he should bring.'   *DO focus movent*

   b.   [$_{CP}$ [njika buna]$_k$ o   ta   **no** o   kwalane mba [$_{CP}$ na   o   mende timba
        which day   you PST NO you tell   me   that you FUT   return
        ___$_k$]]
        'When did you tell me that you would return?'   *ADJ wh-movent*

   c.   moto$_k$ [$_{CP}$ nyena$_i$ na mongele **no** [$_{CP}$ na   o   kwadi [$_{CP}$ na   o   wen ___$_k$]]]
        man   who   I think   NO   that you say   that you see
        'the man who I think that you said that you saw'   *DO relativization*

(77-c) clearly demonstrates that it is not just the case that *no* does not occur in the clause
from which the moved XP originates; instead, it shows that *no* does not appear in any
clause but the one hosting the moved XP.

    Recall that under local subject extraction, *no* cannot surface. Under long extraction
of an embedded subject, however, *no*-marking is triggered in the clause where the subject
lands, cf. (78). Thus, we have a split in subject extraction between local and long-distance
Ā-movement with respect to *no*-marking.

(78)   *Long-subject extraction (Epée 1976*a*: 158)*:

   a.   [$_{CP}$ nja$_k$ wa pula **no** [$_{CP}$ na   a ye   ___$_k$], mba]
        who you want NO   that he come   me
        'Who do you want to come, me?'

---

[33]I will not consider the reflex on the preverbal pronoun in any more detail at this point. See section 3.2.1
for a related phenomenon.

b. *[_CP_ nja_k_ wa pula [_CP_ na a ye   ___ _k_]_,_ mba]
    who you want   that he come   me
    'Who do you want to come, me?'

To summarize, the distribution of *no* in Duala follows pattern II: It occurs only in the clause that contains the moved phrase, but not in any lower clause (cf. Epée 1976*b*: 196, Biloa 1993: 72).

The question might arise at this point, what kind of element *no* is. I will briefly address this issue here. One could wonder whether *no* is a resumptive pronoun. There is strong evidence that is is not: (i) Epée (1976*a*: 155-156) shows that pronouns in Duala agree in noun class with their referring NP, but the particle *no* is invariant for the noun class of the extracted element. (ii) *no* does not occupy the base position of the extracted element. This is evident when we compare (71-a) with (71-d). In the former, a declarative sentence, the indirect object immediately follows the verb and the direct object follows the indirect object. In the latter, the direct object is extracted, but *no* occurs immediately right-adjacent to the verb and before the indirect object, although the base position of the direct object is after the direct object. Instead, it seems to be a clitic. It always follows the first verbal element in the clause. This must not be the main verb, but rather the structurally highest verbal element in the clause, cf. (77-a) where it follows the past auxiliary *ta*. (iii) Biloa (1993: 70) adds that *no* co-occurs with non-referring expressions, which also argues against the hypothesis that it is a resumptive pronoun. Biloa (1993: 72-73) concludes from these facts that *no* is a clitic to the head T that attaches to the right of the verbal element in T, i.e. to an auxiliary or to the main verb which moves to T in the absence of an auxiliary. I will adopt his conclusion.

2.3.2.3   Pronoun choice in Ewe

In Ewe (Niger-Congo, Ghana, Togo), the morphological reflex of movement manifests itself in the form of the 3rd person subject pronoun. Its default form is *é*; if, however, an XP undergoes Ā-movement (wh-movement, focus movement, relativization) through SpecCP, the form of this pronoun changes to *wò* (the acute is used for high tone, the grave for low tone). Ā-movement involves preposing of a phrase to the clause-initial position; in case of focus movement, the focussed phrase is followed by the focus marker ɛ. The examples in (79) illustrate the pronoun choice. In a matrix declarative sentence with basic subject – verb – object word order, the 3rd person subject pronoun can only surface as *é*, cf. (79-a). The same holds for the 3rd person subject of an embedded declarative clauses, cf. (79-b). The change to *wò* under Ā-movement is illustrated in (79-c) for wh-movement to SpecCP of the embedded clause and in (79-d) for relativization (movement of a silent operator).

(79)   *Subject pronoun choice un Ewe, local movement Collins (1993: 157, 177)*:

a.   [é/*wò] fo Kɔsi
      he   hit Kɔsi
     'He hit Kɔsi.'   *declarative*

b.   Kofi gblɔ be   [é/*wò] fo Kɔsi
     Kofi said that   he   hit Kɔsi
     'Kofi said that he hit Kɔsi.'   *embedded declarative*

c.   [_CP_ Kofi biɛ   [_CP_ be lamata_k_ [*é/**wò**] fo Kɔsi ___ _k_]]
          Kofi asked   C why   he   hit Kɔsi
     'Kofi asked why he hit Kɔsi.'   *embedded question*

    d. ga-xe-me   [*é/**wò**] va
          time-in-which he   came
          '(at the time) when he came'   *relativization*

In case of long-distance extraction of an XP to matrix SpecCP, the 3rd person subject pronoun of the embedded clause from which the moved XP originates can be either in the default form *é* or in the special form *wò*. This is illustrated for long object extraction from an embedded clause with focus movement in (80-a) and with wh-movement in (80-b).

(80)   *Long-distance movement in Ewe (Collins 1993: 179)*:

    a. [$_{CP}$ [Kɔsi ɛ   ]$_k$ me gblɔ [$_{CP}$ be [é/**wò**] fo ___$_k$]]
         Kɔsi FOC   I said   C he   hit
         'Kɔsi, I said that he hit.'   *DO focus movement*

    b. [$_{CP}$ me$_k$ e   gblɔ [$_{CP}$ be [é/**wò**] fo ___$_k$]]
         who you say   C he   hit
         'Who did you say that he hit?'   *DO wh-movement*

I will come back to the optionality in these examples in section 2.5.3.2. For the moment is is only important that long-distance movement does not have to lead to the special form in the embedded sentence. If it does not, we get pattern II: The change in the form of the pronoun does not occur in clauses along the movement path in which the extracted XP does not surface.

    The Ā-dependencies that trigger the special form of the subject pronoun involve movement because they are sensitive to islands. This is illustrated in (81) for focus movement from a complex NP island.

(81)   *Island-sensitivity (Collins 1993: 182)*:

    *Kofi$_k$ ɛ   me nya   [$_{NP}$ ame$_i$ [$_{CP}$ [xe$_i$] fo ___$_k$]]
     Kofi FOC I know   person   which hit
     'As for Kofi, I know the person who hit (him).'   *CNP island*

Ewe does not allow for wh-in-situ or partial wh-movement, so we cannot test how the pronoun forms would be distributed in these contexts. But there is evidence that the pronoun form is affected by overt movement across it: If an XP that originates in the matrix clause undergoes Ā-movement, it cannot affect the form of the pronoun in the embedded clause, i.e. the embedded 3rd person subject must occur in the default form *é*. This is illustrated for focus movement of the subject in (82-a) and for movement of the object of a preposition in (82-b).

(82)   *Extraction from the matrix clause (Collins 1993: 180, 181)*:

    a. [$_{CP}$ Kofi$_k$ ɛ   gblɔ ___$_k$ [$_{CP}$ be [é/*wo] fo Kɔsi]]
         Kofi FOC said   C he   hit Kɔsi
         'It is Kofi who said that he hit Kɔsi.'

    b. [$_{CP}$ [Kofi gbɔ]$_k$ ɛ   me se   le   ___$_k$ [$_{CP}$ be [é/*wo] fo Kɔsi]]
         Kofi near FOC I heard from   C he   hit Kɔsi
         'It is from Kofi that I heard that he hit Kɔsi.'

The generalization about the distribution of the special subject pronoun *wò* in Ewe is that it obligatorily occurs only in the clause where the Ā-moved phrase lands; in lower clauses clauses along the path of movement the change in the subject form of 3rd person pronouns is only optional. If the special form *wò* is realized in the non-final clauses

as well, we have an instance of pattern I. If, however, the default form é is used in non-final clauses, Ewe exhibits pattern II: The reflex of movement occurs solely in the topmost clause of the Ā-movement.[34]

2.3.2.4 Focus marking in Indonesian

In addition to *men*-deletion that follows pattern I, Indonesian also exhibits a pattern II reflex of Ā-movement: focus marking by *yang*. It is triggered in the same Ā-dependencies as *men*-deletion, viz. the two reflexes co-occur. We have already seen in section 2.3.1.4 that the dependencies that trigger the deletion reflex involve movement (the ex-situ dependencies are island-sensitive and exhibit weak cross-over effects). Furthermore, we have seen that an overtly (fully or partially) moved wh-phrase must be immediately followed by the marker *yang*, cf. (48-a), repeated in (83-a). In case of partial wh-movement, *yang* shows up in the sentence where the wh-phrase lands but not in a sentence above that position, cf. (52-c), repeated in (83-b). Importantly, *yang* cannot co-occur with a wh-word that stays in-situ, cf. (52-a), repeated in (83-c).[35]

(83) *Distribution of the marker yang in Indonesian*:

    a.   Siapa$_k$ **yang** Sally cintai ___$_k$
         who   FOC Sally love
         'Who does Sally love?'   *full wh-movement*
    b.   [$_{CP}$ Tom (meng)-harap [$_{CP}$ apa$_k$ **yang** Mary (*mem-)beli   ___$_k$]]
         Tom TRANS-expect   what FOC Mary TRANS-bought
         'What does Tom expect Mary bought?'   *partial wh-movement*
    c.   Sally (men)-cintai siapa
         Sally TRANS-love who
         'Who does Sally love?'   *wh-in situ*

Hence, the distribution of *yang* is exactly the same as those of other reflexes of movement we have seen so far. Further evidence that it is a reflex of movement comes from variation: Cole and Hermon (1998) show that for some speakers of the closely related Singaporean Malay *yang* can occur in every clause crossed by overt movement, cf. (84):

---

[34]One might wonder why I include Ewe in the sample of languages because the element that shows the reflex of movement is in a specifier position and not on a head; reflexes in specifiers have been excluded from the discussion because they may not involve the agreement relation that I am interested in, see section 2.1.2. However, the reflex in Ewe is different from the ones in the aforementioned section: In the latter, the reflex occurs in the Ā-position targeted by the movement step that triggers the reflex (such as wh-copying). In Ewe, it is also movement to SpecCP that triggers a reflex, but this reflex occurs not in SpecCP but in SpecTP. Collins (1993: 41f.) argues that the subject pronoun is located in SpecTP because the subject precedes auxiliaries, negation, and aspectual heads (cf. (i-b)):

(i) *Location of the subject (Collins 1993: 40, 161)*:

    … bé mì mé gǎ   dzó o
        that 1PL NEG SUBJ leave PART
        'that I don't leave'

So there is movement to SpecTP, but this is not sufficient to trigger the change in the pronoun form, only Ā-movement to SpecCP does (see e.g. (79-a) with the 3rd person subject pronoun in its default form in SpecTP). Thus, the reflex in Ewe is triggered over a distance, between the element in SpecCP and the element in SpecTP. Therefore, I conclude that an agreement relation is involved in Ewe.

[35]Saddy does not provide an ungrammatical example with *yang* following a wh-word in-situ, but he explicitly states that it is impossible (p. 185).

(84)   *Distribution of yang in Indonesian, (Cole and Hermon 1998: 224)*:

[$_{CP}$ Siapa$_k$ **yang** Bill harap [$_{CP}$ **yang** ___$_k$ akan mem-beli   baju   untuknya]]
      who   FOC Bill hope   FOC   will TRANS-buy clothes for.him
‘Who does Bill hope will buy clothes for him?’

I conclude from these facts that *yang* is a reflex of overt movement. In Malay, it can have a pattern I distribution. However, in the variety of Indonesian described by Saddy (1991), *yang* only occurs in the final clause, cf. (53-b), repeated in (85-a), and the ungrammatical example (85-b).

(85)   *Cross-clausal wh-movement in Indonesian (Saddy 1991: 214, n.7)*:

   a. [$_{CP}$ Siapa$_k$ **yang** Bill (*men-)kira [$_{CP}$ Tom (*men-)harap [$_{CP}$ Fred
         who   FOC Bill TRANS-think   Tom TRANS-expect   Fred
       (*men-)cintai ___$_k$]]]
       TRANS-love
       ‘’Who did Bill think Tom expects Fred loves?
   b. *[$_{CP}$ Siapa$_k$ **yang** Bill tahu   [$_{CP}$ bahwa **yang** Tom citai ___$_k$]]
          who   FOC Bill knows   that   FOC Tom loves
       ‘Who does Bill know that Tom loves?’

Thus, the occurrence of *yang* follows pattern II in Indonesian. It will be important for the upcoming discussion that Indonesian has both a pattern I and a pattern II reflex. The pattern I distribution is Malay will be ignored in what follows.

There is some debate about the nature of the element *yang*: Saddy (1991) argues that *yang* is a focus marker, see also Sabel (2000); Fanselow (2006). Others gloss it as the complementizer ‘that’, cf. Macdonald and Dardjowidjojo (1967) on Indonesian and Cole and Hermon (1998; 2000) on Malay. However, Saddy (1991: 188f.) shows that *yang* is not just the interrogative counterpart of the declarative complementizer *bahwa* because both elements can co-occur in embedded questions, cf. (86):

(86)   *yang co-occurs with the declarative complementizer (Saddy 1991: 188)*:

   a. Bill tahu bahwa Tom men-cintai siapa
       Bill know that   Tom TRANS-love who
       ‘Bill knows who Tom loves.’   *wh-in-situ*
   b. Bill tahu bahwa siapa$_k$ **yang** Tom cintai ___$_k$
       Bill know that   who   FOC Tom love
       ‘Bill knows who Tom loves.’   *wh-ex-situ*

In line with Saddy (1991), I assume that *yang* is a focus marker that realizes the head of the XP in whose specifier the wh-phrase lands, i.e. it is a complementizer-like element. Thus, in contrast to *men*-deletion which takes place in the verbal domain, *yang* is a reflex in the C-domain. Given that *yang* and *bahwa* can co-occur, I assume that Indonesian has CP-recursion.

2.3.3 Pattern III reflexes

In this section we will look at three pattern III reflexes that lead to addition of exponents: complementizer agreement in Wolof *an*-chains, preverbal focus marking in Kitharaka, and *ke*-stranding in Dinka. Furthermore, German exhibits a syntactic reflex of movement that follows pattern III.

2.3.3.1 C-agreement in Wolof

We have seen in section 2.3.1.3 that Wolof has complementizer agreement as a reflex of movement: In *u*-chains with a silent wh-operator, C heads agree in class with the wh-phrase in its specifier. This reflex has a pattern I or pattern II distribution. Apart from *u*-chains, Wolof has another way to form questions: *an*-chains. In these, the wh-expression is overt; it surfaces as *an* and is preceded by a class prefix. (87-a) is the *an*-chain counterpart of (41-b), repeated in (87-b).

(87)   *an-construction (Torrence 2012: 1148, 1152)*:

     a.   Y-an$_k$ l-a   jigéén j-i   togg ___$_k$
          CL-an EXPL-a woman CL-DEF.PROX cook
          'What(pl) did the woman cook?'   *an-chain*

     b.   Ø$_k$ **y**-u jigéén j-i   togg ___$_k$
          Q   CL-u woman CL-DEF.PROX cook
          'What(pl) did the woman cook?'   *u-chain*

     c.   K-an$_k$ l-a   Ayda døør ___$_k$
          CL-an EXPL-a Ayda hit
          'Who did Ayda hit?'   *an-chain*

The complementizer is realized as *a-* instead of *u-*. As shown in (87), the complementizer is preceded by the class prefix *l-*. This marker is invariant since it does not alternate with the noun class of the wh-phrase (compare (87-a) and (87-c)). The class prefix of the wh-phrase cannot cannot surface on the complementizer in *an*-chains. The *l*-prefix is the default class marker that also occurs on complementizers in *u*-chains with a pattern II reflex. Thus, there is no class agreement between the complementizer and the wh-phrase in its specifier in *an*-chains. The Ā-dependency in *an*-chains involves movement because it is island-sensitive and exhibits reconstruction effects, just as the dependency in the *u*-chain. (88-a) shows that movement of the *an*-phrase out of an adjunct island is barred; in (88-b) we see that the *an*-phrase can reconstruct.

(88)   *Wh-movement from islands in Wolof (Torrence 2012: 1157, 1176)*:

     a.   *[$_{CP}$ L-an$_k$ l-a   xale b-i   dem [$_{CP}$ laata Bintë togg-al   Móódu
            CL-an EXPL-a child CL-DEF.PROX leave   before Binte cook-BEN Moodu
            ___$_k$]]
          'What did the child leave before Binte cooked for Moodu?'   *adjunct island*

     b.   L-an$_k$ l-a   [xale b-u nekk] bëgg
          CL-an EXPL-a   child CL-u exist   love
          'Who does every child love?'   *reconstruction: Wh > ∀, ∀ > Wh*

Let me now come to long Ā-movement. As pointed out, there is no class agreement on the complementizer that is adjacent to the wh-expression in *an*-chains; rather, we find the default marker *l-*. Under cross-clausal extraction, however, class agreement with the wh-phrase is possible on all other complementizers crossed by movement, which then surface in the *u*-form. Put differently, there is class agreement in non-final clauses of *an*-chains; agreement is only barred from the final clause, Thus, we get a pattern III reflex, cf. (89-a):

(89)   *Long-extraction in Wolof an-constructions (Torrence 2012: 1173)*:

  a. [_CP_ K-an l-a-ñu   wax [_CP_ **k**-u jigéén j-i   foog [_CP_ **k**-u ma
         CL-an EXPL-a-3PL say   CL-u woman CL-DEF.PROX think   CL-u 1SG
      dóór]]]
      hit
      'Who did they say that the woman thinks that I hit?'   *pattern III*
  b. [_CP_ K-an **l-a**-ñu   wax [_CP_ **l-a**   jigéén j-i   foog [_CP_
         CL-an EXPL-a-3PL say   EXPL-a woman CL-DEF.PROX think
      l-a-a   dóór]]]
      EXPL-a-1SG hit
      'Who did they say that the woman thinks that I hit?'   *pattern IV*

Class agreement in non-final clauses is optional, however. That is, (89-a) alternates with
(89-b) where all complementizers bear the default class prefix (pattern IV). As with *u*-
chains, there is actually even more variation, which I will discuss in section 2.5.3.2.

2.3.3.2 Preverbal focus marking in Kitharaka

In Kitharaka (Niger-Congo, Kenia), we find a pattern III reflex of Ā-movement that in-
volves addition of an exponent, viz. of a focus marker. The basic word order is SVO, cf.
(90-a).[36] An Ā-moved XP appears in clause-initial position, preceding the subject. This is
illustrated in (90-b) for wh-movement. The wh-word can also stay in situ, cf. (90-c). Only
XPs that have undergone overt Ā-movement bear a focus prefix, cf. (90-b) vs. (90-c).[37]

(90)   *Clause-bound wh-movement in Kitharaka (Abels and Muriungi 2008: 692,*
         Muriungi 2005: 45):

  a. Maria a-gur-ir-e   i-buku
      Maria SM-buy-PERF-FV 5-book
      'Maria bought a book.'
  b. I-mbi_k   Maria a-k-ir-e   ___k
      FOC-what Maria SM-build-PERF-FV
      'What did Maria build?'   *wh-ex-situ*
  c. Maria a-k-ir-e   mbi
      Maria SM-build-PERF-FV what
      'What did Maria build?'   *wh-in-situ*

When looking at long Ā-movement, we can see that the prefix *n-* attaches to the verb;
it has an allomorph *i-* before consonant-initial stems. Muriungi (2003; 2005) and
Abels and Muriungi (2008) identify this prefix as a focus marker. In what follows, we are
only interested in this preverbal focus marker and ignore the focus marker on the wh-
phrase. (91) shows that the preverbal focus marker appears in a clause that is crossed by
movement but where the wh-phrase does not land. (91-a) illustrates this for movement
of an argument XP over one clause boundary, whereas(91-b) and (91-c) for argument and

---

[36]Note that the verb always ends in a final vowel that has different surface forms depending on various
factors, e.g. if the sentence is negated or not. The vowel does not play a role for the discussion of reflexes
of movement. Furthermore, Kitharaka is a tone language. Following Muriungi (2005), the tones are not
indicated in the examples because they do not seem to interfere with focus marking.

[37]All the facts that will be presented in this subsection for wh-movement also hold for focus movement in
Kitharaka, including the distribution of the morphological reflex with wh-in-situ because Kitharaka allows
for focus in-situ as well.

adjunct movement over two clause boundaries. The preverbal focus marker cannot appear on the verb of the matrix clause where the wh-phrase surfaces, but must occur in all lower clauses. Thus, focus marking in Kitharaka exhibits pattern III: It only appears in non-final clauses. Muriungi (2005: 67) explicitly states that "... all the verbs in the embedded clauses which the wh-phrase passes through bear the particle *i* or *n*, except the verb of the clause in which the wh-phrase occurs."

(91) *Cross-clausal wh-movement (Muriungi 2005: 47-48, 67-68):*

    a. [$_{CP}$ I-mbi$_k$ g-ug-ir-e   [$_{CP}$ ati John **n**-a-ring-ir-e   ___$_k$]]
        FOC-what 2SG-say-PERF-FV   that John FOC-SM-beat-PERF-FV
        'What did you say that John beat?'   *DO question*

    b. [$_{CP}$ N-uu$_k$ u-ku-thugania [$_{CP}$ ati John **n**-a-ug-ir-e   [$_{CP}$ Lucy
        FOC-who 2SG-PRES-think   that John FOC-SM-say-PERF-FV   Lucy
        n-a-ring-ir-e   ___$_k$]]]
        FOC-SM-beat-PERF-FV
        'Who do you think that John said Lucy beat?'   *DO question*

    c. [$_{CP}$ I-ku$_k$ u-ku-thugania [$_{CP}$ ati chibu **n**-a-ug-ir-e   [$_{CP}$ borisi
        FOC-where 2SG-PRES-think   that chief FOC-SM-say-PERF-FV   police
        n-i-on-ir-e   Lawrence ___$_k$]]]
        FOC-SM-see-PERF-FV Lawrence
        'Where do you think that the chief said the police saw Lawrence?'
                                            ADJ question

This generalization also explains the absence of the preverbal focus marker with local Ā-movement as in (90-b), as well as its absence on the verb of the embedded question in (92). In both cases, the wh-phrase surfaces in the clause in which it originates, hence, there is no reflex of movement.

(92) *Embedded question (Muriungi 2005: 49):*

        [$_{CP}$ n-ti-iji   [$_{CP}$ i-mbi$_k$ Munene a-gur-ir-e   ___$_k$]]
        1SG-NEG-know   FOC-what Munene SM-buy-PERF-FV
        'I don't know what Munene bought.'

There is massive evidence that this preverbal focus marker is indeed a reflex of overt movement. First of all, the movement types that trigger this reflex are island-sensitive. This is shown in (93) for overt wh-movement from a complex-NP-island and a wh-island.

(93) *Island-sensitivity of wh-movement (Muriungi 2005: 63, Abels 2012: 71):*[38]

    a. ??[$_{CP}$ N-ata$_k$ u-ku-ama   [$_{CP}$ kethira **n**-a-kar-ir-e   ___$_k$]]
        FOC-what 2SG-PRES-wonder   whether FOC-SM-behave-PERF-FV
        'How do you wonder whether s/he behaved?'   *wh-island*

    b. *[$_{CP}$ N-uu$_k$ u-ku-thugan-i-a   [$_{CP}$ **i**-n-ding-ir-e   [$_{NP}$ mw-ari$_i$
        FOC-who 2SG-PRES-think-CAUS-FV   FOC-1SG-hit-PERF-FV   CL-girl
        [$_{CP}$ OP$_i$ u-ra   **n**-a-ug-ir-e   [$_{CP}$ ati Peter **n**-a-gur-ir-e
        CL-that FOC-SM-say-PERF-FV   that Peter FOC-SM-marry-PERF-FV
        ___$_k$]]]]]
        'Who do you think that I hit the girl$_j$ who$_j$ said that Peter will marry?'
                                            CNP-island

---

[38]Abels (2012) atributes the CNP-island example to Peter Muriungi (p.c.).

Furthermore, the preverbal focus marker cannot occur on the verb of the embedded clause if a wh-word from this clause stays in-situ but takes matrix scope, see (94-a) (the in-situ counterpart of (91-a)), and (94-b). The marker does not appear because no verb has been crossed by *overt* movement.

(94) *Wh-in situ in Kitharaka (Muriungi 2005: 69-70)*:

    a. [CP G-ug-ir-e   [CP ati John a-ring-ir-e   mbi]]
         2SG-say-PERF-FV   that John FOC-SM-beat-PERF-FV what
         'What did you say that John beat?'

    b. [CP U-ri-thugania [CP ati John a-ug-ir-e   [CP Lucy a-ug-ir-e
         2SG-PRES-think   that John SM-say-PERF-FV Lucy SM-say-PERF-FV
         [CP Pat a-ring-ir-e   uu   ]]]]
         Pat SM-beat-PERF-FV who
         'Who do you think that John said Lucy said Pat beat?'

Kitharaka allows for partial wh-movement. In this context, the preverbal focus marker cannot appear above the overt landing site of the wh-phrase, cf. (95) (note that the wh-phrase follows the embedding complementizer). In (95-a), the reflex can also not show up on the verb of the embedded clause because although this verb lies on the path of overt movement, it is in the clause where the wh-phrase surfaces; recall that the reflex is generally absent from final clauses. In (95-b), where partial wh-movement crosses two clause boundaries, we see the reflex on the most deeply embedded verb because it is a non-final clause.

(95) *Partial wh-movement in Kitharaka (Muriungi 2005: 48, 71)*:

    a. [CP G-ug-ir-e   [CP ati n-uu$_k$ John a-ring-ir-e ___$_k$]]
         2SG-say-PERF-FV   that FOC-who John SM-beat   PERF-FV
         'What did you say that John beat?'

    b. [CP U-ri-thugania [CP ati n-uu$_k$ John a-ug-ir-e   [CP Lucy
         2SG-PRES-think   that FOC-who John SM-say-PERF-FV   Lucy
         n-a-ring-ir-e   ___$_k$]]]
         FOC-SM-beat-PERF-FV
         'Who do you think that John said Lucy beat?'

Thus, the preverbal focus marker in Kitharaka is a reflex of overt movement that instantiates pattern III.

2.3.3.3 *ke*-stranding in Dinka

Dinka (Nilo-Saharan, South Sudan) displays a pattern III reflex of movement in the VP that involves agreement between (an extended projection of) the verb and the moving phrase. In addition, it has a pattern I reflex of the type that occurs in the specifier position a phrase moves through. Although the latter type of reflex is not the focus of this thesis, I will introduce the corresponding Dinka reflex here for two reasons: On the one hand, it will be very helpful to understand the extraction examples, apart from the central pattern III reflex; on the other hand, it shows that Dinka is one of the languages that exhibits multiple reflexes of a single movement operation, which will be important for the discussion of previous approaches to patterns II and III in section 2.4.

Let me begin with some basic facts about Dinka clause structure and argument encoding (see Andersen 1991; 1995; 2002 on details of Dinka morphology, phonology and basic syntax).[39] The basic clause structure is given in (96):

(96)   *Dinka clause structure (Andersen 2002: 6):*
         Topic   Verb$_{fin}$ Subject   Object   Verb$_{non-fin}$ Object   Adverbial

Dinka has the V2-property: The finite verb occupies the second position in the clause. There must be exactly one constituent before the finite verb, in a position that Andersen (2002: 6) calls the topic position. Any kind of argument or adjunct can occupy this position. Ignoring movement of a constituent to the topic position for the moment, the order of constituents after the finite verb is quite rigid: The subject precedes the object which in turn precedes the verbal complex consisting of non-finite verb(s). In a ditransitive context with two objects, only one of the objects precedes the non-finite verb, the other one follows the verb; which of the two objects precedes and which follows the verb is not fixed, the order is free. Adverbials occur in clause-final position. The following examples exemplify the template with a ditransitive verb. In (97-a) the subject is in the topic position, the indirect object precedes and the direct object follows the non-finite verb; in (97-b) the locative adverb is in topic position, the direct object precedes and the indirect object follows the non-finite verb.[40]

(97)   *Clauses with ditransitive verbs in Dinka (van Urk and Richards 2013: 5):*
         a.   Cằn a-bí   Bòl ɣɔ̀ɔc   aléth   rɔ̀ɔk
              Can 3SG-FUT Bol buy.DTR clothes town.LOC
              'Can will buy Bol clothes at the town.'
         b.   Rɔ́k a-bíi   Cán   aléth   ɣɔ̀c   Bòl
              town 3SG-FUT.NS Can.GEN clothes buy.DTR Bol
              'Can will buy Bol clothes at the town.'

Dinka uses both head- and dependent marking to encode arguments. Nouns inflect for number and case (Andersen 2002: 2). Dinka basically exhibits an ergative alignment system: The sole argument of an intransitive verb and the internal argument of a transitive verb are unmarked, whereas the external argument of a transitive verb has a more marked form. The cases are distinguished by tonal differences (Andersen 1991: 272-273). The unmarked case is called absolutive, the marked case is traditionally not called ergative, but rather genitive (Andersen 1991: 273) or oblique case (Andersen 2002: 7) because it is the same case marker that is used with possessors. I adopt the label genitive used in van Urk and Richards (2013) where the following examples are taken from. In addition, there is a special marking for adverbials, the locative case. Dinka has a split in its alignment system: XPs in the topic position are always in the unmarked case (absolutive), even if they would bear a marked case (genitive, locative) in a postverbal position (Andersen 1991: 273, 281). The result is a neutral alignment system in case the transitive subject is in topic position.[41] This effect is exemplified in (97): In (97-b), the subject of

---

[39]The data that exemplify the pattern III reflex are taken from van Urk and Richards (2013), who examine the Bor dialect group; the data in Andersen (1991; 1995; 2002) are based on the Agar dialect group. But for the basic facts of clause structure the two dialects work the same.

[40]The following diacritics are used in the Dinka glosses (cf. van Urk and Richards 2013): The grave represents a low tone and the accute a high tone; the circumflex indicates a falling tone. In addition, Dinka has different voice qualities: underlined vowels are in creaky voice; dieresis below a vowel indicates breathy voice.

[41]See e.g. Siewierska (2004: 50ff.) and Bickel and Nichols (2009) for an overview on alignment types.

the (di)transitive verb is postverbal and thus bears the marked genitive case, whereas it is in the unmarked case in the topic position in (97-a). The locative case is realized on the sentence-final adverb in (97-a) but the adverb is unmarked when in the topic position as in (97-b). This split is not central for the following discussion, but it may help to understand the examples. With respect to head-marking, the finite verb bears a prefix that cross-references the number as well as the grammatical function (subject vs. non-subject) of the XP in topic position.[42]. This marker is present in both (97-a) and (97-b) on the future auxiliary; note the different form depending on whether a subject or a non-subject is the topic. The marker on the finite verb is not central to the following discussion. What is important, however, is head-marking on the non-finite verb, to which I will turn below.

But let us first look at the pattern I reflex of movement in Dinka. As indicated in the template in (96), the topic position and the position immediately before the non-finite verb must be filled by exactly one XP in Dinka. van Urk and Richards (2013) identify Andersen's topic position as the specifier of the head C (see pp. 7-11 for arguments) and the immediately preverbal ($V_{non-fin}$) position as SpecvP (for arguments see pp. 12-13). Since SpecvP and SpecCP must be filled, they assume that v and C have the EPP-property. This means that they inherently bear a structure-building feature, which they call the EPP-feature, that triggers movement of an XP to their specifier. (98) shows that leaving one of these positions empty results in ungrammaticality (van Urk and Richards 2013: 7, 12-13); empty positions are boxed. Any XP (argument or adjunct) can fill SpecCP; SpecvP is more restricted in that it can only host argument XPs but not adjuncts (van Urk and Richards 2013: 13).

(98)   *SpecCP and SpecvP must be occupied (van Urk and Richards 2013: 7, 13)*:

    a. *☐ cí   Bôl   kurá cák
              PERF Bol.GEN bowl make
        'Bol made a bowl.'

    b. *ɣɛn cí   ☐ tíŋ mîir
        I   PERF   see giraffe
        'I saw a giraffe.'

    c. *ɣɛn cí   ☐ yiɛ́n Ayén kitàp
        I   PERF   give Ayen book
        'I gave a book to Ayen.'

With this background, let us look at movement dependencies, exemplified with wh-movement (although the same holds for other Ā-movement types such as topicalization and relativization). There are three strategies to form wh-questions: Either the wh-word stays in-situ, or it moves to its scope position SpecCP, or it undergoes partial movement to a SpecCP position which is below its scope position. The ex-situ strategy is obligatory if the wh-word combines with the focus particle *ye-* (van Urk and Richards 2013: 10-11). In-situ questions freely alternate with ex-situ questions (van Urk and Richards 2013: 10, fn.6). The ex-situ strategy clearly involves movement because the wh-dependency is island-sensitive:

(99)   *Island-sensitivity in Dinka (Coppe van Urk, p.c.)*:

    a. *[CP Yeŋú$_k$ cíi   Ádìt jàạl [CP wuín cíi   Máyèn ___$_k$ kuèm]]
           what PERF.NS adit leave   when PERF.NS Mayen   break
        'What did Adit leave because Mayen broke?'   *adjunct-island*

---

[42]Number marking only occurs in declarative clauses and only if the XP is 3rd person: *a-* is used for 3sg nouns, *aa-* for 3pl nouns, cf. Andersen (1995: 7).

b. *[_CP Yeŋú_k cíi   Máyòk ràan   [_CP lói ___k] tíŋ]
       what PERF.NS Mayok person   do   see
       'What did Majok see someone that makes?'   *CNP-island*

Interestingly, SpecCP and SpecvP need not be filled when an XP is moved through them. If, for example, the indirect object is questioned and undergoes movement to SpecCP, SpecvP must be empty, cf. (100). Although it is in principle possible that the direct object occupies SpecvP, it cannot do so if the indirect object is moved to SpecCP (and vice-versa if the direct object moves).

(100)   *Obligatory empty Specv under object extraction (van Urk and Richards 2013: 15)*:

a. Yeŋà_k cíi   mòc ☐ yiɛ̬n ___k kitàp
     who   PERF.NS man   give   book
     'Who did the man give the book to?'

b. *Yeŋà_k cíi   mòc kitàp_n yiɛ̬n ___k ___n
     who   PERF.NS man book   give
     'Who did the man give the book to?'

Hence, the generalization that SpecCP and SpecvP must be filled is not surface-true under movement, an instance of opacity. van Urk and Richards (2013) propose that this pattern is derived if the EPP is checked derivationally and if movement applies successive-cyclically through every SpecvP and SpecCP: In (100-a), the EPP is checked by the indirect object when it moves to SpecvP. This is masked by the subsequent movement of the indirect object to a structurally higher position (SpecCP). The same effect obtains for SpecCP if an XP undergoes cross-clausal Ā-movement: In this case, the embedded SpecCP position must be empty on the surface (van Urk and Richards 2013: 16-17). Again, this is because the XP checks the EPP of the embedded C when moving to its specifier; further movement to the matrix SpecCP obscures that the embedded SpecCP had been filled. Thus, obligatory emptiness of positions that otherwise must be occupied is a reflex of movement in Dinka. To be more precise, it is a pattern I reflex because it occurs along the path of movement (every SpecCP and SpecvP on the path of long Ā-movement must be empty). It could be classified as a deletion reflex like e.g. *men*-deletion in Malay, though in specifiers and not on a head. Of course, it does not literally involve deletion, but on the surface there is less material in the specifier then would be without movement.

With this much background on the syntax of extraction in Dinka, let me finally come to the crucial data. Apart from the obligatory emptiness, there is another (morphological) piece of evidence for successive-cyclic movement through SpecvP, introduced in van Urk and Richards (2013): If a plural XP moves through SpecvP, the morpheme *ke* occurs at the left edge of the vP. van Urk and Richards (2013) refer to this process as *ke*-stranding. This stranding is obligatory, *ke* cannot be omitted, cf. (101-a) vs. (101-b). Under long-distance extraction, *ke* must occur in every SpecvP crossed by movement, cf. (101-c); omitting one or both *ke*s leads to ungrammaticality (van Urk and Richards 2013: 19-20). Extraction of a singular XP does not trigger the occurrence of an (overt) morpheme, cf. (101-d).

(101)   *ke-stranding in Dinka (van Urk and Richards 2013: 18-19)*:

a. Yèyîŋa_k cíi   Bôl   **ké** ☐ tíŋ ___k
     who.PL PERF.NS Bol.GEN PL   see
     'Who all did Bol see?'   *ke-stranding*

b. *Yèyîŋa_k cíi   Bôl   ☐tíŋ ___k
   who.PL PERF.NS Bol.GEN   see
   'Who all did Bol see?'   *no ke-stranding*

c. [_CP Yèyîŋà_k yé   **ké**☐tàak, [_CP ☐cíi   Bôl   **ké**☐tíŋ
      who.PL IMPERF.2SG PL   think   PERF.NS Bol.GEN PL   see

   ___k]]
   'Who all do you think Bol saw?'   *ke-stranding in every vP*

d. Yeŋà_k cíi   Bôl   ☐tíŋ ___k
   who PERF.NS Bol.GEN   see
   'Who did Bol see?'   *sg-XP extracted*

van Urk and Richards (2013) do not provide an analysis of the distribution of *ke*. They only state that this reflex of movement provides additional evidence for successive-cyclic movement through SpecvP. However, their paper contains a number of examples from movement of plural XPs in declarative sentences (provided to illustrate properties other than *ke*-stranding) from which we can gain a better understanding about the distribution of *ke*. (102-a) shows that what we have already seen with wh-movement in (101-b) also applies to EPP-driven movement: Movement of a plural XP to SpecCP leads to *ke*-stranding in SpecvP. Interestingly, *ke*-stranding does not apply when the plural XP stops in SpecvP. This situation arises if another constituent satisfies the EPP-property of C, cf. (102-b) (recall that adjuncts cannot satisfy the EPP of v so that an object can occur in SpecvP although the adjunct is extracted). Furthermore, *ke* does not occur if the plural XP stays in its VP-internal base-position, cf. (102-c). The plural XP is in italics in (102).

(102)   *The distribution of ke in declarative clauses (van Urk and Richards 2013: 5)*:

   a. *Aléth_k* aa-bíi   Cán   **ké**☐ɤɔ̀c   Bòl ___k rɔ̀ɔk
      clothes 3PL-FUT.NS Can.GEN PL   buy.DTR Bol   town.LOC
      'Can will buy Bol clothes at the town.'   *pl-XP: SpecCP*

   b. Rɔ̂_k a-bíi   Cán   *aléth_n* ɤɔ̀c   Bòl ___n ___k
      town 3SG.FUT.NS Can.GEN clothes buy.DTR Bol
      'Can will buy Bol clothes at the town.'   *pl-XP: SpecvP*

   c. Càn_k a-bí   ___k Bòl_n ɤɔ̀c   ___n *aléth* rɔ̀ɔk
      Can 3SG-FUT   Bol buy.DTR   clothes town.LOC
      'Can will buy Bol clothes at the town.'   *pl-XP: in the VP*

As Coppe van Urk informs me (p.c.), the same abstract pattern holds e.g. for wh-movement. Wh-words can be left in situ. If a plural wh-object stays in the VP, there can be no *ke*-stranding. And if a plural wh-word ends up in SpecvP (partial wh-movement), *ke* cannot co-occur with it.

These facts suggest that *ke*-stranding has a pattern III distribution: *ke* only occurs in the vP if there is *overt* movement to SpecvP (i.e. not with plural XP in-situ) and if the XP that moves to SpecvP uses this position as an intermediate landing site; *ke* cannot co-occur with the XP, viz. if SpecvP is the final landing site of XP. But this means that *ke*-stranding is actually a misnomer: *ke* is not stranded in the sense that quantifiers are stranded in the analysis by Sportiche (1988), cf. section 2.1.2. Stranding in Sportiche's terms means that a phrase XP and the stranded element form a constituent in the base position of the XP. XP can move alone and leave behind the other element in the base position, i.e. XP strands the other element; alternatively, both elements can move together and then co-occur in the landing site of XP. *ke* does not behave like this: It can never

co-occur with the plural XP, neither in the base position of the associated XP nor in its landing site. Hence, *ke*-stranding does not involve stranding. *ke* is a morphological reflex of movement that involves addition of material and follows pattern III. I will nevertheless continue to use the term *ke*-stranding introduced by van Urk and Richards (2013), just to have a label for the phenomenon. I suggest that v agrees in number with the XP moved through SpecvP; if it acquires the value *plural*, this is realized by the exponent *ke*. That *ke* encodes the value plural is evident from the fact that it also occurs as a suffix marking 3rd person plural subjects in verbal inflection (see Andersen 1995: 20) as well as in the possessive paradigm where it indicates the plurality of the possessum (see Andersen 2002: 17).

2.3.3.4 Obligatory extraposition in German

German (Indo-European) has a syntactic reflex of movement that follows pattern III: obligatory extraposition. Let me first illustrate optional extraposition. The basic word order in German is SOV. It surfaces in embedded declarative clauses; the object DP cannot follow the verb, cf. (103). The verb-final order is obscured in main clauses due to the V2-property which requires that the finite verb moves to C and that exactly one constituent is moved to SpecCP; this can lead to the occurrence of the object DP in postverbal position.

(103)   *SOV word order in embedded clauses:*[43]

    a. daß er   Claudia   liebt
        that he.NOM Claudia.ACC loves
        'that he loves Claudia'

    b. *daß er   liebt Claudia
        that he.NOM loves Claudia.ACC
        'that he loves Claudia'

Although objects precede the main verb in embedded clauses, sentential objects are preferably extraposed to the postverbal position. Hence, CP extraposition is optional in declarative clauses in German, cf. (104).

(104)   *Optional extraposition of declarative sentential objects (Müller 1999: 376)*:

    a. daß er [$_{CP}$ daß Fritz   Claudia   liebt] gesagt / nicht gewußt hat
        that he   that Fritz.NOM Claudia.ACC loves said   / not   known has
        'that he said / did not know that Fritz loves Claudia'   *non-extraposed*

    b. daß er ___$_k$ gesagt / nicht gewußt hat [$_{CP}$ daß Fritz   Claudia
        that he   said   / not   known has   that Fritz.NOM Claudia.ACC
        liebt]$_k$
        loves
        'that he said / did not know that Fritz loves Claudia'   *extraposed*

The same holds for embedded interrogative clauses (although extraposition is even more preferred than in declarative clauses, therefore the question mark in (105-a)).

(105) *Optional extraposition of interrogative sentential objects (Müller 1999: 376)*:

    a. ?daß er [$_{CP}$ wen$_k$ Fritz   ___$_k$ liebt] nicht gesagt / gewußt hat
        that he   whom.ACC Fritz.NOM   loves not   said / known has
        'that he did not say / know who Fritz loves'   *non-extraposed*

---

[43]I made up the examples in (103) on the basis of Müller's (1998) examples.

 b. daß er ___$_k$ nicht gesagt / gewußt hat [$_{CP}$ wen$_n$ Fritz ___$_n$ liebt]$_k$
   that he not said / known has whom.ACC Fritz.NOM loves
   'that he did not say/know who Fritz loves'   *extraposed*

Stechow and Sternefeld (1988) observe that extraposition of a sentential object becomes *obligatory* if wh-extraction has taken place from that CP, i.e. if the CP contains an intermediate landing site of the moved wh-phrase but not the wh-phrase itself, as it was the case in (105-a) (see also Müller 1999: 377):

(106) *Obligatory CP-extraposition under wh-extraction (Müller 1999: 377)*:
 a. *(Ich weiß nicht) [$_{CP}$ wen$_n$ er [$_{CP}$ ___$_n$ daß Fritz ___$_n$ liebt]
   I know not whom.ACC he that Fritz.NOM loves
   gesagt hat]
   said has
   'I do not know who he said that Fritz loves.'   *non-extraposed*
 b. (Ich weiß nicht) [$_{CP}$ wen$_n$ er ___$_k$ gesagt hat [$_{CP}$ ___$_n$ daß Fritz
   I know not whom.ACC he said has that Fritz.NOM
   ___$_n$ liebt]$_{k]}$
   loves
   'I do not know who he said that Fritz loves.'   *extraposed*

Müller (1998) identifies obligatory extraposition as a reflex of wh-movement. Crucially, it is a pattern III reflex because if the CP contains the final landing site of the wh-phrase, extraposition is not compulsory. That movement is involved is clear from the fact that wh-movement in German is island-sensitive.[44] This is shown in (107) for movement from a complex NP island and an adjunct island.

(107) *Wh-extraction out of strong islands (Müller 1995: 59, 84)*:
 a. *Wen$_k$ hast du [$_{NP}$ einen Artikel [$_{CP}$ der ___$_k$ beeindruckt hat]]
   who.ACC have you an article which.NOM impressed has
   geschrieben
   written
   'Lit: Whom did you write an article that impressed?' *CNP island*
 b. *Was$_k$ ist sie losgefahren [$_{CP}$ nachdem sie ___$_k$ getrunken hat]
   what.ACC is she driven after she drunk has
   'Lit: What did she drive after she drank?'   *adjunct island*

In addition, wh-movement induces strong cross-over effects (whereas weak cross-over effects are less strong or even absent):

(108) *Strong cross-over under wh-movement (Müller 1995: 163)*:
 *[$_{CP}$ Ich weiß nicht [$_{CP}$ [wen$_i$]$_k$ sie$_i$ ___$_k$ getroffen hat]]
  I know not whom.ACC she.NOM met has
  'I don't know whom she met.'

As for wh-in situ, we cannot test whether the reflex occurs or not because German does not allow for wh-in-situ with a matrix or embedded question interpretation. Another construction that can be used to show that the reflexes are tied to *overt* movement is partial

---

[44]Island-sensitivity holds at least for strong islands; weak island violations such as extraction from wh-islands is quite acceptable.

movement. Indeed, German allows for what looks like partial wh-movement, though we will see that it is different from the instances of partial movement in the other languages discussed in this section. In the German construction, the wh-phrase can surface in an embedded SpecCP position that is not its scope position; additionally, the scope position is filled by the scope marker *was*, cf. (109-a) with full wh-movement vs. (109-b) with partial wh-movement:

(109) *Full vs. partial wh-movement in German (Müller 1999: 381)*:

    a. (Ich weiß nicht) [$_{CP}$ wen$_n$ er ___$_k$ gesagt hat [$_{CP}$ ___$_n$ dass Fritz
       I   know not   whom.ACC he   said has   that Fritz.NOM
      ___$_n$ liebt]$_k$]
          loves
      'I do not know who he said that Fritz loves.'   *full movement*
    b. (Ich weiß nicht) [$_{CP}$ was$_n$ er ___$_k$ gesagt hat [$_{CP}$ wen$_n$ (dass)
       I   know not   SCO he   said has   whom.ACC that
      Fritz   ___$_n$ liebt]$_k$]
      Fritz.NOM   loves
      'I do not know who he said that Fritz loves.'   *partial movement*

There is a debate about the nature of partial wh-movement: It could be full movement of the wh-phrase to its scope position but without pied-piping the whole wh-phrase; part of it it stranded in a lower SpecCP, where the wh-word surfaces. Another view is that wh-movement is indeed only partial; the wh-phrase moves to a SpecCP position below its scope position and the scope marker is base-generated in the scope position (see Müller 1999: 382 for references). Müller (1999: 383) argues for the first view based on the interaction of partial wh-movement with extraposition. (109-b) vs. (110-a) shows that extraposition of the CP containing the partially moved wh-phrase is obligatory. It thus behaves like a CP with an intermediate wh-trace and not like one in which the wh-phrase has reached its scope position; in the latter case, extraposition would be optional.

(110) *Interaction of partial wh-movement and extraposition (Müller 1999: 383f.)*:

    a. *(Ich weiß nicht) [$_{CP}$ was$_n$ er [$_{CP}$ wen$_n$ (dass) Fritz   ___$_n$ liebt]
       I   know not   SCO he   whom.ACC that Fritz.NOM   loves
      gesagt hat]
      said   has
      'I do not know who he said that Fritz loves.'
    b. ?dass ich [$_{CP}$ was$_n$ er ___$_k$ gesagt hat [$_{CP}$ wen$_n$ Fritz   ___$_n$ liebt]$_k$]
      that I   SCO he   said has   whom.ACC Fritz.NOM   loves
      gar   nicht wissen möchte
      PART not   know want
      'that I really do not want to know who he said that Fritz loves.'

In addition, the CP containing the scope marker *was* does not have to undergo extraction, it is only optional, cf. (110-b). The CP with the scope marker thus behaves like a CP with a wh-phrase in its final landing site. This provides evidence for the hypothesis that there is indeed movement of the wh-phrase to its scope position in the "partial" wh-movement construction' – just without pied-piping of all features of the wh-phrase. Hence, we are actually dealing with full wh-movement in the construction with the scope marker. Therefore, it is not surprising that German "partial" wh-movement behaves differently from

partial movement in the languages we have seen so far. In the other languages, the clause in which the partially moved phrase surfaces behaves like a final clause with respect to the reflex of movement. In German, it behaves like a non-final clause.[45]

To summarize, German obligatory extraposition is a reflex of movement that exhibits pattern III (with partial movement behaving like full movement): Extraposition is compulsory only for CPs whose SpecCP position is an intermediate landing site for wh-movement, but not if the wh-phrase surfaces in SpecCP. Finally, note that we find the same abstract pattern in Basque: A CP through whose SpecCP a wh-phrase has moved must be obligatorily extraposed; usually, extraposition is optional and it remains optional for CPs in whose specifier the wh-phrase surfaces (cf. Ormazabal et al. 1994).

2.3.4   Data summary

The following table summarizes the data presented in this section. It contains the phenomena, the reflex type, the reflex pattern(s) they follow, as well as their properties: Is the Ā-dependency that triggers the reflex sensitive to islands? Does it exhibit strong or weak cross-over effects? A checkmark indicates that relevant data have been presented and that the movement types do exhibit island effects and cross-over effects, respectively. As for partial movement, in-situ constructions (no overt movement at all) and movement that solely applies within the final clause, the question is whether the reflex occurs in clauses that are not crossed by overt movement. A checkmark indicates that the reflex behaves as expected if the reflexes are tied to overt movement only, i.e. that no reflex occurs in clauses that are not on the path of movement. An asterisk indicates that the data are not available to me; a hyphen indicates that the language does not have the relevant construction. As for Wolof complementizer agreement, additional evidence for movement comes form reconstruction (and not from cross-over effects).

We can see that in all languages, the reflex occurs when the Ā-dependency involves movement (island-sensitivity, cross-over effects or reconstruction). In addition, it is clearly tied to overt movement in the languages where we find the relevant constructions to test this (partial movement and in-situ-constructions).

---

[45]Note that in the other languages, there was no scope marker in partial movement constructions. It would be interesting to see whether the presence / absence of a scope marker correlates with full vs. partial movement of the wh-phrase that does not surface in its scope position. That is, is what looks like partial wh-movement actually full movement in all languages with a scope marker? And is partial movement "real" partial movement in all languages without a scope marker? The distribution of reflexes of movement provides a test case in languages that morphologically distinguish final and intermediate movement steps.

language	phenomenon	reflex type	pattern	properties			
				island-sensitive	cross-over	in-situ	partial mvt.
Chamorro	wh-agreement	morphological replacement	I	✓	✓ (strong)	–	–
	complementizer agreement	morphological addition	II	✓	✓ (strong)	–	–
Irish	complementizer selection: *aL*-chains	morphological replacement	I	✓	✓ (weak)	–	–
Wolof	complementizer agreement: *u*-chains	morphological addition	I / II	✓	✓ (reconstr.)	–	–
	complementizer agreement: *an*-chains	morphological addition	III / IV	✓	✓ (reconstr.)	–	–
Indonesian	*men*-deletion	morphological deletion	I	✓	✓ (weak)	✓	✓
	focus-marking	morphological addition	II	✓	✓ (weak)	✓	✓
Kikuyu	downstep deletion	phonological deletion	I	✓	*	✓	✓
Spanish	inversion	syntactic	I	✓	✓ (weak)	✓	–
Duala	*no*-marking	morphological addition	II	✓	*	✓	–
Ewe	subject pronoun choice	morphological replacement	I / II	✓	*	–	–
Kitharaka	preverbal focus marking	morphological addition	III	✓	*	✓	✓
Dinka	*ke*-stranding	morphological addition	III	✓	*	✓	✓
German	extraposition	syntactic	III	✓	✓ (strong)	–	–

Table 2.1: Data summary: reflexes of successive-cyclic movement

2.4   Previous analyses of pattern II

Languages with pattern I, in which the reflex occurs in every clause along the path of overt movement, have been taken as strong evidence that Ā-movement applies in a successive-cyclic fashion with intermediate stops in every clause crossed by movement. So far, I have implicitly been assuming in the presentation of the data that patterns II and III involve successive-cyclic movement as well. However, especially with respect to pattern II, one might question this assumption since we do not see a reflex of putative intermediate movement steps. The aim of this section is to justify this assumption. I will provide evidence that – at least in some languages – even movements that produce the reflex patterns II or III apply successive-cyclically. In sections 2.4.1 and 2.4.2 I provide empirical evidence against an alternative analysis of pattern II that denies the existence of intermediate movement steps or agreement in intermediate positions. Hence, I argue that we are dealing with opacity: The generalization that movement of XP to SpecHP results in a morphological reflex on the head H is not surface-true; in pattern II and pattern III languages, the reflex occurs only on some of the heads H to whose specifier XP is moved, see section 2.4.3. Sections 2.4.4 and 2.4.5 discuss previous analyses of pattern II (all couched within the Government and Binding framework) that make crucial use of levels of representation and enriched representations to resolve opacity. We will see that they have a number of empirical and conceptual problems.

2.4.1   No reflex – no movement?

Following standard assumptions, I assume that morphological reflexes of movement are the result of an agreement relation between a head H and an XP moved to SpecHP. The existence of pattern I has led researchers to the conclusion that cross-clausal Ā-movement makes intermediate stop-overs. In Wolof u-chains, for example, a C head agrees in class with an XP moved to SpecCP. Since we find complementizer agreement in every clause along the path of movement, we can conclude that the moving XP moves successive-cyclically (at least) through every SpecCP; depending on the location of the reflex, we can make the same conclusion for other heads. So, the generally accepted idea is that whenever there is a morphological reflex on a head H, there must be an intermediate landing site in SpecHP. Simply put, the implication in (111) holds:

(111)   If there is a reflex of movement in (the vicinity of) a position P, there must have been movement through P.

This might seem trivial, but we will see that things are not as simple in the reverse direction. Let us look at pattern II: In languages of this type, the morphological reflex only occurs in the clause where the Ā-moved XP surfaces. The abstract pattern is repeated in (112) (for wh-movement and a reflex R on the C head):

(112)   *Pattern II*:
            $[_{CP1}$ wh-XP$_k$ $[_{C'}$ **C$_1$-R** ... $[_{CP2}$ ___$_k$ $[_{C'}$ C$_2$ ... $[_{CP3}$ ___$_k$ $[_{C'}$ C$_3$ ... ___$_k$ $]]]]]]$

How can we analyze the distribution of the reflex? An obvious assumption would be to interpret the correlation between movement and the occurrence of a morphological reflex in (111) as a bi-conditional. The reverse implication would be the following:

(113)   If there is movement to a position P, there must be a reflex of movement in (the vicinity of) P.

Applied to pattern II, (113) would amount to say that there is no reflex of movement in non-final clauses simply because the XP that surfaces in the topmost clause of the dependency does not make a stop-over in non-final clauses. There are two ways to implement this (henceforth, I will label the different hypotheses about the emergence of pattern II as (HI), (HII), etc.):

(HI) XP is base-generated in its surface position.

(HII) XP moves to its surface position, but it does so in one fell swoop, without any stop-overs.

I will address each of them in turn.

2.4.1.1 Base-generation

Indeed, there is independent evidence for analysis (H1) from Irish. We have seen in section 2.3.1.2 that Irish has a reflex of Ā-movement that instantiates pattern I: The form of the complementizer changes from *go* to *aL* in all clauses along the movement path, see (114-a). That this involves movement is clear from the fact that the Ā-dependencies that trigger this alternation are island-sensitive and show weak cross-over effects. The relevant examples are repeated in (114-b) and (114-c).

(114) *Irish aL-chains*:

    a. [an t-úrscéal]$_i$ [$_{CP}$ [OP$_i$]$_k$ **a** mheas mé [$_{CP}$ **a** dúirt sé [$_{CP}$ **a**
        the novel   *aL* thought I   *aL* said he   *aL*
        thuig   sé ___$_k$]]]
        understood he
        'the novel that I thought he said he understood'   *long relativization*

    b. *an fear$_i$ [$_{CP}$ [OP$_i$]$_k$ **a** phóg mé an [$_{NP}$ bhean$_j$ [$_{CP}$ [OP$_j$] **a** phós
        the man   *aL* kissed I   the   woman   *aL* married
        ___$_k$]]]
        'the man who I kissed the woman who married'   *CNP island*

    c. *[fear$_i$] [$_{CP}$ [OP$_i$]$_k$ **a** d'fhág a$_i$ bhean ___$_k$]
        man   *aL* left   his wife
        'a man that his wife left'   *WCO*

The complementizer in Irish can take another special from beside *aL*: *aN* (inducing nasalization on the following segment). *aN* also occurs in Ā-dependencies such as relativization, cf. (115-a). But in contrast to sentences with the *aL* complementizer, the *aN* complementizer (a) requires a resumptive pronoun instead of a gap in the theta-position of the operator, and (b) can only occur in the clause where the operator lands; C heads in non-final clauses are realized as *go,* just like in clauses without any Ā-dependency. Hence, the distribution of *aN* follows pattern II. Examples of long relativization with *aN* is given in (115-b) and (115-c).

(115) *aN-chains in Irish (McCloskey 2001:67, McCloskey 2002:190, McCloskey 1990:75)*:

    a. [an ghirseach]$_i$ [$_{CP}$ OP$_i$ **ar** ghoid na síogaí í$_i$]
        the girl   *aN* stole the fairies her
        'the girl that the fairies stole (her) away'

    b.   an [t-ór seo]$_i$ [$_{CP}$ OP$_i$ **ar** chreid   corr-dhuine [$_{CP}$ **go** raibh sé$_i$ ann]]
       the gold DEM   *aN* thought some.people   *go* was it there
       'this gold that some people thought there was'

    c.   an rud$_i$ [$_{CP}$ OP$_i$ **ar** dúirt sé [$_{CP}$ **go** gcoinneodh sé ceilte é$_i$]]
       the thing   *aN* said he   *go* keep.COND he hidden it
       'the thing that he said he would keep (it) hidden'

Interestingly, *aN*-chains differ in another important way from *aL*-chains: In contrast to the latter, the former are neither island-sensitive (cf. (116-a)) nor do they exhibit weak cross-over effects (cf. (116-b)). The two complementizers in (116-a) are both written as *a* but they can be distinguished by the mutations they induce on the following segment (lenition vs. nasalization).

(116)   *Properties of aN-chains (McCloskey 1979: 33-34, McCloskey 1990: 110)*:

    a.   sin teanga$_i$ [$_{CP}$ OP$_i$ **a** mbeadh meas   agam ar [$_{NP}$ duine$_j$ ar bith [$_{CP}$
       that language   *aN* would.be respect at.me on   person any
       OP$_j$ **a** tá ábalta ___$_j$ í$_i$ a labhairt]]]
          aL is able   it to speak
       'That is a language that I would respect anyone who could speak it.'
                                            CNP-island

    b.   fear$_i$ [$_{CP}$ OP$_i$ **ar** fhág a$_i$ bhean é$_i$]
       man   *aN* left his wife   him
       'a man that his wife left'   *WCO configuration*

McCloskey (2002 and previous work) concludes from these facts that there is no Ā-movement in *aN*-chains; rather, the operator is base-generated in its surface position. The different reflexes *aN* and *aL* in Ā-dependencies result from the operation that puts the XP into its surface position: When *aL* shows up, the XP is moved to SpecCP (successive-cyclically); when *aN* occurs, the XP is base-generated in the specifier of the C head realized as *aN*. Hence, Irish provides evidence for the assumption that pattern I is triggered by successive–cyclic movement, whereas pattern II arises because there are no intermediate movement steps in intermediate SpecCP positions: There is no movement in the first place (note that McCloskey does not make any claims about the nature of pattern II beyond Irish).[46]

---

[46]The same argument has been made on the basis of the alternation between pattern I and the absence of any reflex of movement: We have seen that Chamorro and Ewe have pattern I under Ā-movement: In Chamorro, wh-agreement occurs on all verbs along the path of movement, and in Ewe the 3rd person subject pronoun takes a special form under extraction. This holds e.g. for wh-movement. There is evidence from island-sensitivity that the reflexes indeed involve movement (see the data in section 2.3.1.1 and 2.3.2.3). In both languages, however, the morphological reflex must not occur in any clause under topicalization: In (i-a) from Ewe, the special subject pronoun *wò* cannot occur in the embedded clause (note that topics are followed by the topic marker ɖe and a resumptive pronoun is present at the thematic base position of the topic phrase). In (ii-a) from Chamorro, there is the regular subject-verb-agreement morphology on the verb under topicalization; recall that in other Ā-dependencies wh-agreement occurs and suppresses the regular agreement morphology.

(i)   *Topicalization in Ewe (Collins 1993: 182)*:

    a.   Kofi$_i$ ɖe me gblɔ be   [é/*wo] fo-e$_i$
       Kofi TOP I   said that he   hit-him
       'As for Kofi, I said that he hit him.'

    b.   Kofi$_i$ ɖe me nya ame   [$_{CP}$ xe   fo-e$_i$]
       Kofi TOP I know person   which hit-him
       'As for Kofi, I know the person who hit him.'   *CNP-island*

Can this analysis be generalized to all instances of pattern II? The answer is no. In some languages like Irish, pattern II indeed arises because there is no movement involved, but this does not hold for all pattern II languages. As I have shown in section 2.3.2, the movement types that trigger the respective reflex in Chamorro, Duala, Ewe, Malay and Wolof are island-sensitive. Let me repeat an example of extraction from a complex NP island in Ewe:

(117)   *Island-sensitivity in Duala*:

   *moto$_i$ [$_{CP}$ [nyena$_i$]$_k$ na neimbi **no** [$_{NP}$ mbo$_j$ [$_{CP}$ [nyenat$_j$]$_n$ ___$_n$ e kuko
   man   who   I recognize NO dog   which   SM bit

   ___$_k$]]]
   'the man who I recognized the dog which bit (him)'   *CNP island*

To conclude, a base-generation analysis is not applicable to these languages: Pattern II does not simply follow from the absence of movement.

2.4.1.2   Movement in one fell swoop

Let me now turn to hypothesis (HII): We have seen that the Ā-dependency that triggers a pattern II reflex involves movement in some languages, but maybe movement applies in one fell swoop from the base position of the XP to its surface position. Since there are no intermediate landing sites, there can be no reflex of movement in non-final clauses (no agreement can be established between a head and an XP in an intermediate landing site). Indeed, such an analysis has been proposed for some of the languages with pattern II in section 2.3.2: Epée (1976*a;b*) argues on the basis of the distribution of *no*-marking in Duala that movement is not successive-cyclic; Martinović (2013:10, fn.8) mentions long movement without intermediate stop-overs as a possible analysis for Chamorro complementizer agreement.[47]   Similarly, Takahashi and Gracanin Yuksek (2008) speculate that

---

(ii)   *Topicalization in Chamorro (Chung 1998:266-267)*:

   a.   Pues [i gubietnu giya Guam] ha-konni' pro$_i$ todu tatti i   taotao
        so   the governor LOC Guam AGR-take   all   back the person
        'So the governor on Guam took all the peoply back.'

   b.   I che'lu-n Maria$_i$ taya'   [$_{NP}$ [$_{CP}$ Op$_j$ ma'a'ñao-ña   pro$_i$ ___$_j$] médiku$_j$]
        the sibling-L Maria AGR.not.exist   WH.OBL.afraid-AGR   doctor
        'Maria's sister, there are no doctors she's afraid of.'   *CNP-island*

Collins (1993) and Chung (1998) provide evidence that topicalization does not involve movement in Ewe and Chamorro, respectively, because it is not island-sensitive: In (i-b) and (ii-b) we see that the alleged gap is located inside a complex NP island. Nevertheless, the sentences are grammatical. This suggests that the operator is base-generated in its surface position. Crucially, other types of Ā-dependencies that trigger a pattern I reflex in Ewe and Chamorro are island-sensitive. As in Irish, the absence of a morphological reflex is thus an indicator of the absence of movement in topicalization in Chamorro and Ewe. The difference is that Irish shows an alternation between patterns I and II, whereas Chamorro and Ewe exhibit an alternation between pattern I and the absence of any reflex (pattern IV). Thus, in Irish the complementizer can Agree with a base-generated operator; in Chamorro and Ewe, it cannot do so. See section 4.1 for an account of such asymmetries in terms of the different timing of external Merge relative to Agree.

[47]Postal (1972) also proposes that the non-occurrence of a reflex of Ā-movement in intermediate positions provides evidence that Ā-movement is not successive-cyclic: English allows for preposition stranding under movement of an NP that is the complement of a preposition. However, stranding is impossible in positions that have been argued to be intermediate landing sites for cross-clausal Ā-movement, such as SpecCP of an embedded clause. I do not discuss the English data here in detail for two reasons: First, the reflex is not a pattern II reflex; it only occurs in the clause from which the extracted element originates and not in the final clause. See section 4.3 for discussion of such reflexes. Second, the reflex occurs in an XP-position that is part

the absence of a morphological reflex of Ā-movement on intermediate C heads in Haitian is due to the fact that cross-clausal Ā-movement does not make a stop-over in the embedded SpecCP position. But this is not because Ā-movement applies in one-fell swoop per se, rather, they assume that embedded clauses lack the CP-projection and hence the potential intermediate landing site – the effect is the same as if movement was not successive-cyclic. Takahashi and Gracanin Yuksek (2008: 240) do not provide independent evidence from Haitian Creole for their claim. It is clear, however, that this analysis cannot be transferred to every language with pattern II: We have seen data from several languages with pattern II where the complementizer surfaces in non-final clauses, for instance in Duala (see section 2.3.2.2). For these languages, it is clear that a CP – and thus a potential intermediate landing site – is projected; nevertheless, there is no reflex in non-final clauses. Therefore, I put aside the proposal by Takahashi and Gracanin Yuksek (2008) in what follows.

The question is whether the proposal by Epée (1976*a;b*) can be generalized: Does Ā-movement in all pattern II languages apply in one fell swoop, accounting for the absence of reflexes in intermediate positions? Again, I claim that the answer is no. I will present two arguments for this claim.

The first argument is based on the observation that some languages have more than one reflex of movement simultaneously, and the distinct reflexes exhibit different distribution patterns. Of particular interest are those languages that have both a pattern I and a pattern II reflex triggered by the same Ā-dependency. In a nutshell, the argument goes as follows: Remember that I adopt the standard view that if there is a morphological reflex of movement on a head H, there has been movement through SpecHP. In order to account for pattern I, we have to assume that movement makes a stop-over in every clause along the path of movement, i.e. that it is successive-cyclic. Furthermore, under the current hypothesis, the absence of a reflex implies the absence of movement, i.e. pattern II follows if movement applies in one fell swoop. Now, if pattern I and pattern II co-occur in a single language and if they are both triggered by the same Ā-dependency, we encounter a paradox: One and the same movement step would have to apply successive-cyclically (for pattern I) and in one fell swoop (for pattern II) at the same time, which is impossible.

We have encountered two languages that exhibit this configuration: Chamorro and Indonesian. Chamorro shows a pattern I reflex on the verb (wh-agreement) and a pattern II reflex on the complementizer, cf. sections 2.3.1.1 and 2.3.2.1. They are both triggered e.g. by wh-movement. This is abstractly depicted in (118):

(118)   *C-agreement and wh-agreement under long Ā-movement in Chamorro*:
          [$_{CP1}$ wh-XP$_k$ [$_{C'}$ **C-R** ... **v-R** ... [$_{CP2}$ C ... **v-R** [$_{CP3}$ C ... **v-R** ... ___$_k$]]]]

The same holds for Indonesian, more precisely for those varieties in which the focus marker *yang* has a pattern II distribution, cf. section 2.3.2.4. In addition, there is a pattern I reflex on the verb, viz. *men*-deletion, cf. section 2.3.1.4. Crucially, both *men*-deletion and focus marking are triggered by the same types of movement, e.g. by wh-movement. Thus, given the assumption that a morphological reflex of movement indicates a landing site of the moving XP, the pattern I reflex of Ā-movement in Chamorro and Malay / Indonesian clearly shows that movement applies in a successive-cyclic fashion. The pattern II

---

of the movement chain and not on a head that projects this position (see section 2.1.2 for more reflexes of this type). Hence, it is not clear whether agreement between a head and the moving XP is involved. As noted above, this thesis is about the interaction of agreement and movement (Agree and Merge), therefore, I ignore reflexes in intermediate specifier positions.

distribution of the other reflex can thus not be attributed to the absence of intermediate stop-overs, falsifying hypothesis (HII).

Apart from this empirical argument, there is a conceptual problem when we adopt hypothesis (HII): Since there are numerous languages with pattern I reflexes that show that Ā-movement applies in small steps, we would have to assume that languages can differ in the locality of Ā-movement: In languages with pattern II, Ā-movement is unbounded, it can span a potentially unlimited number of clauses; in languages with pattern I, however, Ā-movement is subject to strict locality conditions and thus applies in a sequence of local steps. Given that research in the last decades has shown that many typologically diverse languages are subject to the same locality restrictions, viz. islands, I take this to be undesirable.

Before closing this section, let me discuss a potential objection. There may be a way to uphold hypothesis (HII) for Chamorro and Malay / Indonesian, once we recognize that the pattern I and the pattern II reflex occur on different heads: The former occurs in the vP-domain on a head I will identify as v, and the latter on the complementizer C. We could say that cross-clausal Ā-movement goes through every SpecvP along the path of movement, but not through intermediate SpecCP positions. This is abstractly depicted in (119): XP originates in the VP of the embedded clause CP_2 and undergoes movement to SpecCP of the matrix clause CP_1. It makes a stop-over in every SpecvP on its way, but not in the embedded SpecCP position.

(119)   *Successive-cyclic Ā-movement through SpecvP*:
         $[_{CP_2} XP_k C \dots [_{vP} \rule{1em}{0.4pt}_k v \dots [_{CP_2} C \dots [_{vP} \rule{1em}{0.4pt}_k v \dots [_{VP} \rule{1em}{0.4pt}_k]]]]]$

Movement through every SpecvP would result in a pattern I reflex on v, and skipping of intermediate SpecCP leads to a pattern II reflex on the complementizer.[48]

Indeed, such a derivation without stop-overs in intermediate SpecCP positions has been proposed by Rackowski and Richards (2005) and in a number of papers by den Dikken (den Dikken 2009*a*;*b*; 2012) for cross-clausal Ā-movement. The conclusions against the need for intermediate movement through SpecCP drawn in Rackowski and Richards (2005) from Tagalog have been revisited in van Urk and Richards (2013). They provide an empirical argument from Dinka that successive-cyclic movement *does* go through intermediate SpecCPs after all, and they propose an analysis with intermediate movement steps through all SpecvP and SpecCP positions on the path of movement that captures both the Dinka and the Tagalog facts. We have seen the empirical evidence for stop-overs in these positions (obligatory emptying) in section 2.3.3. As for den Dikken's approach, the challenge is to account for the occurrence of reflexes of movement in intermediate SpecCP positions although Ā-movement does not stop in these positions. den Dikken proposes the following solution: In case of (apparent) wh-movement from an embedded to the matrix clause as in the Belfast English example in (120-a), the wh-phrase indeed only moves to the embedded SpecCP. This movement is terminal, the wh-phrase cannot move on to the matrix clause, cf. (120-b). The wh-phrase we see in the matrix clause is only a scope marker (Sco), as in the scope marking strategy with partial movement in German, see (109-b). In contrast to the German scope marker, however, the wh-element in the matrix SpecCP looks identical to the real wh-phrase in the embedded SpecCP due to a process of concord between the scope marker and the real wh-phrase. den Dikken calls this strategy 'full-concordial scope marking'.

---

[48]An empirical argument against this alternative analysis would be languages where the pattern I and the pattern II reflex both occur on the same head. I am not aware of such a language, however.

(120)   *Full-concordial scope marking (den Dikken 2009b: 90,93)*:

    a.   Who did Mary claim did they meet?

    b. $[_{CP}$ (Sco+FF$_{DP}$)=who did Mary $[_{VP}$ claim $[_{CP}$ DP$_{wh}$ did they meet DP$_{wh}$]]]

                 ❶ concord —————— ❶ Move ——

Only the highest wh-phrase in the matrix SpecCP position is pronounced. The other wh-phrase in the "intermediate" position (which is actually the final position in den Dikken's analysis) remains silent because it is identical to the structurally highest wh-phrase (the scope marker that undergoes full concord). Since there is an occurrence of the wh-phrase in the intermediate SpecCP position, it can trigger a morphosyntactic reflexes such as inversion in Belfast English; but this wh-phrase does not *move through* this position, it terminates there. Abels (2012: 51ff.) convincingly argues against den Dikken's proposal. In a nutshell, Abels shows that full-concordial scope marking patterns with movement rather than with scope marking or concord with respect to linearization, the ability to trigger pied-piping and locality. He concludes that "[a]ll of this raises the suspicion that 'full-concordial scope marking' is no more than a new label for movement through Spec,CP." (ibid.) In addition, Abels (2012: 51, fn.25) points out data that call into question the initial assumption that movement to SpecCP is always terminal (i.e. criterial in Rizzi's (1997) sense). Finally, let me emphasize again that the massive empirical evidence from morphosyntactic as well as semantic reflexes of movement in the C and V/v domain has led researchers in the field to agree on the conclusion that if there is a position through which successive-cyclic movement applies, then it is SpecCP and SpecvP. In general, the debate is about whether other heads are the target of such movements, too, but SpecCP and SpecvP are by and large undisputed. Hence, I take a derivation with successive-cyclic movement through SpecvP but not through SpecCP not to be tenable for Chamorro and Malay / Indonesian.

      To summarize the discussion so far, we have seen one piece of empirical evidence against the claim that pattern II arises because movement applies in one fell-swoop. The evidence was that the co-occurrence of patterns I and II triggered by a single Ā-movement dependency.

      In the remainder of this section, I will provide a second argument against hypothesis (HII). This argument is based on pattern III reflexes (although the hypothesis has never been pursued for pattern III reflexes). Can we transfer the hypothesis that the non-occurrence of a reflex implies the absence of movement to pattern III? That is, could the absence of a reflex in the clause where the Ā-moved XP surfaces be attributed to the absence of movement to the respective position in that clause? The answer is very clearly no. First, there is evidence from island-sensitivity and / or weak cross-over that the movement types resulting in a reflex of pattern III actually involve movement and not base-generation. We have seen the relevant data from Wolof, Kitharaka, Dinka, and German in section 2.3.3. Second, remember that I am assuming that the occurrence of a reflex of movement on a head H implies movement through / to SpecHP (cf. (111)). Since we do see reflexes in all non-final clauses in languages with pattern III, we can conclude that movement applies successive-cyclically. Assume we have a pattern III reflex on C heads as in (121). From what has been said so far we know that steps (a) and (b) involve movement because there is a reflex and the dependency is island-sensitive.

(121)   *Pattern III: Long Ā-movement*:

$[_{CP_1}$ wh-XP$_k$ $[_{C'}$ C$_1$ ... $[_{CP_2}$ ___ $_k$ $[_{C'}$ **C$_2$-R** ... $[_{CP_3}$ ___ $_k$ $[_{C'}$ **C$_3$-R** ... ___$_k$ $]]]]]]$

              ? (c)   Move (b)   Move (a)

The question is what kind of dependency step (c) is. Plausibly, it is movement as well. What other dependency should it involve? A concordial scope marking analysis with base-generation of an operator in SpecCP$_1$ + concord with the XP in SpecCP$_2$ may come to mind.[49] But first of all, if full concordial scope marking is nothing else than movement anyway, this mechanism does not tell us anything about how the final step differs from the intermediate steps. In addition, there is empirical evidence that the dependency in (c) involves movement: In Kitharaka, we have seen examples in which there is an island between CP$_1$ and CP$_2$ and the operator surfaces in SpecCP$_1$. These sentences are ungrammatical, cf. (93) in section 2.3.3.2; an example is repeated in (122). If (c) involves movement out of CP$_2$, the ungrammaticality follows automatically (movement is island-sensitive). If it involved concord, den Dikken would have to assume that concord is island-sensitive (which would, however, make scope marking indistinguishable from movement, see Abels' 2012 argumentation).

(122)   *Wh-island in Kitharaka*:

??$[_{CP_1}$ N-ata$_k$ u-ku-ama $[_{CP_2}$ kethira **n**-a-kar-ir-e   ___$_k$ $]]$
        FOC-what 2SG-PRES-wonder   whether FOC-SM-behave-PERF-FV
'How do you wonder whether s/he behaved?'
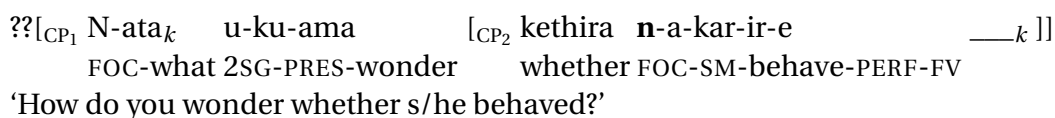
Hence, there is clearly movement to SpecCP in the final clause, but still there is no reflex of this movement. The existence of pattern III thus also shows that the absence of a reflex of movement does not imply that there has been no movement.

To summarize, there is empirical evidence against the hypotheses (HI) and H(II). Even in pattern II languages movement applies successive-cyclically, but it does not trigger a reflex in all clauses crossed by movement.

2.4.2 No reflex – no Agree?

In the previous subsection, I have argued that the absence of reflexes in intermediate positions in pattern II languages is not (generally) due to the absence of movement to these positions. An alternative analysis that comes to mind is the following:

(HIII) There are intermediate movement steps in pattern II languages, but there is no agreement relation between the head H and the XP in SpecHP if SpecHP is an intermediate landing site.

I will model agreement through the operation Agree. As introduced in section 1.2.3.1, Agree is triggered by probe features on H; these features are unvalued and seek for a value on an XP. Thus, one could assume that in pattern II languages, the heads that trigger intermediate movement steps simply lack a probe feature to Agree with XP and hence, there

---

[49]An analysis along these lines might indeed be a possibility for Irish which allows for a mixture of *aL*-chains (involving movement) and *aN*-chains (involving base-generation), so-called hybrid chains: The operator is moved to an embedded SpecCP (which triggers *aL*-marking) and is bound by an operator base-merged in a structurally higher clause (with the *aN*-complementizer). However this derivation cannot be transferred to those pattern II languages which display island-sensitivity for the step between CP$_1$ and CP$_2$, as I argue in this subsection.

is no reflex of movement; only heads that project a final landing site bear a probe feature. As for pattern III languages, we would have to assume that the reverse holds: Only heads that project intermediate landing sites bear the relevant probe features. This analysis can easily be applied to languages with a non-zero / zero marker alternation, i.e. to languages in which an overt reflex alternates with the absence (a zero marker) of any marker in final vs. non-final clauses: In Duala, for instance, the marker *no* occurs on the finite verb in the final clauses of an Ā-dependency; in non-final clauses, *no* is not replaced by a different marker, rather, there is no marker at all on the finite verb, viz. *no* is not replaced, it is absent, cf. (123):

(123)   *Long-distance movement in Duala*:

    [$_{CP}$ [njika buna]$_k$ o   ta   **no** o   kwalane mba [$_{CP}$ na o   mende timba ___$_k$]]
        which day   you PST NO you tell   me   that you FUT   return
    'When did you tell me that you would return?'

It is difficult to find arguments against this analysis for languages like Duala.[50] However, this solution is less plausible for pattern II languages like Wolof, for instance, where we find a non-zero / non-zero alternation of exponents: In final clauses of *u*-chains, which can optionally exhibit pattern II, there is an overt marker on the C head that indicates agreement with the moved XP, whereas in non-final clauses the marker is not absent, but rather replaced by a default marker; cf. the example repeated in (124):

(124)   *Long-distance wh-movement in Wolof u-chains*:

    [$_{CP}$ Ø$_k$ k-u   Kumba wax [$_{CP}$ ne l-a   Isaa defe [$_{CP}$ ne **l**-a   Maryam dóór
        Q   CL-u Kumba say   FRC EXPL-a Isaa think   FRC EXPL-a Maryam hit
    ___$_k$]]]
    'Who did Kumba say that Isaa thought that Maryam hit?'

The same default marker is found in Wolof an-chains that follow pattern III: The C heads in non-final clauses exhibits full class agreement with the wh-phrase, but the C head in the final clause bears the default class marker (cf. section 2.3.3.1). The presence of the default marker indicates that there is a probe on the complementizer that attempts to get valued, but if valuation through copying of a value from the XP in SpecHP fails (for whatever reason), a default value is inserted to satisfy the needs of the probe (otherwise, the derivation would crash, see section 2.5.2.1 for technical details). If there was no probe on heads that project intermediate landing sites, why should a default marker be inserted in the first place? What should it stand in for? I conclude that the absence of a reflex of movement on a head H cannot generally be the result of the lack of an agreement trigger on H, contra hypothesis (HIII).

2.4.3   It's opaque!

Let me summarize the discussion so far. I have investigated whether the absence of reflexes in intermediate positions in pattern II languages is due to the absence of intermediate landing sites. There are no intermediate landing sites if (i) there is no movement in the first place – rather, the seemingly Ā-moved XP is base-generated in its surface

---

[50]Note that the analysis that assumes that there is no Agree between H and an XP moved to SpecHP is highly problematic for approaches in which Agree between H and XP is a prerequisite for movement of XP to SpecHP.

position (HI); or (ii) because movement applies in one fell swoop from the base to the surface position of XP (HII). I have provided empirical evidence from island-sensitivity and the co-occurrence of patterns of reflexes that at least in some of the languages with pattern II the Ā-dependencies that trigger the reflex do involve movement and do apply successive-cyclically, falsifying hypotheses (HI) and (HII).[51] Furthermore, the existence of default markers argues against (HIII), the absence of probe features on heads that project intermediate landing sites in pattern II languages, and on heads that project final landing sites in pattern III languages. Thus, hypothesis (HI) – (HIII) are empirically falsified.

The important conclusion we can draw from this is that **patterns II and III are opaque**: Ā-movement is successive-cyclic in languages with pattern I, II, and III (contra (HI) and (HII)). And usually, Ā-movement of an XP leaves a reflex on H if XP targets SpecHP because H always bears a probe feature (contra (HIII)). However, **the generalization** that movement to SpecHP triggers agreement with H (which results in a morphological or syntactic reflex) **is not surface-true** in languages with pattern II or pattern III: We see a reflex only on some of the heads H that are present in the structure, but not on others. In Chamorro, for example, movement to SpecCP **feeds agreement** with C. However, under cross-clausal Ā-movement, this agreement only surfaces in the clause where the operator lands, but not on the C head of non-final clauses, although there is movement through intermediate SpecCPs. Hence, **agreement** between the operator in SpecCP and C **underapplies** in intermediate SpecCPs in Chamorro; movement to intermediate SpecCPs **counter-feeds agreement** with C. In languages with pattern III, agreement between a head and the XP moved to its specifier underapplies in the final landing site of XP.

The insight that languages with patterns II and III, respectively, exhibit an opaque interaction of movement and agreement will be the basis for the further discussion in this thesis. I summarize the main points in (125):

(125)   *Transparent and opaque patterns of reflexes of movement*:

> (a) Ā-movement is subject to the same locality restrictions in languages with pattern I, II, and III. Ā-movement applies successive-cyclically.
>
> (b) Observation: Movement of XP to SpecHP can in prinicple feed agreement between H and XP in SpecHP. Result: a reflex on H.
>
> (c) Pattern I (transparent): Every movement step to SpecHP *feeds* agreement between H and XP.
>
> (d) Pattern II (opaque): The final movement step of XP to SpecHP *feeds* agreement between H and XP. Intermediate movement steps *counter-feed* agreement. Agreement *underapplies* in intermediate positions.
>
> (e) Pattern III (opaque): Intermediate movement steps of XP to SpecHP *feed* agreement between H and XP. The final movement step *counter-feeds* agreement. Agreement *underapplies* in the final position.

---

[51]Boeckx (2003; 2008*b*) argues that resumption as in Irish *aN*-chains is the result of successive-cyclic movement (even in the absence of island effects) and does not involve base-generation of the phrase in the position immediately preceding *aN*, as is standardly assumed. If one adopts this view, the absence of *aN* in non-final clauses can no longer be attributed to base-generation, i.e. the lack of intermediate landing sites. Thus, the problem of how to account for pattern II also arises for languages in which the respective Ā-dependency shows no island-sensitivity.

The overall conclusion is that the presence of a morphological reflex of movement on a head H can be taken as strong evidence for the presence of an XP in SpecHP (externally or internally merged) which agrees with H and thereby triggers the reflex. However, the reverse does not hold: **The absence of a reflex of movement on H does not automatically indicate the absence of movement to SpecHP.** Interestingly, the same conclusion has been drawn for semantic reflexes of successive-cyclic movement (cf. 2.1.2.2): The absence of evidence for reconstruction to an intermediate position SpecHP does not necessarily imply that there has been no successive-cyclic movement through this position, see e.g. Anand and Nevins (2005); Bobaljik and Wurmbrand (2005); Boeckx (2008*b*), contra Abels (2003; 2012).[52] It is thus not surprising to find the same effect for morphosyntactic reflexes of movement.

2.4.4   Opacity and levels of representation

In the previous subsections, I have argued against two hypotheses according to which pattern II is not opaque because there is either no movement to positions where we do not find a reflex, or no attempt to Agree. In this and the following subsection, I will discuss three analyses of pattern II from the literature that acknowledge the opacity in the sense that it is assumed that movement is successive-cyclic, even in pattern II languages (though without any arguments).[53] Hence, the question arises why the reflex does not occur in intermediate landing sites.

   In this section, I will be concerned with analysis (AI): reference to levels of representation. Recall that in Chamorro, complementizer agreement is a reflex of movement that follows pattern II. Chung (1998: 230), working within the Government and Binding framework, proposes the following constraint on complementizer agreement in Chamorro ([N], [O], and [local] are Chung's abbreviations for the features involved in the agreement relation):

(126)   *Operator-C Agreement (holds at S-structure)*:
         C^0 and an Associate that is both its specifier and an operator must have compatible values for [N], [O], and [local].

Reformulated as a rule / operation, (126) says that C agrees with an operator in its specifier in certain features. Crucially, this constraint applies at S-structure, i.e. the output structure of syntax proper, after all movement operations have taken place (cf. figure 1.6

---

[52]Boeckx (2008*b*: 58) is very explicit about the issue: "Whereas the copy theory of movement readily accounts for reconstruction by involving the interpretation of unpronounced copies, we cannot conclude from this that if no reconstruction effect is found, no copy is available at the relevant site. All we can conclude from the absence of reconstruction is either that there is no copy present, or that a copy was created, but for some (perhaps interpretive) reason cannot be interpreted in the relevant position."

[53]In fact, Biloa (1993) is the only one who argues for successive-cyclic Ā-movement in a pattern II language, namely in Duala (contra the claim in Epée 1976*a*;*b*). His argument is based on the observation that Ā-movement is sensitive to wh-islands and complex NP-islands, from which he concludes that the bounding nodes in Duala are NP and IP. Hence, cross-clausal Ā-movement must apply successive-cyclically at least through every SpecCP position in order not to violate the Subjacency Condition in (2). This reasoning is based on the assumption that there is only a single specifier per phrase. As for extraction from wh-islands, the wh-phrase in the embedded SpecCP occupies the only escape hatch, and hence, extraction of another wh-phrase is blocked. In Minimalism, however, a phrase can have multiple specifiers. So, island-sensitivity only provides evidence that a dependency involves movement, but it does not say anything about the locality of movement. Insofar, Biloa does not provide an argument for *successive-cyclic* Ā-movement in Duala either.

in chapter 1). At S-structure, an Ā-moved XP is in its final landing site (S-structure feeds spell-out at PF). Consequently, (126) predicts that operator-C-agreement can only apply with the C head in whose specifier XP surfaces, deriving pattern II. Actually, there is no opacity in the analysis: We look at the output structure with the generalization in mind that agreement applies between C and an XP in SpecCP. On the surface, this configuration only holds in the final landing site of XP, i.e. the generalization is surface-true. Thus, postponing agreement until S-structure yields pattern II: The XP is not in intermediate positions anymore at S-structure, and hence it cannot trigger agreement in these positions. However, more needs to be said, the constraint in (126) does not yet do what it is supposed to do.

To uncover the complication, let us see how variation is accounted for. We know that Chamorro also has a pattern I reflex (wh-agreement), i.e. there is agreement with XP also in intermediate positions. But since at S-structure the Ā-moved XP only surfaces in its final landing site, how can there be agreement in these positions? The solution is to assume enriched representations: Movement of XP leaves behind an empty element in every intermediate landing site, viz. a trace. The agreement rule at S-structure refers to all occurrences of XP, including traces of XP. And indeed, Chung's (1998) constraint for wh-agreement in Chamorro makes reference to traces (it involves an agreement relation between representations of XP and the functional head I):[54]

(127)   *Wh-agreement (Chung 1998: 257)*:
        I^0 and an A-bar-bound trace that is free within I^0's minimal m-command domain must have compatible values for [Case].

This in turn raises the question why there is no agreement with traces in pattern II languages. The assumption seems to be that traces vary in their ability to be controllers for agreement. If pattern II is to be derived, traces must not be taken into account; if , however we want to derive pattern I, traces must be controllers. So what is missing in (126) is the explicit statement that traces do not count as Associates (= controllers for agreement) for operator-C-agreement. To summarize, for Chung (1998), movement is subject to the same locality conditions in pattern I and pattern II languages (which is desirable given the evidence in section 2.4.1), and agreement applies at S-structure; variation between languages reduces to the status of traces as agreement controllers: In some languages they are, in others they are not.[55] In this way, the two patterns can be derived, but the central assumption strikes me as questionable for languages like Chamorro that have both pattern I and pattern II triggered by a single movement operation. Traces left by one and the same instance of Ā-movement must count as agreement controllers (for wh-agreement) and at the same time they must not (for operator-C-agreement); it is at best odd that an element (the trace) behaves so differently with regard to the same process (agreement) within a single language.[56, 57]

---

[54]See Georgopoulos (1991: 105) for a related proposal for pattern I wh-agreement in Palauan.

[55]Note that it would in principle be possible to account for pattern III with agreement at S-structure if the agreement constraint exclusively held for traces and not for their XP antecedent.

[56]It is not completely impossible to make the analysis work for languages like Chamorro because the reflexes occur on different heads (I vs. C). One could thus say that traces in SpecIP are agreement controllers whereas traces in SpecCP are not, but this seems very unattractive as it comes close to a restatement of the facts.

[57]An alternative way to account for variation would be to say that traces are always agreement controllers but that some movement types leave traces while others do not. Those that do not can only trigger pattern II reflexes and those that do have pattern I reflexes. But the problem remains the same: How can a single

2.4.5 Opacity resolved by enriched representations

In this subsection I present two other types of approaches that assume that movement is successive-cyclic in pattern II languages, (AII) and (AIII). The common background assumption is that the locality of movement does not vary between languages. Since movement must be successive-cyclic in languages with pattern I, this must also be the case in languages with pattern II. Hence, there is opacity with pattern II (although the term *opacity* is never used): Some movement steps feed agreement with the moved element and others don't. The work I will discuss in this section is actually the only work I know of that is seriously concerned with variation, i.e. the authors are aware of the existence of other patterns beside the classic pattern I and try to account for it.[58] As pointed out in section 2.2.5, more recent (Minimalist) work on reflexes of successive-cyclic movement by and large ignores other patterns. The only exception is the paper by Reintges et al. (2006), which I will discuss in section 2.5.3.1.

The approaches that will be presented are an instance of the enriched representations strategy to opacity, as introduced in section 1.2.2. Two subtypes of this strategy can be distinguished:

(AII) A special type of empty element is left in intermediate positions.

(AIII) There are constraints on empty elements in intermediate positions.

I will address each of them in turn.

2.4.5.1 Quasi-operators

Let me begin with analysis (AII), versions of which are proposed in Haïk et al. (1985); Tuller (1986); Haïk (1990); Biloa (1993); Ouhalla (1993). These are approaches that analyse morphological reflexes of Ā-movement in terms of Ā-binding (for the concept of Ā-binding see Aoun 1985; 1986; Finer 1985; Aoun and Li 1989). In a nutshell, the binding principles are extended to the Ā-domain: The binder occupies an Ā-position. The special morpheme that occurs under Ā-movement, e.g. *no* in Duala, is an anaphor that requires a local binder in an Ā-position. If there is movement of an XP to the local SpecCP, the moved XP ends up in an Ā-position and can thus act as a binder for the anaphor, whose occurrence is thereby licensed. If, however, Ā-movement does not apply, the anaphor is not locally bound and ungrammaticality results. In this way, the presence of a special morpheme is tied to Ā-movement. The anaphor is in complementary distribution with a pronominal element (which is zero in Duala). This pronominal element must not be locally bound from an Ā-position (Principle B). Hence, it can only occur if there in no movement to SpecCP and thus no local Ā-binder; if Ā-movement applies, a Principle B violation obtains, resulting in ungrammaticality.[59] Note that under such an approach, the morpheme that shows up under movement is not literally a *reflex* of movement because its occurrence is not caused by movement. Rather, allomorph choice and the availability of movement conspire: The allomorph (anaphor vs. pronoun) can be freely chosen in any derivation, there is no trigger for the choice. The completely independent application of movement or its omission

---

[58]Indeed, they only talk about patterns I and II, but this might simply be due to the fact that the existence of pattern III was unknown at that time. All examples of pattern III that I know of were obtained from work published since 2005 (with the exception of the syntactic reflex in German), cf. section 2.3.3.

[59]See also Epée (1975); Clements (1984*a*); Chung and McCloskey (1987); Schneider-Zioga (1995) for Ā-binding approaches to morphological reflexes of movement. In contrast to the work discussed here, these authors do not address cross-linguistic variation in the distribution of morphological reflexes.

will have consequences for the licensing of the chosen allomorph – it either licenses the allomorph or it doesn't, depending on the nature of the allomorph chosen.[60]

Let me come to the central question now: How is the difference between pattern I and II derived? Haïk et al. (1985); Haïk (1990); Biloa (1993); Ouhalla (1993) all work within the GB framework. Binding applies at S-structure. They assume that intermediate positions in a movement chain are occupied by traces, i.e. representations are enriched. In languages with pattern I, traces of an $\bar{\text{A}}$-moved XP can act as local binders, just like the XP in its final $\bar{\text{A}}$-position (at S-structure). If $\bar{\text{A}}$-movement makes a stop-over in every SpecCP position, the anaphor (= the morphological "reflex" of movement, abbreviated as 'R') is licensed in every clause crossed by movement because it is locally bound by the XP in its final landing site or a trace left by the XP, cf. (128):

(128)   *Pattern I: traces are $\bar{\text{A}}$-binders*:
$[_{CP}$ XP$_k$ $[_{C'}$ ... $\mathbf{R}_{+anaph}$ ... V $[_{CP}$ t$_k$ $[_{C'}$ ... $\mathbf{R}_{+anaph}$ ... V t$_k$ $]]]]$
         └─ $\bar{\text{A}}$-binding ─┘   └─ $\bar{\text{A}}$-binding ─┘
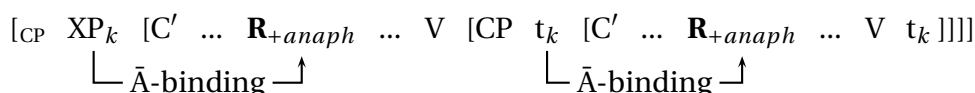
In languages with pattern II, however, it is stipulated that traces cannot act as binders, only the XP itself can do so. Consequently, the anaphor is only licensed in the clause where the XP has its final landing site, cf. (129), resulting in a pattern II reflex.

(129)   *Pattern II: traces are not $\bar{\text{A}}$-binders*:
$[_{CP}$ XP$_k$ $[_{C'}$ ... $\mathbf{R}_{+anaph}$ ... V $[_{CP}$ t$_k$ $[_{C'}$ ... $\mathbf{R}_{+anaph}$ ... V t$_k$ $]]]]$
         └─ $\bar{\text{A}}$-binding ─┘   └─ $*\bar{\text{A}}$-binding ─┘
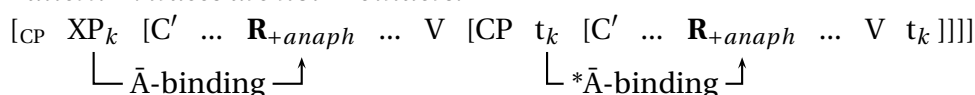
Traces that cannot bind anaphors are called "quasi-operators" (Haïk 1990 attributes the term to Tuller 1986). Hence, cross-linguistic variation arises because there are two types of traces: One type can be a binder and the other one cannot; languages choose between the two types of traces.[61] This account is highly problematic for languages like Chamorro in which a pattern I and a pattern II reflex occur simultaneously, triggered by a single instance of $\bar{\text{A}}$-movement: The traces left by movement must be binders (to derive pattern I) and must not be binders (to derive pattern II) at the same time, which is paradoxical.

Although pattern III is not treated in this work, we can imagine what the analysis would have to look like: We would have to assume that only traces can act as binders, but the XP in SpecCP cannot do so. However, this is at odds with the general practice: That traces are deficient in some sense is a standard assumption in GB (they do not seem to possess all of the features their antecedent bears); this is reflected in the fact that traces

---

[60] An important question is how general the $\bar{\text{A}}$-binding approach to reflexes of movement is: Can it handle all types of reflexes of movement, viz. addition, replacement, and deletion? Take a language where the morphological reflex is not an additional, more marked morpheme (as in the languages which the $\bar{\text{A}}$-binding approach has been applied to), but rather subtractive in nature (the disappearance of a morpheme under $\bar{\text{A}}$-extraction). Usually, the pronominal allomorph is taken to be the default, phonologically zero allomorph, and the anaphor is the special form, i.e. the more marked, non-zero allomorph. But in languages with subtractive reflexes, the zero / less marked morpheme must be the non-default anaphor and the overt / more marked morpheme is the default form. This is questionable from an iconicity point of view where the morphologically more marked form corresponds to the more marked syntactic choice; nevertheless, it would be feasible to apply the $\bar{\text{A}}$-binding approach to these languages. What is, however, highly problematic are *syntactic* reflexes of movement such as inversion in Spanish. What is the anaphoric or the pronominal element in this case? The morphological form of the verb does not change, only its syntactic position. I don't see a simple way to subsume syntactic reflexes under the $\bar{\text{A}}$-binding approach.

[61] This is similar though not identical in spirit to the proposal by Chung (1998), introduced in section 2.4.4. In Chung's work, all traces are of the same type, but not all of them are taken into account for certain agreement rule.

(and other empty elements) are subject to special licensing constraints that ensure that properties required e.g. for interpretation of the empty element can be recovered from the antecedent: As for traces, the relevant constraint was the *Empty Category Principle* (ECP, Chomsky 1981); one condition of the ECP was that traces that are not in a position that is selected by a lexical item, must be governed by their antecedent. In light of this, it is not an odd assumption of the Ā-binding approaches to postulate that traces can be deficient when it comes to binding. But if anything is in principle able to be a binder for a trace (under the right structural relations), then it is the antecedent of the trace (the occurrence in the final landing site at S-structure). So, the assumption that antecedents cannot be binders in some languages, which is required under the Ā-binding approach to account for pattern III, is questionable.

To summarize, the "quasi-operators" approach requires that (a) occurrences of a moved XP in intermediate and final positions can be distinguished (trace vs. antecedent), and crucially, (b) that two different types of traces can be distinguished (binders vs. non-binders). We will see that the analysis I propose in section 2.5 can dispense with assumption (b).

2.4.5.2   Constraints on empty elements

Let me now come to analysis (AII) of opacity in pattern II languages: It is also couched within the GB-framework and relies on enriched representations. In addition to the postulation of traces in intermediate positions of a movement chain, there are operations and constraints that apply to traces. There are two versions of this analysis: Either traces initially share a certain class of features with their antecedent (those that are involved in the reflex) or they do not. Let us assume for the sake of illustration that we want to analyze a morphological reflex on complementizers: A C head indicates the ϕ-features of the phrase moved through its specifier. In the first version of the approach, traces bear (at least) the ϕ-features of their antecedent. Hence, what we initially expect is to find pattern I because intermediate C heads can agree in ϕ-features with the trace in their specifier. In order to derive pattern II, there must be a constraint in the respective languages that demands the deletion of the ϕ-features in traces, or of traces in general before the C head tries to agree in ϕ-features with them (deletion of traces was independently required for completely different reasons, cf. Lasnik and Saito 1984). If traces are deleted before C tries to establish agreement with them, trace deletion bleeds agreement. Since traces only occur in intermediate positions, we derive pattern II: Bleeding only affects intermediate positions, but we still find the reflex on the C head in the clause where the moved phrase lands: The antecedent is not subject to deletion because it is not a trace. This is proposed by Haïk (1990) as one possible solution to the opacity problem with pattern II.

The second version goes in the opposite direction; it is advanced in Collins (1993: 187ff.): The presupposition here is that, initially, traces do not share the agreement-relevant features of their antecedent (ϕ-features in our current abstract example). Thus, we would expect a pattern II reflex: Only the antecedent in the final landing site bears the relevant features and can thus trigger a reflex on the head in whose specifier it sits. In order to derive pattern I, there must be an operation that enforces copying of the agreement-relevant features from the antecedent to the trace(s) in the chain it heads. If agreement between C heads and the element in their specifier applies after copying, we get pattern I: Copying to traces feeds agreement between C and the trace. In languages with pattern II, however, there is a constraint that prohibits this copying operation.[62] In order to integrate

---

[62]Since Collins (1993) is concerned with Ewe subject pronoun choice, which is optional in non-final

pattern III languages into the system, there must be an operation that deletes the relevant features from the antecedent of the traces in Haïk's (1990) system, which would be feasible. Collins's (1993) analysis requires more operations to capture pattern III than it needs for patterns I and II: First, copying must apply to traces, and afterwards the features must be deleted on the antecedent. While feasible as well, the account of pattern III shows that Collins would need both copying *and* deletion; Haïk only requires deletion to capture all three patterns.

The problems with this approach are the same as in the quasi-operators approach (cf. section 2.4.5.1): If a language has both patterns I and II triggered by a single instance of Ā-movement, deletion or copying must both apply and not apply to traces of the same chain, which is paradoxical. From a conceptual point of view, the approach not only needs to postulate traces, as (AI), but requires that operations and constraints apply to traces. The analysis I will propose dispenses with such operations and constraints.

2.4.5.3 A Minimalist reanalysis

In the following paragraphs, I want to investigate whether the GB approaches to opacity can be successfully reformulated within (a representational version of) Minimalism. Two issues arise when we want to translate (AI) – (AIII) into Minimalism: First, there is no distinct S-structure level anymore; second, trace theory has been abandoned. In paragraph 2.4.5.3.1, I will briefly summarize the Minimalist approaches to movement that have replaced trace theory. These will be applied to (AI) – (AIII) in the subsequent paragraphs. Another general issue is the following: Approaches (AI) – (AIII) all assume that agreement between a head and the element in its specifier (antecedent or trace) applies late, viz. after all movement operations have applied. However, Minimalism assumes a derivational model of syntax where movement (internal Merge) and agreement (Agree) are interleaved, according to standard assumptions. Furthermore, structure-building must be in accordance with the Cycle. Applying agreement in the syntax but after all movement operations would therefore be counter-cyclic. If every XP is a cyclic domain, this is the case at least in intermediate positions: Agreement between the head H^0 and the phrase in its specifier only affects HP, which is already dominated by another phrase. So, is it possible at all to translate analyses (AI) – (AIII) into a Minimalist framework? It is, in fact. It has repeatedly been proposed in the Minimalist literature that agreement / Agree is in fact a postsyntactic PF-operation (cf. Marantz 1991; Ackema and Neeleman 2004; Fuß 2007; Bobaljik 2008; Miyagawa 2010 among others). Hence, it applies after all syntactic operations (including movement). Thus, late Agree is not completely incompatible with Minimalist assumptions. For the sake of the argumentation, I will adopt PF-Agree in the following Minimalist reanalyses.

2.4.5.3.1 Excursus: intermediate occurrences in Minimalism

In the Government and binding era, intermediate positions in a movement chain are occupied by traces. Traces encode the position of the moved element in what would correspond to previous stages of the derivation. We have seen in section 1.2.2 that constraints could refer to traces, which was necessary (among other things) to derive transparent and opaque interactions. Minimalism dispenses with traces because traces are not part of the numeration; hence, the insertion of traces in intermediate positions violates the *Inclusiveness Condition*:

---

clauses (pattern I vs. pattern II), he proposes that the copying operation is optional. But if a language only exhibits pattern II, copying must be strictly prohibited.

(130) *Inclusiveness Condition (cf. Chomsky 1995:228)*:
       Material that is not part of the numeration is inaccessible throughout a derivation.

For the same reason, other abstract material whose main purpose is to encode previous stages of a derivation is abandoned as well (indices, etc.). A desideratum is to derive restrictions on the distribution of elements by derivational constraints on operations that only require information from the local context; in this way, information can be successively forgotten, minimizing the computational burden (cf. Chomsky 2001 et seq.). Representational or global constraints (filters) that need to take into account large parts of the structure (and thus require information on previous locations of an XP) are to be avoided.

Various proposals have been made about the kinds of elements that can occupy intermediate landing sites in a movement chain:

(a) the copy theory of movement (Chomsky 1995)

(b) no intermediate representation (Lasnik 1999*a*, Epstein and Seely 2002)

(c) the multi-dominance approach (McCawley 1982; Gärtner 2002; Starke 2001; Frampton 2004; Abels 2004; 2012; Citko 2005; Wilder 2008; Gracanin-Yuksek 2009; de Vries 2013; Johnson 2009)

Under the copy theory of movement an XP is base-merged in a position P_1. If XP moves to a position P2, a copy of XP in P_1 is made; this copy is then merged in P_2, and so forth for all subsequent movement operations. Thus, final and intermediate landing sites are occupied by copies of XP. This is illustrated in (131) for movement of the internal argument of a transitive verb, labeled XP, to SpecCP with an intermediate step in SpecvP (there may be more intermediate steps but the exact number and locus of intermediate steps does not matter for the present discussion). Normally, only one of the copies is pronounced (usually the highest one):

(131)   *Copies in intermediate positions*:
          $[_{CP}$ **XP** $[_{C'}$ C $[_{TP}$ T $[_{vP}$ **XP** $[_{v'}$ DP_{ext} $[_{v'}$ v $[_{VP}$ V **XP** $]]]]]]]]$

The approach in (b) assumes that movement of XP does not leave behind anything in intermediate positions. When XP moves to SpecvP, it is present in this position when we look at the vP-stage of the derivation. Then it moves on to SpecCP; when we look at the vP after the final movement step to SpecCP has taken place, there is nothing in SpecvP. As a result, the only occurrence of XP on the surface is in its final landing site, cf. (132). For illustration, I box the positions that XP occupied at various stages of the derivation, but it should be kept in mind that there is literally nothing in these positions; there once was, but is not anymore. As a consequence, movement destroys sisterhood relations: XP is the sister of v′ at the vP-level, but when looking at the CP, XP is no longer the sister of v′ because there is no occurrence of it left in this position.

(132)   *Nothing left behind in intermediate positions*:
          $[_{CP}$ **XP** $[_{C'}$ C $[_{TP}$ T $[_{vP}$ ☐ $[_{v'}$ DP_{ext} $[_{v'}$ v $[_{VP}$ V ☐ $]]]]]]]]$

Under a multi-dominance approach to movement, there is only a single occurrence of the moved XP (no copies etc.) that is merged in various positions in the structure. This means that XP is immediately dominated by more than one node. When XP moves from within the VP to SpecvP, it is remerged in SpecvP but still remains the sister of V; in addition, it becomes the sister of v′. I will call the branches from XP to the positions it is linked to by remerger *chain links* in what follows, represented by dashed lines. In (133) three chain links are created: XP is linked to the complement position of V, to SpecvP and to SpecCP.

(133)   *Multi-dominance approach to movement*:

```
                    CP
                   ⁝ ╲
                   ⁝    C′
                   ⁝   ╱  ╲
                   ⁝  C    TP
                   ⁝      ╱  ╲
                   ⁝     T    vP
                   ⁝         ⁝ ╲
                   ⁝         ⁝  v′
                   ⁝         ⁝ ╱ ╲
                  XP ⋯⋯ DPₑₓₜ    v′
                    ⁝            ╱  ╲
                    ⁝           v    VP
                    ⁝ ⋯⋯⋯⋯⋯⋯⋯⋯       ╲
                                        V
```
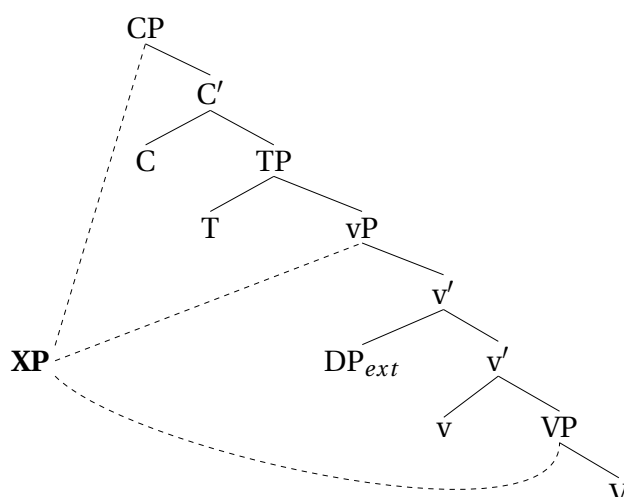
Each of these hypotheses has its advantages and drawbacks, but I will not evaluate them here. For the following discussion, it is sufficient to know how movement is modeled in these approaches. The analysis I will propose in section 2.5 is compatible with each of them anyway. With this much background on the nature of occurrences in intermediate positions, we can try to reformulate analyses (AI) – (AIII) in Minimalism.

2.4.5.3.2 Levels of representation in Minimalism Chung's (1998) analysis crucially refers to S-structure. As pointed out above, there is no distinct S-structure anymore that the agreement rule / operation can refer to (cf. figure 1.7 in chapter 1). However, one could mimic the basic idea by applying agreement at PF. Under this assumption, pattern II arises if agreement only holds between a head and the occurrence of the moved phrase in the position where it is phonologically realized. As for pattern I, all occurrences of the operator, including the ones in intermediate positions, are agreement controllers. However, if one adopts the proposal that movement does not leave behind anything in intermediate positions, pattern I (and also pattern III) cannot be produced under Chung's late agreement analysis because there is nothing to Agree with in intermediate positions at PF; only pattern II could be derived. As for the copy theory and the multi-dominance approach to movement, there are occurrences of the moved XP in intermediate positions; however, it is not trivial to distinguish them from the occurrence in the final position, which is a prerequisite for Chung's approach though to account for patterns II and III. The distinction between final and intermediate occurrences will be the topic of the next paragraph. Note finally that the general problem with this approach also arises in the Minimalist version: In languages in which patterns I and II co-occur, occurrences in intermediate positions must and must not count as agreement targets at the same time.

2.4.5.3.3 Quasi operators in Minimalism In this paragraph, I investigate whether the Ā-binding analysis can be transferred to Minimalism. First of all, we need to assume that binding applies late, after all movement operations (as we did for Agree). The approach relies on empty elements (traces) left by movement. Recall that trace insertion violates the *Inclusiveness Condition* in (130). Hence, we first need to replace traces by copies, real gaps, or an XP that is linked to several positions (multi-dominance). If movement does not leave anything behind, the Ā-binding approach cannot derive patterns I and III: An

anaphor does not have a local binder if there is a gap in the closest intermediate landing site that c-commands it ("nothing" cannot bind anything). Pattern II can be derived, however: The occurrence in the final position can be a binder for the anaphor in the final clause, but not for those in non-final clauses (anaphors need to be locally bound, viz. within the minimal clause). A different kind of problem arises under the copy theory of movement and the multi-dominance approach: The occurrences of the moved XP in intermediate position may not be distinguishable from the occurrence of XP in its final landing site. Under Chomsky's (1995) view, copies of XP are completely identical to the element they are copies of.[63] And given multi-dominance, there is literally the same element in all landing sites because there is only a single occurrence of the moved XP in the first place. But a distinction between final and intermediate movement steps is required for patterns II and III in the quasi-operators approach: Binding of the moved XP applies only in one of these position, hence, binding must somehow make reference to the distinction. Note that this problem did not arise in GB, where final and intermediate landing sites were, by assumption, occupied by different elements: The latter are occupied by traces and the former is occupied by the moved XP itself. Therefore, it was easy to make the distinction. A rule / constraint that refers exclusively to traces automatically only affects intermediate positions; a rule / constraint that refers to XP only affects the final landing site. Since this distinction is lost in the copy theory of movement and the multi-dominance approach, there is a serious problem. The same holds for the approach by Chung (1998), cf. paragraph 2.4.5.3.2.

This problem can be solved, however: Proponents of multi-dominance approaches or the copy theory of movement must also distinguish between occurrences in final and intermediate positions when it comes to linearization at PF and interpretation at LF: Usually, only one occurrence of an element in a movement chain is pronounced, often this is the one in final position; the occurrence in the final position is also privileged when it comes to the calculation of scope, but the ones in intermediate positions are usually not. Therefore, there are suggestions for how to distinguish final and intermediate positions in these two approaches to movement. Let me give one example for each of the two approaches: Abels (2012: 144f.), working within a multi-dominance approach, defines the final position of a remerged XP as follows: The highest occurrence of XP is the one whose mother dominates all other occurrences of XP. As for the copy theory of movement, Nunes (2004: 70ff.) proposes that feature checking (or feature valuation) affects only one copy in a chain, and not all members of it (contra Chomsky 1995, cf. footnote 63). Feature-checking happens in final positions. The copy in the lower position still contains an unchecked feature. Thus, final and intermediate positions can be distinguished by making reference to the status of their features (checked vs. unchecked). Chains are unified only at LF.

So, at least if traces are replaced by copies or by a multi-dominance approach to movement, there might be a way to translate the Ā-binding approach into Minimalism, but it is impossible to keep the analysis if movement doesn't leave behind anything. What remains for the Ā-binding approach, however, regardless of whether we use copies or chain links (multi-dominance), is that we need to distinguish two types of them: those that can act as binders and those that cannot. In pattern II languages, occurrences in intermediate positions cannot act as binders, while in pattern I languages they can, just like the occurrence in the final position; in pattern III languages only occurrences in inter-

---

[63]Chomsky (1995: 381, n.12) assumes that if one copy in a chain is affected by an operation, all other members of that chain are affected as well. In this way, all copies of a XP are always identical.

mediate position can be binders. Note that this leads to the same problem as in Chung's (1998) approach when it comes to languages where a single instance of Ā-movement has both a pattern I and a pattern II reflex, as e.g. Chamorro: The occurrences of the moved element left in intermediate positions must be able to act as binders and at the same time must not be able to do so, which is paradoxical.

2.4.5.3.4   Constraints on empty elements in Minimalism   Let me finally come to analysis (AIII) and its Minimalist reformulation. According to (AIII), operations apply to empty elements in intermediate positions (copying / deletion of features). First, if movement of an XP does not leave behind anything in the position XP is moved from, the operations cannot apply to intermediate positions: It is impossible to delete or copy features from / to nothing. Consequently, patterns I and III cannot be derived. As for the copy theory and the multi-dominance approach to movement, it is first of all necessary to make a distinction between occurrences in final and intermediate positions. Let us assume for the sake of the argument that this can be done (see the last paragraph for a distinction in terms of structural relations or feature content). In addition, operations must apply to occurrences in intermediate positions. Let us see how (AIII) fares under the copy theory of movement. Copying to occurrences in intermediate positions in not necessary under Collins's (1993) approach because copies already contain the features of their antecedent (since they are copies). Under Haïk's (1990) approach, we simply need to add a deletion operation to intermediate copies. Things are worse for the multi-dominance approach if features need to be deleted on intermediate occurrences in order to derive pattern II (cf. Haïk 1990): Since there is only a single occurrence of XP that is linked to several positions in the structure, deleting features on that occurrence will automatically affect *all* positions in the chain. Thus, the distinction between occurrences in final and intermediate positions required for patterns II and III cannot be made. To conclude, Minimalist implementations of approaches that apply operations to occurrences in intermediate positions only work under the copy theory of movement.

I will briefly summarize the Minimalist reanalyses: When transferred to Minimalism, analyses (AI) and (AII) are only compatible with the copy-theory or the multi-dominance approach to movement; analysis (AIII) is even more restricted as it can only be formulated under the copy theory. A general requirement for all three analysis is that occurrences in final and intermediate positions can be distinguished, which may not be trivial under the copy theory and the multi-dominance approach. If movement does not leave behind anything, however, none of the analyses can capture patterns I and III. In addition, all approaches run into a paradox in languages where a pattern I and a pattern II reflex co-occur because the occurrences in intermediate positions are then subject to two conflicting requirements.

2.4.6   Interim conclusions

In this section I have discussed three hypotheses ((HI) – (HIII)) that deny opacity for pattern II languages: (HI) and (HII) assumed that there is no reflex in intermediate positions in pattern II languages because there is no movement through these positions; and according to (HIII) there is no reflex because there is no agreement in intermediate positions. All hypotheses were shown to be empirically untenable: Island-sensitivity suggests that movement is involved in the first place (at least in some pattern II languages), and the co-occurrence of patterns I and II argues against movement in one fell swoop without intermediate stop-overs; default markers in non-final clauses suggest that there has been

an attempt to Agree. Hence, we are dealing with opacity in pattern II (and pattern III) languages: A head H can in principle enter into agreement with an XP in its specifier, and movement applies successive-cyclically. However, agreement underapplies in certain positions. Not all specifiers of HP to which an XP is moved agree with H, only some do. The crucial distinction is between final and intermediate landing sites in a movement chain.

Approaches (AI) – (AIII), formulated within the GB framework, acknowledge the opacity of pattern II. The absence of agreement in some positions is accounted for by postponing agreement to S-structure. In addition, occurrences in intermediate positions are either ignored, or different types of intermediate occurrences are postulated; alternatively, operations and constraints apply to these occurrences before agreement takes place. These three approaches encounter several problems when transferred to Minimalism: First of all, agreement would have to be counter-cyclic unless it is postponed to PF; second, it is not clear whether a distinction between final and intermediate occurrences of a moved XP – inherent in GB's trace theory of movement – can be made under the copy theory and the multi-dominance approach to movement. The approaches cannot derive all patterns if movement does not leave behind anything or if the multi-dominance approach is adopted. Each analysis requires additional assumptions / machinery beside the final / intermediate occurrence distinction: different types of intermediate occurrences (cf. approaches (AI) and (AII)), or operations and constraints on intermediate occurrences (cf. approach (AIII)).

In the next section, I will present a unified analysis in which agreement is a syntactic operation and applies in accordance with the Cycle. It does not require any of the additional assumptions of the previous analyses, neither different types of intermediate occurrences nor operations on them. In addition, this analysis will be compatible with all three approaches to movement in Minimalism (copy theory, multi-dominance, gaps).

2.5   An ordering approach to opacity

The desideratum of this section is to provide a uniform Minimalist analysis for all patterns of reflexes of successive-cyclic movement. We have seen that patterns II and III are opaque: Agreement between a head H and a phrase moved to its specifier underapplies in some positions.[64] The state of the art for pattern II languages are approaches that make use of enriched representations. However, these approaches have a number of empirical and conceptual problems: Some of them cannot account for all patterns of reflexes, and they cannot be easily upheld under more recent theoretical assumptions, especially concerning movement. What I propose is that these problems can be avoided if we replace the enriched representations approach by the traditional transformational grammar analysis of opacity: sequential rule application / rule ordering. Sequentiality can be straightforwardly formulated within Minimalism because it is a derivational framework. In this section, I introduce the core idea of the analysis, the assumptions, as well as the

---

[64]There is an important distinction between *local* and *global opacity*: Pattern II, for example, is locally opaque, i.e. at the phrase level: When looking at an HP with an XP moved to its specifier (intermediate movement step) but without agreement between H and XP, it is unclear why movement to SpecHP does not feed agreement with H (counter-feeding); the same holds for the final movement step in pattern III languages. However, pattern II is not opaque when we evaluate opacity globally, i.e. at the clause level: Movement of an XP feeds Agree with a head H if XP is in SpecHP; at the CP-level, this is only the case in the final clause, so pattern II is expected. Pattern I, however, is globally opaque because there is agreement between XP and H in intermediate clauses although XP is not present in intermediate SpecHPs at the CP-level; the same holds for pattern III. Here, I am concerned with local opacity.

derivations for patterns I – IV. In addition, I argue that the ordering analysis is superior to the enriched representations approaches because it provides a simple and elegant way to capture the co-occurrence of different patterns and optionality of patterns without the additional assumptions required by previous approaches.

2.5.1 Core idea of the analysis

Reflexes of movement are the result of a successful Agree relation between a head H and the XP moved to SpecHP. Patterns II and III suggest that there is an important distinction between final and intermediate movement steps: In pattern II languages, only the former feed Agree with H, whereas the latter do not, they counter-feed Agree. The reverse holds for pattern III. The patterns can be derived as follows:

(134)   *Core idea of the analysis*:
        Final and intermediate movement steps apply at different points of the deriva-
        tion. In particular, they may be interleaved with Agree when movement and
        Agree operations are triggered by the same head H. The movement step that ap-
        plies before Agree feeds Agree because it puts the XP in the right structural posi-
        tion to be an accessible agreement controller for H. Movement steps that apply
        after Agree come too late to feed Agree, viz. they counter-feed it; the XP is not in
        the right structural position to be accessible for the Agree-triggering feature on H
        at the point of the derivation where H seeks for an agreement controller. The se-
        quence of operations is encoded by the order of the operation-inducing features
        on H. Cross-linguistic variation arises through reordering of these features.

In light of the discussion in the previous section that it is not trivial to distinguish between final and intermediate movement steps by referring to the position a moved XP occupies (at least under the copy theory and the multi-dominance approach to movement), I as- sume that the difference rather lies in the *trigger* of final and intermediate movement steps. That is, there is not a single feature that triggers Merge; rather, different types of Merge triggers must be distinguished: a trigger for final and a trigger for intermediate steps. Consequently, the two types of Merge-inducing features can be ordered differently relative to Agree, one type applying before and the other after Agree. The result is opacity: Some movement steps feed Agree, others counter-feed it. That intermediate movement steps have designated triggers is in itself not a new proposal (see e.g. the notion of *edge feature* in Chomsky 2000 et seq. that triggers intermediate steps). What is new is that the the triggers for final and intermediate movement steps can apply at different points of the derivation. The timing is detectable from the success or failure of the Agree relation that interacts with the two types of movement. Thus, variation in the patterns of reflexes of movement offers morphosyntactic evidence for the split of Merge types.

     In section 2.5.2 I present my assumptions about syntactic structure-building, agree- ment, and the syntax-morphology interface. Section 2.5.3 contains abstract deriva- tions that illustrate how the interactions of Merge and Agree predicted by reordering of operation-inducing features derive all the patterns of movement reflexes. Finally, in sec- tion 2.6, I present detailed analyses of the data introduced in section 2.3.

2.5.2 Assumptions

The analysis is couched within the Minimalist Program (Chomsky 1995 et seq.). As intro- duced in 1.2.3, the basic architecture of grammar is the following:
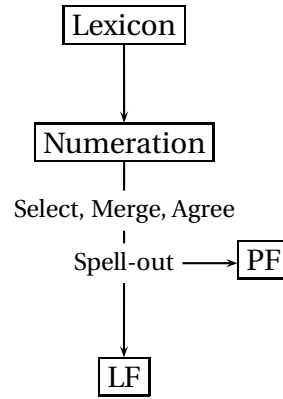
Figure 2.1: The Y-model in Minimalism, extended

Items required for the derivation are taken from the lexicon and are assembled in the numeration. In the syntax, Merge builds up complex syntactic objects out of lexical items. The operation Select applies in the numeration and marks an item for being merged next. Select and Merge apply until the numeration does not contain any elements anymore. The structure thereby created is the input for the PF- and the LF-interface. In what follows, I summarize my assumptions about the presyntactic numeration, the postsyntactic morphological component, and structure-building in syntax.

2.5.2.1 Syntax

2.5.2.1.1 Clause structure   I adopt the basic clause structure in (135):

(135) $[_{\text{CP}}$ C $[_{\text{TP}}$ T $[_{\text{vP}}$ DP$_{ext}$ $[_{\text{v'}}$ v $[_{\text{VP}}$ DP$_{int}$ $]]]]]$

The internal argument (DP$_{int}$) is merged as the sister of V. The external argument (DP$_{ext}$) is introduced in the specifier of the functional head v that takes the VP as its complement (Chomsky 1995; Kratzer 1996). There are two functional projections above vP: TP and CP. SpecCP is the target for (final) Ā-movement. We will see in the analyses of the individual languages that a few more functional projections within the CP-domain might be needed (CP-recursion or a split CP à la Rizzi 1997). The structure unfolds step by step in a bottom-up fashion by alternating applications of the basic operations Merge and Agree (incremental structure-building).

2.5.2.1.2 Operations   The definitions of the two basic operations Merge and Agree are repeated in (136):

(136)   *The elementary operations Merge and Agree (cf. Chomsky 2000; 2001)*:
  a.   Merge:
      (i)   Merge is a structure-building operation that takes two items α and β and creates a new item γ from α and β: $[_{\gamma}\ \alpha\ \beta]$.
      (ii)   Merge is triggered by structure-building features [•F•] on a head.
  b.   Agree (preliminary version):
      (i)   Agree copies feature values from an element α (the goal / controller) to an element β (the probe / target): The probe P is initially unvalued and receives a value under Agree with the goal G if P and G are in a c-command relation.
      (ii)   Agree is triggered by probe features that seek for a value: [∗F:□∗].

The syntactic structure grows by the application of Merge. Merge comes in two types: external Merge and internal Merge (Chomsky 2004). As for external Merge, at least one of the elements that is to be merged is taken from the workspace, i.e. it is independent from the element it will be merged with. If none of the elements is taken from the workspace, we speak of internal Merge: One of the two elements is c-commanded by the other before they Merge; thus, internal Merge is movement. The operation Merge is subject to the *Extension Condition* (Chomsky 1995: 190): Every Merge operation must extend the tree at the root of the current phrase marker, viz. there can be no counter-cyclic Merge.

Agree is an operation that relates two elements in the structure and transfers feature values from one of the elements to the other. Agree is triggered by probe features that are unvalued. Probes are located on heads and seek for a value on an XP. I use Agree to model agreement between a phrase moved to SpecHP and the head H. The value acquired by H under Agree is realized morphologically. Agree is subject to the following conditions:

(137)   *Definition of Agree (based on Chomsky 2000; 2001)*:
        Agree between a probe P and a goal G applies if

        a.   P and G are in an asymmetric c-command relation,
        b.   P has an unvalued feature [∗F:□∗] and G has a matching valued feature [F],
        c.   G is the closest matching goal for P.
        d.   Result: G values P.

The definition of asymmetric c-command is given in (138):

(138)   *Asymmetric c-command*:
        An element α asymmetrically c-commands an element β iff β is included in the sister of α.

The condition in (137-a) leaves open whether the probe must c-command the goal (downward Agree) or whether the goal must c-command the probe (upward Agree):

(139)   *Probe c-commands goal (downward Agree)*:
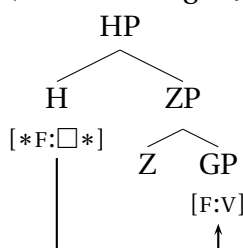
(140)   *Goal c-commands probe (upward Agree)*:
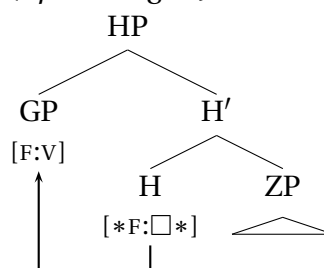
I left the direction of Agree in (137) unspecified because I follow Baker (2008*b*) in assuming that both options exist (see also Carstens 2012). He proposes that there is a parameter on the direction of Agree and languages choose one of the options ('F' corresponds to a functional head with a probe feature):[65]

---

[65]The standard assumption in GB and early Minimalism was that agreement applies in a specifier-head configuration, i.e. upward (Chomsky 1986*a*; Kayne 1989; Kinyalolo 1991; Koopman 1992; Chomsky 1995, see Assmann et al. (2012) for more recent arguments for the privileged status of Spec-head Agree, though downward Agree is not excluded). Chomsky (2000) et seq. holds that Agree solely applies downwards, which has been the prevailing view for about a decade. In recent years, however, a lively debate about

(141)   *The Direction of Agreement Parameter (Baker 2008b: 215)*:

     a.   F agrees with DP / NP only if DP / NP asymmetrically c-commands F (upward Agree), *or*

     b.   F agrees with DP/NP only if F c-commands DP/NP (downward Agree).

The evidence Baker (2008*b*) provides for the necessity of having both upward and downward Agree is the following (see also Collins 2004): In some languages, agreement is sensitive to the post-movement configuration and in others it is not. Put differently, movement feeds new agreement relations in the former but not in the latter. For example, both English and some Bantu languages have agreement between T and the XP in SpecTP; this XP is usually the logical subject (the agent) of an active verb. However, these languages also exhibit locative inversion. In this construction, it is the locative XP that moves to SpecTP. Whereas in the Bantu languages that have this construction, movement of the locative feeds agreement between the locative XP and T, it does not do so in English. Rather, we still get agreement with the logical subject, which stays within the vP. Similarly, some Bantu and some Indo-European languages show complementizer agreement in class and ϕ-features, respectively. But whereas the form of the complementizer covaries with the class of the XP in SpecCP in the Bantu languages, it is constant in Indo-European languages: In the latter, the complementizer always agrees with the subject of the embedded clause in its c-command domain, regardless of the features of the XP in SpecCP. In addition, there is variation with respect to in-situ constructions: In Kinande, for example, operators (question words, focussed XPs) agree with the complementizer in class, but only when they are in SpecCP. Kinande also allows for wh-in-situ, but then the complementizer cannot agree with the operator, suggesting that the complementizer can only look upwards for a goal. As pointed out above, the complementizer in the respective Indo-European languages can very well agree with the subject in its c-command domain, suggesting that it looks downwards. Baker concludes that if movement does not change agreement relations and if agreement with an XP can be established without movement of XP (XP in-situ), Agree applies downwards (probe c-commands goal). If, however, movement of XP to SpecHP feeds agreement between H and XP and if this agreement is excluded if XP stays in the c-command domain of H (in-situ), Agree applies upwards (goal c-commands probe).[66] We have seen that in all the languages introduced in section 2.3, the

---

the direction of Agree has arisen. Koopman (2006); Wurmbrand (2012); Zeijlstra (2012) argue that Agree is only upwards (Koopman restricts this to Spec-head agreement, Wurmbrand and Zeijlstra also allow for upward Agree at a distance if DPs act as probes). This view has been challenged by van Koppen (2011); van Koppen et al. (2011); Preminger (2013) who argue that downward Agree is needed after all, at least for ϕ-Agree. Béjar and Řezáč (2009) claim that the direction of Agree changes in the course of the derivation: Probes initially search downwards (first cycle Agree), but can look upwards into their specifier if no goal is found on the first Agree cycle.

[66]It has been proposed in the literature that upward Agree is actually downward Agree between H and XP plus obligatory subsequent movement of XP to SpecHP. This is implemented as follows: The probe on H is accompanied by an EPP-feature that triggers movement of the agreed-with XP (see e.g. Carstens 2005; Reintges et al. 2006). As Řezáč (2004) points out, there is a general problem with this set of assumptions: How can the EPP-feature "know" which XP was the goal of Agree triggered by the probe feature it accompanies? Some kind of memory needs to be introduced in order to store the relevant information, e.g. an index on the goal. The EPP must be sensitive to this index. However, elements like traces and indices are to be avoided in Minimalism whenever possible. Another, perhaps more serious problem is raised by Baker (2008*b*: 173). He points out that there are upward Agree relations in predicate adjective constructions where the goal is outside the projection of the head that bears the probe, viz., the goal has never been in the c-command domain of H at any point of the derivation. This can be derived under pure upward Agree (not restricted to Spec-head Agree), but not under the probe+EPP-analysis for upward Agree.

morphological / syntactic reflex is tied to overt movement; in particular, it does not occur if the operator stays in-situ (in case the languages allow this). Following Baker's argumentation, I conclude that agreement in these languages applies upwards, which immediately explains the absence of a reflex with in-situ operators (Indonesian, Kikuyu, Spanish, Duala, Kitharaka, Dinka).[67] For the languages that do not allow for XP-in-situ constructions, I also conclude that Agree applies upwards when the agreement exponent on H covaries with the features of the XP in SpecHP (Wolof, Chamorro). For the remaining languages without XP-in-situ (Ewe, Irish, German) I will also assume that Agree applies upwards, although there is no morphological evidence; but in light of the problems of the alternatives discussed in footnote 66 (downward Agree plus EPP-movement) I consider this a reasonable assumption. The focus of the present chapter is thus on reflexes of movement based on upward Agree; in chapter 3 I will present reflexes that result form downward Agree.

The definition of Agree contains a minimality condition, see (137-c). It will be relevant for downward Agree discussed in chapter 3. Closeness is defined in terms of asymmetric c-command: α is closer to β than γ if β asymmetrically c-commands both α and γ, and if α asymmetrically c-commands γ.

It has been assumed that internal Merge (movement) is parasitic on Agree, i.e. an Agree relation between H and an XP is a prerequisite for Merge of XP with H' (cf. Chomsky 2000; 2001; 2004). I do not adopt this assumption. Internal Merge probably requires matching between the structure-building feature and the corresponding feature of the to-be-moved XP, but not Agree in the sense of feature transfer from XP to H (for the independence of internal Merge from Agree see e.g. Nevins and Anand 2003; Boeckx 2003; Bobaljik and Wurmbrand 2005; Baker 2008*b*, and Richards 2009 making that claim for a certain type of movement). If internal Merge (movement) was parasitic on Agree, Agree could feed Merge, but not the other way around; however, we have seen in the data that movement can feed Agree as well.

What ultimately triggers the operations is that structure-building features and probe features are uninterpretable features that cannot be interpreted at the interfaces; therefore, they need to be discharged in the course of the derivation (*Full Interpretation*, Chomsky 1995: 27). A structure-building feature [•F•] is discharged if the head on which it is located merges with an element that has a matching feature [F]. A probe feature [∗F:□∗] is discharged if it receives a value v from a matching feature on a goal: [F:v]. Discharged features are invisible for further syntactic operations (but maybe not for postsyntactic operations). With respect to structure-building features, I assume that Merge with an element that bears a matching feature is the only way to discharge the feature. As for probe features, however, I follow Řezáč (2004); Anand and Nevins (2005); Preminger (2011*b*) in assuming that they need not necessarily be valued to get discharged, it is sufficient if they attempt to find a goal. That means, Agree is not optional; a probe must seek for a goal. But if there is no goal with a matching valued feature in the search space of the probe, the failure to receive a value does not lead to a crash of the derivation. There are two ways

---

[67]Note that there is variation when it comes to in-situ constructions and the occurrence of a reflex of movement. In the languages I introduced in section 2.3, there is evidence that the wh-phrase really stays in situ; for instance, there is no island-sensitivity (when the scope position and the base position are separated by an island). Consequently, there is also no reflex of movement. In contrast to this, Reintges et al. (2006) show that wh-in-situ does trigger a morphological reflex in Passamaquoddy (pattern I) and Egyptian Coptic (pattern II). However, they argue that the wh-phrase in these languages does undergo movement before spell-out (because movement is island-sensitive), but that only the lowest copy is pronounced. Hence, these data are not problematic for the present approach. As long as there is movement in the syntax, it can feed upward Agree. Which occurrence in the movement chain is pronounced is determined at PF and does not interfere with the establishment of Agree relations in syntax.

to get an unvalued probe feature discharged (applying immediately after the proe feature has tried to find a goal): Either, it is deleted by default (as a last resort to prevent the crash of the derivation), or it is assigned a default value. Languages choose between these options. They correspond to the effects we have seen in the data: If there is no agreement with the moved XP, we either get no marker at all or a default exponent.

Apart form the conditions under which structure-building and probe features can be discharged, syntactic operations are subject to two other requirements. The first is the *Earliness Principle* (Pesetsky 1989) in (142), which demands that operations apply as soon as their context is met:

(142)   *Earliness Principle*:
         An operation-inducing feature must be discharged as soon as possible.

The second requirement is that syntactic operations apply in accordance with the *Strict Cycle Condition*, a stricter version of the *Cycle* (see section 1.2.1.3).[68]

(143)   *Strict Cycle Condition (SCC, based on Chomsky 1973: 243)*:
         a.   No operation can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.
         b.   Every XP is a cyclic node.

We have seen in the discussion of the Cycle in section 1.2.1.3 that the smaller the cycles the more orders of operations can be predicted. The strictest version would thus be one in which every projection was a cyclic node, as proposed by McCawley (1984; 1988). I will justify in section 3.3 why assuming that every projection is a cyclic domain is too strong; I thus make use of the slightly weaker version where the relevant cyclic domain is the XP.

2.5.2.1.3   Locality   We have seen in sections 2.1 and 2.3 that there is massive evidence that movement applies successive-cyclically. I enforce this by the *Phase Impenetrability Condition* (PIC):

(144)   *PIC (Chomsky 2000: 108)*:
         In phase α with the head H, the domain of H is not accessible to operations outside α; only H and its edge are accessible to such operations.

The domain of a phase head H is the complement of H; the edge comprises specifier(s) of HP as well as adjuncts to HP. The PIC requires that an XP that is to discharge a structure-building feature on a head Y moves through the edge (i.e. the specifier) of every phase that intervenes between the base position and the final landing site (SpecYP) of XP in order to remain accessible for Y. If it did not move to the edges of intervening phases, XP would be trapped inside the domain of a phase dominated by YP. As a consequence, the structure-building feature on Y could not be discharged and hence, *Full Interpretation*

---

[68]One might wonder whether there is any difference between the *Strict Cycle Condition* and *Cycle*. As pointed out by Perlmutter and Soames (1979: 146f.) and revisited in Pullum (1992: 210f.), the *Cycle* does not *per se* exclude that a rule can apply to a cyclic domain embedded in another cyclic domain; such an application is indeed possible if the context for the application of a rule R_1 in the lower domain is created by the application of another rule R_2 in the higher domain, and as long as R_1 has not already applied in the larger domain. The *Strict Cycle Condition* excludes such cases; in fact, it prohibits any application of a rule in a lower domain if it is contained in a larger domain. See Chomsky (1995); Bošković and Lasnik (1999); Collins (1997); Freidin (1978; 1999); Kitahara (1997); Watanabe (1995) for attempts to derive the SCC.

would be violated. There is massive evidence from reflexes of movement that CP and vP are intermediate landing sites (and thus phases in the present system); I will argue in chapter 3 that SpecTP and Spec*n*P (a DP-internal head) are also stop-over positions. For concreteness, I thus assume that *every phrase* is a phase. The PIC then requires that a phrase makes a stop-over in the edge of every XP on the way to its final landing site. While Merge is subject to the PIC (to enforce intermediate movement steps), I follow Bošković (2007*b*) in assuming that Agree is not.[69] This will be important when we look at downward Agree, which can apply at a (certain) distance.

2.5.2.1.4 Triggers for final vs. intermediate steps I assume that all syntactic operations are feature-driven (see Chomsky 2008 for a different view). Thus, intermediate movement steps must be triggered as well. The crucial assumption for the analysis is that **intermediate movement steps have different triggers than final movement steps**. The **designated triggers of intermediate steps** are the edge features [•EF•] (Chomsky 2000; 2001; 2008). Intermediate movement steps apply only for locality reasons, viz. to keep a potential goal accessible for a structure-building feature from outside (cf. the PIC in (144)). I assume that a head does not bear the triggers for intermediate steps inherently, but only if required by the context (see section 2.5.2.3 for details). Final movement steps, on the other hand, do not apply for locality reasons. Rather, they apply in order to check a specific feature such as the wh-feature on interrogative C or the EPP on T, and to get an XP into a certain structural position thereby (compare e.g. the Wh-criterion that demands that wh-phrases surface in SpecCP). I will call these features *final features* in what follows. In contrast to final features, edge feature are assumed to be (categorially) underspecified structure-building features.

At this point, one might think of final movement steps as criterial movement steps in the sense of Rizzi (2004; 2007). Rizzi (2004: 1) defines a a criterial position as "[...] a position dedicated to the expression of some scope-discourse properties. [...] criterial positions terminate chains: a phrase meeting a criterion is frozen in place, and its chain cannot extend further (Criterial Freezing)." However, final positions in this thesis do not have the same properties as Rizzi's criterial positions, and therefore, I will not use the label *criterial*. There are two respects in which final positions are different form criterial positions. For one, a final position does not have to be terminal. If XP moves to SpecHP to discharge a final movement trigger, I do not want to exclude per se that it can move on to a structurally higher position to check another final movement trigger, although this will not be crucial for the analysis of reflexes of movement.[70] Indeed, the assumption that criterial movement is always terminal causes problems for languages in which T has the EPP-property. The XP that checks the EPP must not be moved to SpecCP, which it should be, however, if it is e.g. a wh-phrase that needs to check [•WH•] on C. See Rizzi and Shlonsky (2004) for discussion of this problem. The second and more important reason why movement to a final position is not identical to movement to a criterial position is that a criterial position is associated with scope-taking and discourse-properties. However, we have seen in the data in section 2.3 that the last movement step of an XP that undergoes partial

---

[69]See also Bobaljik and Wurmbrand (2005) on different locality domains for Agree and (A-)movement.

[70]Such a derivation might be required for scrambled wh-words in German: Clause-bound wh-movement does not exhibit superiority effects, which can be attributed to the independently available scrambling operation: A wh-word may scramble above another and is then closer to the [•WH•]-feature on C than the one it has been moved across, giving rise to what looks like a minimality violation on the surface. If the scrambling-triggering feature is a final feature as well, final scrambling must not block subsequent checking of the final wh-movement trigger.

movement always behaves like a 'real' final movement step to the scope position under full movement with respect to the morphological reflex it triggers. Hence, the trigger for the partial movement step must be a final feature. However, it could not be a criterial movement step: Partial movement of XP terminates in a position (in overt syntax) that is not its scope position. In case of full movement to the scope position, the final position in my system is identical to Rizzi's criterial position, but as for partial movement, it is not. Thus, the label *final* is not to be confused with *criterial*.

Let me add a few remarks on partial wh-movement. Since the last movement step to the surface position of a wh-phrase in a partial movement construction behaves like a final movement step, it must be triggered by a final feature in the analysis I will present. What kind of feature is this? Two options come to mind: Either, non-interrogative C heads (that do not project the scope position of a wh-phrase) can also bear a [•WH•]-phrase; or there must be another final feature that is distinct from [•WH•] and from the edge feature. As for the former solution, scope taking cannot be associated with [•WH•]-checking anymore if non-interrogative C heads can bear this feature as well. With respect to the second option, Sabel (2000) proposes that partial wh-movement is triggered by a focus feature [+FOC] ($\cong$ [•FOC•]) on the C head that hosts the wh-phrase on the surface, whereas the final movement step is triggered by a wh-feature on the C head (in addition to a FOC-feature). I will adopt the general idea: The last movement step under partial movement is triggered by a final feature (e.g. [•FOC•]) that is different from the final feature that triggers the final step under full movement ([•WH•]), and different from intermediate triggers.[71]

2.5.2.2 Morphology

I assume a postsyntactic realizational model of morphology: That is, syntax solely operates on abstract feature bundles and the morphology interprets the output of syntax by linking these abstract features with phonological information. For concreteness, I adopt the framework of Distributed Morphology (DM, Halle and Marantz 1993; 1994; Harley and Noyer 1999). In DM, the linking elements are the vocabulary items (VIs) which are pairings of morphosyntactic features (in square brackets, to the right of the arrow) with phonological features (in slashes, to the left of the arrow), cf. (145); in addition, A VI can contain morphosyntactic features that restrict its insertion context (context features, to the right of the slash).

(145)   *Abstract representation of a VI*:
         /X/ ↔ [A, B, C] / ___ [D]

VIs are inserted into terminal nodes of the syntactic structure. In order to be inserted, the morphosyntactic features of the VI have to be a subset of those of the terminal node. Since VIs may be underspecified with respect to their morphosyntactic features, several VIs may compete for insertion into a terminal node. This competition between VIs is resolved by the *Specificity Principle*, cf. (146). According to this principle, the most specific VI is chosen, viz. the VI which shares the most features with the terminal node.

(146)   *Subset Principle*:
         A VI α is inserted into a terminal node T, iff (i) and (ii) hold:

         (i)   The morphosyntactic features of α are a subset of the features of T, and
         (ii)   there is no other VI β that fulfills (i) and is more specific than V.

---

[71]See also Abels (2012: 70) for the assumption that partial wh-movement has a trigger that is different from the [•WH•]-feature that triggers full wh-movement.

DM has a number of operations that manipulate the abstract feature bundles before vocabulary insertion applies. This may lead to the choice of an unexpected VI, deriving syntax-morphology mismatches. The operation that I will make use of is impoverishment (Bonet 1991; Noyer 1992; 1998; Halle and Marantz 1993; Bobaljik 2002; Frampton 2002; Nevins 2003): Impoverishment rules delete features on terminal nodes prior to vocabulary insertion. They take the form in (147), where a feature [F] is deleted in the context of a feature [K]. The context feature and the feature undergoing deletion must be in a sufficiently local configuration. I assume that they must be located on the same terminal node.

(147)   *Abstract impoverishment rule*:
         [F] → Ø / __[K]

To summarize, DM assumes the following architecture of grammar (ignoring the presyntactic numeration and the semantic component branching from the syntactic component):
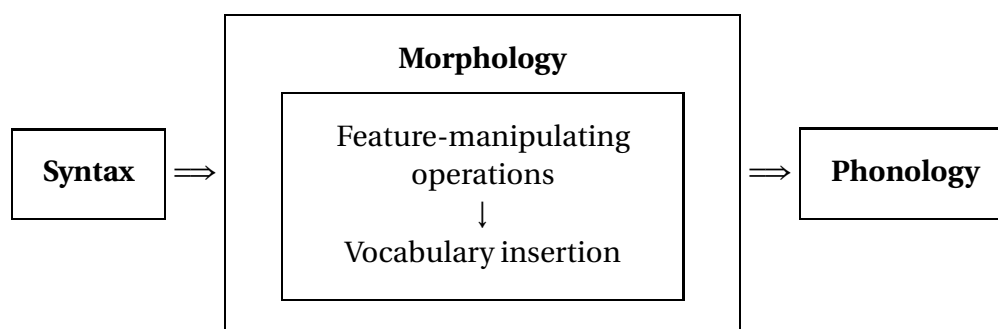
Figure 2.2: Architecture of the postsyntactic morphological component

2.5.2.3   Numeration

Following the Minimalist model of grammar, the functional and lexical elements used in the derivation are taken from the lexicon and are assembled in the presyntactic numeration. The element that will be merged next in the derivation is selected before it actually enters the derivation (Chomsky 1995: 248). At the very first step of the derivation, two elements in the numeration are selected for Merge (both elements that take part in Merge are still in the numeration at this point).

I assume that operations can apply to the numeration. I make use of a feature-insertion operation because I assume that edge features [EF] are not freely available on heads, i.e. they are not part of the lexical entry of a head in the lexicon. Instead, I assume that they are inserted on a head when they are needed. Edge feature insertion upon requirement in the numeration has two advantages: First, applying it in the numeration avoids an overgeneration problem. If edge features were freely available, it could be that intermediate movement steps are triggered for which there is no final trigger, i.e. superfluous and thus prohibited movement steps could apply (that would have to be rules out by transderivational economy). Second, edge feature insertion in the syntax violates *Inclusiveness* (cf. (130)) because an element is assigned a feature it did not have in the numeration; if the edge feature is assigned in the numeration, *Inclusiveness* is trivially fulfilled (cf. Biskup 2011: 37ff.). Furthermore, we will see evidence in the analyses of Dinka (cf. section 2.6.1.3), Hungarian and Icelandic (cf. chapter 3) that edge feature insertion applies

very early, before the head that bears the edge feature has triggered any non-edge feature-driven movement. This is guaranteed if edge feature insertion applies in the numeration.

The question is how exactly edge features are inserted; how can it be determined in the numeration whether they are required for the derivation? Here, I adopt a proposal by Heck and Müller (2000; 2003) who introduce the constraint *Phase Balance* (see also Fischer 2004 and the related constraint *Feature Balance* in Biskup (2011: 48)). I propose the modified version in (148):[72]

(148)   *Phase Balance (based on Heck and Müller 2000; 2003, applies in the numeration)*:
An edge feature [•EF•] is inserted on the selected phase head H for every feature [•F•] on a head Y in the numeration if:

a. $Y \neq H$,
b.   there is no accessible matching feature [F].

(149)   *Accessibility*:
A feature [F] is accessible if it is part of the workspace and not selected (workspace: lexical items in the numeration, previously generated trees that are unconnected to the current phrase marker).

Since *Phase Balance* makes reference to selected heads, it applies after the operation Select. Recall that I assume that every phrase is a phase, thus every head is a phase head and *Phase Balance* thus applies to every head. According to (144), an edge feature is inserted on a head H selected for Merge if there is another head Y in the numeration that has a structure-building feature [•F•] (a final movement step trigger), and if there is no element E with a matching feature [F] left in the workspace; rather, the element E is already part of the current phrase marker M. What edge feature insertion guarantees is that E bearing [F] will be attracted to SpecHP (the edge of the phase) in order to remain accessible for the structure-building feature [•F•] on the head Y that will be merged at a later stage of the derivation. If no edge feature was inserted on H, the element E, included in M, would end up in the domain of H because M becomes the complement of H. Thus, it would not be accessible for [•F•] and the derivation would crash (cf. the PIC in (144)). The condition in (148-a) avoids edge feature insertion on H for the feature [•F•] if H bears [•F•] itself: The final movement step to SpecHP will be triggered by this feature, so an edge feature would be superfluous. Hence, *Phase Balance* triggers intermediate movement steps through every XP.

Let us look at an example of *Phase Balance* at work. Assume we want to derive a clause where the internal argument of a transitive verb is questioned, as in (150):

(150)   What do you read?

For the initial Merge operation, two items are selected in the numeration: the verb and the internal argument. At this stage, there is a structure-building feature on a head in the numeration that is not equal to the selected head V: the [•WH•]-feature on C. In addition, there is no accessible element with a matching feature [WH]. Though the DP with [WH] is still in the workspace, it is not accessible because it is selected for Merge, cf. (191). Hence, an edge feature is inserted on V.[73] Consequently, the internal argument will be

---

[72] *Phase Balance* in Heck and Müller (2000; 2003) is a violable OT-constraint which is checked at the phrase level; it thus applies in the syntax. In contrast, my version of their constraint is not violable; it applies in the numeration (just like Biskup's *Feature Balance*) and thus below the phrase level. Because of these differences, some modifications are required.

[73] The exception in (191) that selected items do not count as accessible is necessary to enforce edge feature

moved to SpecVP to discharge the edge feature after DP$_{int}$ merged with V.[74]Edge feature insertion also applies to v and T: When v or T is selected for Merge, there is still the C-head with the structure-building feature in the numeration, and the element with the matching feature [WH] is not accessible because it is already part of the current phrase marker. In this way, the wh-phrase moves from phrase to phrase and thereby remains accessible. Edge feature insertion does not, however, apply when the head which projects the final landing site of the wh-phrase is selected: At this point, there is still a head in the numeration with a structure-building feature, namely C itself. However, since the selected head is equal to the head which bears the structure-building feature in question, no edge feature is inserted, cf. (190-a). C attracts the wh-phrase and thereby discharges its [•WH•]-feature.

Finally, there is a last operation in the numeration: ordering of features. If Select and *Phase Balance* have applied and a head ends up with more than one operation-inducing feature, these features need to be ordered according to their timing of application. Ordering, in particular of Agree and Merge triggers, is necessary because we get transparent interactions between them: Movement to SpecHP can feed Agree between H and the XP in SpecHP. As discussed at length in chapter 1, only sequential application can produce transparent interactions, simultaneous application cannot. I assume that the operation-inducing features are ordered on a stack of which only the topmost feature is accessible. The second topmost feature only becomes accessible when the topmost feature is discharged and so forth. The effect is the same as that of the *Strict ordering hypothesis* (see section 1.2.1.2): Every operation-inducing feature can only attempt to trigger an operation at a certain point of the derivation, viz. when it is on top of the stack. It cannot trigger an operation before that point because the feature is not accessible when another feature is on top of it on the stack. The feature can also not trigger an operation at a later point, even if its context would be given at that stage of the derivation, because it will have been discharged already (by default or through matching / valuation); if it wasn't discharged, the derivation would be doomed to crash anyway because *Full Interpretation* is violated.

The order of features is determined by language-specific *ordering statements* that are part of the grammar. They take the form in (151) (where [•F•] stands for any final movement trigger and [•EF•] for an edge feature that triggers intermediate movement steps):

(151)   *Abstract ordering statement*:
         [•F•] ≻ [∗F∗] ≻ [•EF•]

---

insertion on the first merged head, such as V in the example. Without it, the wh-phrase merged with the first head would count as accessible because it is still part of the workspace when *Phase Balance* is checked, and therefore the wh-phrase would end up in the complement of V, i.e. in V's domain. Note that even with this condition, edge feature insertion on a selected head H is still suppressed if there is an additional [WH]-feature in the numeration in addition to the selected item with the same feature [WH]. This is necessary in multiple questions (in languages without multiple wh-movement) where the later merged wh-phrase will ultimately satisfy the [•WH•]-feature on C, making successive-cyclic movement of the first merged wh-phrase unnecessary. See Heck and Müller (2003) for discussion of configurations of this type.

[74]It is necessary to ensure that the edge feature on V (or any other movement trigger) is not checked in-situ by the complement of V. To enforce movement of DP$_{int}$ to the specifier of V, I assume that a structure-building feature can only be discharged by an XP that asymmetrically c-commands the H that bears the edge feature; in this way, an edge feature on a head H triggers movement to SpecHP. The wh-DP$_{int}$ in the abstract example in the text does not *asymmetrically* c-command V, and hence, it must undergo movement to SpecVP. I thank Martin Salzmann for pointing this problem and the solution out to me.

It has to be read as follows: If a head H has a subset of the features [•F•], [•EF•], [∗F∗] they are ordered on a stack such that the feature to the left is on top of the stack, the second feature from the left is immediately below, and the right-most feature on the list is the bottom-most feature on the stack. In prose, the ordering statement in (151) says that final movement steps apply before Agree, and intermediate movement steps apply after Agree when they are triggered by the same head H. The ordering statement includes all operation-inducing features a head may ever bear and it specifies their relative order. But in practice, a head will usually only bear a subset of the features mentioned in the ordering statement. What matters is the order of the features relative to each other. For example, if a C head in a language with the ordering statement in (151) bears a [•WH•]-feature and a [∗ϕ∗]-probe, the former is on top of the stack (being a trigger for final movement steps) and the latter is at the bottom:

(152)   Ordering on C:

$$C \begin{bmatrix} [\bullet\text{WH}\bullet] \\ [\ast\phi\ast] \end{bmatrix}$$

After the features of a head are ordered on a stack in the numeration, the selected head is merged into the syntactic structure and discharges its operation-inducing features step by step, from the top to the bottom of the stack.

To summarize, I assume the following sequence of operations in the presyntactic component:
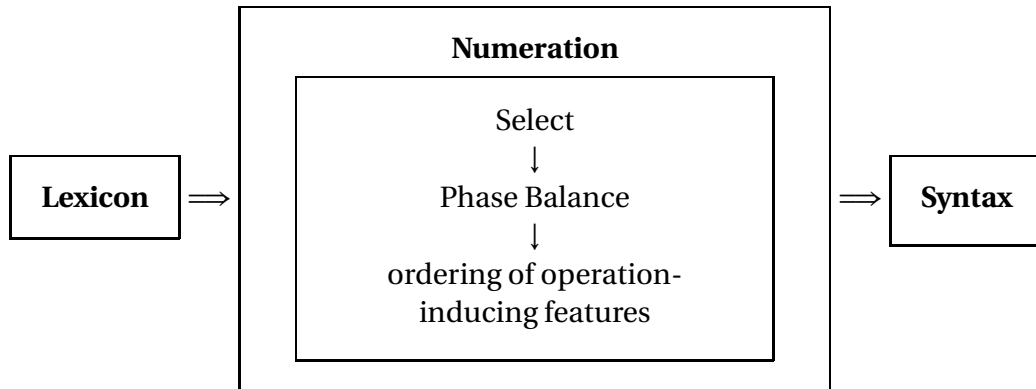
Figure 2.3: Architecture of the presyntactic numeration

2.5.3   Abstract analyses

In this section I go through the abstract derivations for patterns I to IV that arise through the interaction of Merge and upward Agree. I will argue that in order to derive the cross-linguistic variation, two types of Merge triggers need to be distinguished: triggers for final and for intermediate movement steps (final features vs. edge features). Note that traces in the trees are inserted only as mnemonic devices; they do not have any theoretical status.

2.5.3.1   Patterns I – IV

I assume that reflexes of movement on a head H are the result of an Agree relation between H and the XP moved to SpecHP. In the cases at hand, Agree applies upwards. Due to stacking of operation-inducing features and feature-discharge from top to bottom, upward Agree automatically reduces to Spec-head Agree: When a head H probes upwards,

it either finds an XP in the specifier of HP or not. If there is an XP and it is a suitable goal according to the restrictions in (137), the probe on H agrees with it and is thereby discharged. If XP is not a suitable goal (e.g. no matching features) or if there is no XP at all in SpecHP, the probe on H is discharged by default (default valuation of default deletion). Then, the next feature on the stack of H is discharged. Now, imagine Agree initially fails but the relevant Agree configuration arises at a later stage of the derivation, either because a goal XP is merged to SpecHP after Agree attempted to apply, or because a goal XP is merged in the specifier of a projection that dominates HP. In both cases, Agree between H and XP cannot be established because it is too late: The probe was discharged at an earlier point of the derivation; we cannot "keep it alive" and wait whether it will find a goal at a later stage of the derivation because without immediate discharge, the next feature on the stack of H could not be accessed, viz. structure-building could not even continue. Thus, when I talk about upward Agree, it always implies Spec-head Agree and not upward Agree at a distance.[75]

With this in mind, let us look at how Merge can interact with Agree if both operations are triggered by a single head H. In general, a Merge operation that applies *before* Agree feeds Agree between H and the XP merged in SpecHP. This is so because at the point when H probes upwards, there is an XP in SpecHP, cf. (153) (in the examples, Agree is triggered by a ϕ-probe; v is a variable over ϕ-values).

(153) *Internal Merge feeds Agree*:

a. Initial structure

$$H'$$

$$
\begin{array}{ccc}
& H & WP \\
& \begin{bmatrix} [\bullet F\bullet] \\ [{*}\phi{:}\square{*}] \end{bmatrix} & \text{... DP } [\phi{:}v] \text{ ...}
\end{array}
$$

b. First step: Merge, [•F•] discharged

$$HP$$

$$
\begin{array}{ccc}
\text{DP } [\phi{:}v] & & H' \\
& H & WP \\
& \begin{bmatrix} [\not\bullet\not F\not\bullet] \\ [{*}\phi{:}\square{*}] \end{bmatrix} & \text{... } t_{DP} \text{ ...}
\end{array}
$$

❶ Merge

---
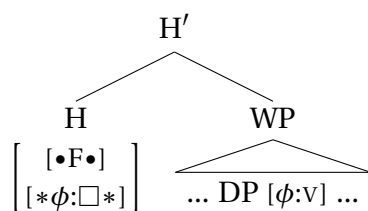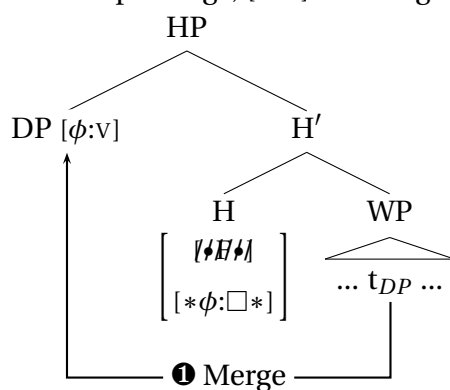
[75]Baker (2008*b*: 192f.) provides an empirical argument that upward Agree is at least sometimes restricted to Spec-head Agree: In Kinande, a preposition P can only agree with its complement in class if this complement is extracted from the PP (surfacing in a projection that dominates PP), but not if the complement stays in-situ. This implies that P probes upwards. But if it can look upwards, the question arises why P cannot agree with another DP that c-commands it, e.g. the subject of an active transitive verb in a sentence like "Kambale is speaking with the old people". P is c-commanded by the subject DP in this example and should thus be able to agree with it. Baker proposes that these facts are derived if PP is a phase and if upward Agree is restricted to Spec-head Agree. Movement of the complement of P out of the PP must go through SpecPP, and in this position it can trigger agreement on P. But the subject DP that is generated outside of PP cannot trigger agreement on P because it is not in SpecPP at any stage of the derivation.

c. Second step: Agree

```
              HP
          ┌────┴────┐
     DP [φ:V]       H′
        ↑       ┌────┴────┐
        │       H         WP
        │    [∗φ:□∗]    ┌──┴──┐
        └─❷ Agree ─┘    ... t_DP ...
```

d. Result: probe valued

```
              HP
          ┌────┴────┐
     DP [φ:V]       H′
                ┌────┴────┐
                H         WP
             [φ:V]    ┌──┴──┐
                      ... t_DP ...
```
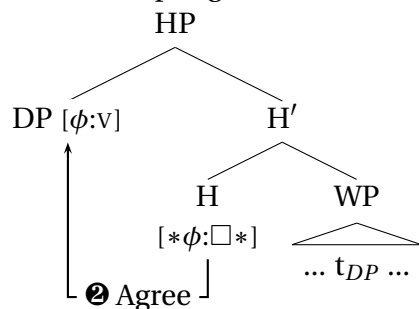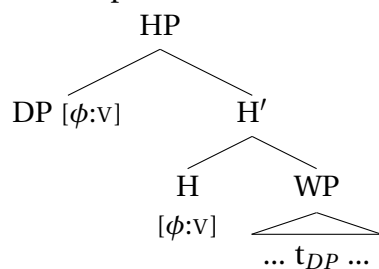
A Merge operation that is triggered *after* Agree counter-feeds Agree: At the point of the derivation where the probe attempts to Agree, there is no XP in SpecHP. The probe is thus discharged by default. A potential goal XP is moved to SpecHP only afterwards, but this is *too late* to feed Agree – the probe has been discharged before, cf. (154) (with default deletion of the probe; alternative: default valuation).

(154) *Internal Merge counter-feeds Agree*:

a. Initial structure

```
              H′
          ┌────┴────┐
          H         WP
      ⎡[∗φ:□∗]⎤  ┌──┴──┐
      ⎣ [•F•] ⎦   ... DP [φ:V] ...
```
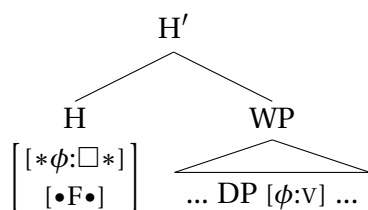
b. First step: Agree, no goal found

```
    ┌──────────────→ H′
    │            ┌────┴────┐
❶ Agree          H         WP
    │        ⎡[∗φ:□∗]⎤  ┌──┴──┐
    └────────⎣ [•F•] ⎦   ... DP [φ:V] ...
```
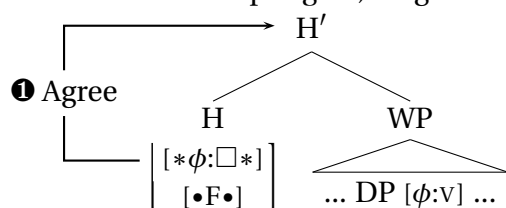
c. Last resort: probe deleted by default

```
              H′
          ┌────┴────┐
          H         WP
      ⎡[∗̶φ̶:̶□̶∗̶]⎤  ┌──┴──┐
      ⎣ [•F•] ⎦   ... DP [φ:V] ...
```

d. Second step: Merge, [•F•] discharged

$$
\text{HP} \\
\text{DP } [\phi\text{:v}] \qquad \text{H}' \\
\text{H} \qquad \text{WP} \\
\textit{[•H•]} \\
\text{... t}_{DP} \text{ ...} \\
\text{❷ Merge}
$$
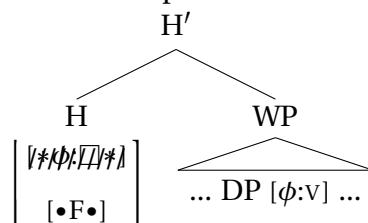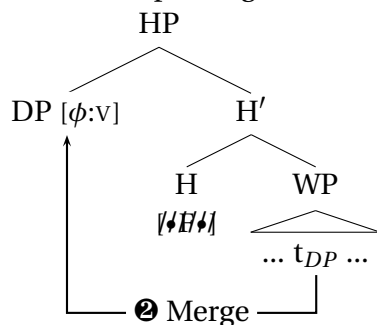
On the surface, the configuration in (154-d) is opaque: Assume that XP has a matching feature value for the probe on H. When looking at the HP, XP is in the right structural configuration to be a goal for the probe on H, nevertheless, Agree fails. Put differently, it does not apply although its context is given; Agree is counter-fed / underapplies. If the XP is not moved at all to SpecHP, we also do not get Agree, just as in (154); the only difference is that the XP is not in SpecHP at any point of the derivation.

This approach presupposes ordering of Merge relative to Agree, which has been argued to be required for transparent interactions of the two operations before (see section 1.2.3 for references). As for the reflexes of movement, we thus expect a reflex whenever Agree is fed by movement, as in (153) (movement applies early enough), and we expect a reflex to be absent if Agree is counter-fed by movement, as in (154) (movement applies too late). This derives patterns I and IV. Now, in languages with pattern II or pattern III, where some movement steps feed Agree and others don't, a problem arises. We would need the following order of operations:

(155)   *Order of Merge and Agree in languages with pattern II or III*:
        Merge > Agree > Merge

But this order is symmetric; Merge cannot be predicted to apply both before and after Agree if the parameter choice is Merge > Agree or Agree > Merge. This paradox is resolved if the Merge operations that apply before Agree are of a different type than the ones applying after Agree. That is, instead of (155), we have (156):

(156)   *Order of Merge$_1$ and Merge$_2$ to Agree in languages with pattern II or III*:
        Merge$_1$ > Agree > Merge$_2$

If the ordering statement can make reference to this difference, it does not only order Merge relative to Agree but rather Agree relative to Merge$_1$ and Merge$_2$. What would be the difference between the two types of Merge? We have seen that in pattern II and pattern III languages, it is about whether the movement step is final or intermediate. Thus, I propose that we can derive the opacity in these languages if we distinguish subtypes of Merge, viz., final vs. intermediate movement steps. They will be distinguished by their triggers, as pointed out in the assumptions: Intermediate movement steps are triggered by edge features, final movement steps have different triggers. Consequently, the two types of internal Merge operations can apply at different points relative to Agree; thus, one movement type feeds and the other counter-feeds Agree. The four patterns of reflexes of movement arise by reordering of operation-inducing features on the head H that triggers Agree and Merge, see the options in 2.2.

I represent the order of features on H in the form of ordering statements; a comma between operation-inducing features means that they are not ordered to each other.[76]

order of features	interactions		pattern
	final steps	intermediate steps	
a. [•F•], [•EF•] > [∗F:□∗]	feed Agree		P I
b. [•F•] > [∗F:□∗] > [•EF•]	feed Agree	counter-feed Agree	P II
c. [•EF•] > [∗F:□∗] > [•F•]	counter-feed Agree	feed Agree	P III
d. [∗F:□∗] > [•F•], [•EF•]	counter-feed Agree		P IV

Table 2.2: Orders of operation-inducing features: final vs. intermediate internal Merge

Note that on the surface, the output structures look the same under all four orderings: There is an XP in SpecHP, moved there to discharge a Merge triggering feature on H. But movement only feeds Merge if it applies before Agree. If all movement steps apply early or late, movement uniformly triggers or doesn't trigger Agree; if Agree applies in between the two types of Merge, only one type of movement feeds Agree, giving rise to opacity: In principle, Agree between H and XP is possible, but its does not surface in all HPs that XP is moved through.

To summarize, the logically possible reorderings of a probe feature and the two types of internal Merge triggers (edge feature vs. final feature) result exactly in the four patterns of reflexes of movement we find cross-linguistically.[77]

Before closing this paragraph, I will compare the present approach with the proposal by Reintges et al. (2006), the only explicit Minimalist analysis of the variation between patterns I and II. Reintges et al. (2006) also assume that a reflex of movement is the spell-out of features acquired by a head via Agree with a phrase that undergoes movement. However, they make use of downward Agree. The effects of what I refer to as upward Agree are reanalysed as downward Agree between H and an XP plus subsequent EPP-driven movement of XP to SpecHP (cf. footnote 66). Pattern I arises if the head H (present in the final and in non-final clauses) agrees with the (copy of) the moved XP in its c-command domain. In pattern II languages, it is also the case that the H agrees with XP in final and non-final clauses; but in contrast to pattern I languages, the features copied onto H are not morphologically realized on every head H. They are only spelled out on the head that bears an operator feature. And this is the head in whose specifier the operator surfaces, viz. where it has its final landing site. Assume we are looking at a case of wh-movement. The wh-phrase XP bears an operator-feature which Reintges et al. call [OP]. The head H that triggers the final movement step bears a corresponding uninterpretable feature [uOP] plus a feature [UX]. All other heads H in non-final clauses only bear [UX]. The morphological reflex of movement is sensitive to successful [uOP]-checking, hence,

---

[76]Usually, each head will have only one of the two types of internal Merge triggers under discussion: an edge feature if it projects an intermediate landing site, or a structure-building feature that triggers the final movement step of an XP. The respective Merge trigger is discharged either before or after Agree, giving rise to a feeding or a counter-feeding interaction of the two operations; the ordering statement contains all features that could ever need to be ordered.

[77]There are more logically possible orderings: If [•F•] and [•EF•] apply both before or after Agree, the former could precede the latter or vice versa. But all that matters for the occurrence or absence of the morphological reflex (as a result of Agree) is the order of structure-building features relative to the probe feature. The order of [•F•] and [•EF•] is irrelevant when they both apply before or after Agree. The two possible suborders will not lead to different results with respect to Agree.

it only occurs on the highest head H. This analysis is similar in spirit to mine: A special type of feature is checked in the final position; this feature is not checked in intermediate positions. In contrast to Reintges et al. (2006), however, I do not assume that there is successful Agree in non-final clauses in pattern II languages (that just happens not to be spelled out). Reintges et al. do not include pattern III in their analysis, but it seems to me that it is not straightforward how to derive it in their analysis. According to the logic of their proposal, one would probably want to say that only [UX]-checking is indicated in pattern III languages, excluding [UOP]-checking in final position. However, the features checked in the final position are a superset of those checked in intermediate positions; [UX] is also checked in final position. Consequently, one cannot exclusively refer to intermediate positions with [UX]-checking; this will automatically include the final position and hence, pattern III is not derived. Since the features checked in final and intermediate positions are not in a superset relation in my analysis, this problem does not arise. What would be required to save Reintges et al.'s (2006) analysis would be something like the following: Realize a checked [UX]-feature on H by the exponent M unless there is a checked [UOP]-feature on the same head. But it is not trivial to implement this in a realizational model of morphology: Exponents are specified for the context they can be inserted into, but not for what context they can*not* realize. A possible solution would be to say that a checked [UX]-feature is deleted in the context of a checked [UOP]-feature prior to vocabulary insertion, making vocabulary insertion on the head projecting the final landing site impossible. Reintges et al. (2006) do not address partial movement, optionality of reflexes (pattern I or pattern II as e.g. in Ewe) or the co-occurrence of multiple reflexes (as in Chamorro), and I can thus not evaluate their proposal with respect to these issues.

2.5.3.2 Optionality through partial ordering

As pointed out in the data section 2.3, the reflex of movement is optional in many languages, i.e these languages show free alternation between different patterns. Let me repeat some examples: In Spanish, inversion is obligatory, but the corresponding reflex Stylistic inversion in French is optional. It either applies to all verbs along the path of Ā-movement or to none of the verbs. Hence, pattern I alternates with pattern IV. A different kind of alternation is found in Ewe and Wolof: Pronoun choice in Ewe and complementizer agreement in Wolof *u*-chains exhibits an alternation between patterns I and II: The reflex of movement is obligatory in the clause where the Ā-moved XP lands, but it is optional in non-final clauses. If it is present in the latter, the reflex occurs in all clauses along the movement path; if it is absent, it only shows up in the topmost clause of the dependency. In Wolof *an*-chains, pattern III and IV alternate freely: Either the agreeing complementizer occurs in all non-final clauses crossed by movement (but not in the final clause), or there is no agreeing complementizer at all along the path of movement.

In his work on reflexes of successive-cyclic movement, Boeckx (2008*b*: 35, fn.2) notes that optionality of the reflex is a property of many languages, but he does not show how optionality can be formally implemented. In general, it is simply mentioned as a fact in the literature (see e.g. Torrence 2012 on Wolof who presents an analysis for pattern I but not for the other patterns). At most, it is said that the operations involved are optional. Collins (1993), for example, proposes for Ewe that the copying rule that is responsible for bringing the features involved in Agree from the antecedent in final position to the traces in intermediate positions is optional.

Moreover, modeling optionality in Minimalism is not trivial because operations cannot be optional: Once an operation-inducing feature is present on a head, it needs to get checked (at least it needs to attempt to do so). What can be optional is the presence of the feature in the first place. If it is absent, the operation cannot apply. Against this background, we could derive optionality in two ways: Either the movement triggering features on a head is optional – where there is no movement to SpecHP there can be no Agree with SpecHP – or the probe feature is optional – although there is movement to SpecHP, there is no probe on H that could Agree with SpecHP. Let us see how far we get with these possibilities. The first is clearly not an option. In Indonesian, *men*-deletion can be absent in all positions, including the final position. If this was due to the absence of the Merge-triggering features, the final feature would have to absent as well, which would result in wh-in-situ. But crucially, *men*-deletion can be absent under full overt wh-movement. And if the wh-phrase is ex-situ there must clearly be a final and intermediate movement triggers. Thus, the absence of the reflex cannot simply be due to the absence of movement triggers. The second option is to have optional probe features on certain heads instead of optional structure-building features. Agree with SpecHP cannot apply if the probe feature is absent. For pattern I vs. pattern II we would have to say that the probe on a head that bears an edge feature is only optionally present. While feasible, this is only a restatement of the facts. To conclude, optionality of features is not a very insightful way to handle alternations between the patterns of reflexes of movement.

Fortunately, the ordering approach to reflexes of movement offers a simple and elegant way to derive optionality without postulating optional features: Optionality is accounted for if the orderings in the ordering statements are not total but *partial*. Thus, I propose that optionality is a consequence of underspecified ordering statements. When the features on a selected head are ordered in the numeration, the underspecified ordering statement needs to be turned into a total ordering. This is necessary to determine the location of the unordered features on the stack. If two features $[F_1]$ and $[F_2]$ are unordered relative to each other in a partial ordering statement, the total ordering derived from this can either order $[F_1]$ before $[F_2]$ or $[F_2]$ before $[F_1]$, both options are compatible with the partial order. The fact that two total orders can be created gives rise to optionality. Let me illustrate the proposal for the alternation between patterns I and II in Ewe and in Wolof *u*-chains. It follows if only the order between the probe feature and the feature triggering the final movement steps is fixed; the edge feature is not ordered relative to these features, cf. (157). Note that I use abstract features, the actual nature of the probe and structure-building features does not matter for the purpose of this discussion; the feature to the right of the vertical line '|' is not ordered relative to the rest of the features.

(157)   *Partial ordering for pattern I / II alternation*:
         [•F•] ≻ [∗F:□∗] | [•EF•]

Due to the fixed order of [•F•] and the probe feature, final movement steps will always feed Agree. The ordering of the edge feature relative to the two other features is variable. The partial order in (157) can be resolved into the following two total orders:[78]

---

[78]There are more possibilities but for the occurrence of the reflex of movement it only matters whether the edge feature is ordered before or after Agree; whether it applies before or after the final feature when both Merge triggers are ordered on the same side of Agree is irrelevant for the reflex.

(158) *Total ordering: pattern I / II alternation:*

    a.   [•F•] ≻ [∗F:□∗] ≻ [•EF•]   *intermediate steps apply after Agree*

    b.   [•F•], [•EF•] ≻ [∗F:□∗]   *intermediate steps apply before Agree*

Under the first order, the edge feature is ordered after the probe feature in the numeration. The result will be pattern II where only final movement steps feed Agree; intermediate steps come too late and counter-feed Agree. Under the second option, the edge feature is ordered before the probe feature. As a consequence, both types of movement steps feed Agree, giving rise to pattern I. Under the partial ordering approach to optionality, we do not have to say that operations or the operation-inducing features are optional, which is problematic anyway (see the discussion above). The operations always apply; what is optional is *the point of the derivation at which they apply*. In the case of Merge, the timing of this operation has consequences for the interaction with Agree (feeding or counter-feeding) and therefore, we get different patterns of reflexes.

    Before I continue with the analysis of optionality between other patterns, let me come to an interesting kind of variation among languages that show the pattern I / pattern II alternation. Subject pronoun choice in Ewe behaves just as expected and predicted under the partial order in (157): This reflex can either appear in all clauses along the path of movement or only in the topmost clause of the dependency. In Wolof, however, patterns I and II are only a subset of the actual possibilities. Look at the examples of *u*-chains in (159) where a silent wh-word is extracted from the most deeply embedded clause across two clause-boundaries. (159-a) shows the classic pattern I where all complementizers along the movement path show class agreement with the wh-phrase; (159-b) illustrates the classic pattern II configuration with default agreement (class marker *l-*) on intermediate and full agreement on the topmost complementizer; we have already seen these examples in section 2.3.1.3. However, there are two more grammatical sentences: In (159-c) and (159-d) only one of the complementizers in the embedded clauses agrees in class, the other one shows default agreement; it does not matter in which order the agreeing and the non-agreeing intermediate complementizers appear. The only thing that is not variable is the form of the complementizer in whose specifier the wh-phrase lands: In *u*-chains, it must show class agreement (independently of the distribution of agreeing and non-agreeing complementizers in the embedded clauses), cf. (159-e). In a nutshell, only the topmost complementizer has to Agree in class with the wh-phrase in its specifier, the intermediate complementizers on the movement path can occur in any possible combination of agreeing and default forms (Torrence 2012: 1174).

(159) *Patterns of complementizer agreement in Wolof u-chains (Torrence 2012: 1173)*:

    a. [_CP_ Ø_k_ **k**-u Kumba wax [_CP_ ne **k**-u Isaa defe [_CP_ ne **k**-u Maryam dóór
          Q CL-u Kumba say   FRC CL-u Isaa think   FRC CL-u Maryam hit
        ___*k*]]]

    b. [_CP_ Ø_k_ **k**-u Kumba wax [_CP_ ne **l**-a   Isaa defe [_CP_ ne **l**-a   Maryam
          Q CL-u Kumba say   FRC EXPL-a Isaa think   FRC EXPL-a Maryam
        dóór ___*k*]]]
        hit

    c. [_CP_ Ø_k_ **k**-u Kumba wax [_CP_ ne **l**-a   Isaa defe [_CP_ ne **k**-u Maryam
          Q CL-u Kumba say   FRC EXPL-a Isaa think   FRC CL-u Maryam
        dóór ___*k*]]]
        hit

    d.   [$_{CP}$ Ø$_k$ **k**-u Kumba wax [$_{CP}$ ne **k**-u Isaa defe [$_{CP}$ ne **l**-a   Maryam
            Q   CL-u Kumba say   FRC CL-u Isaa think   FRC EXPL-a Maryam
            dóór ___$_k$]]]
            hit

    e.   *[$_{CP}$ Ø$_k$ **l**-a   Kumba wax [$_{CP}$ ne **k**-u Isaa defe [$_{CP}$ ne **l**-a   Maryam
            Q   EXPL-a Kumba say   FRC CL-u Isaa think   FRC EXPL-a Maryam
            dóór ___$_k$]]]
            hit
            'Who did Kumba say that Isaa thought that Maryam hit?'

How can we account for this variation? Are there actually more patterns than the ones identified so far? I suppose that the answer is no. The core asymmetry is still between the final and intermediate C-heads: The former must Agree, the latter can Agree. That the final complementizer must Agree is derived by the partial order in (157) that also underlies optionality in Ewe. I propose that the only difference between Ewe and Wolof is the mode in which the partial ordering is turned into a total ordering: In Ewe, the partial order in (157) is converted into one of the two total orders in (158) just *once*, before the derivation starts. The chosen order is then *fix for the whole derivation*, i.e. it holds for any head in the numeration on which final and intermediate movement steps need to be ordered. In this way, we can get either the classic pattern I (every edge features is ordered before the probe feature, feeding Agree) or the classic pattern II (every edge features is ordered after the probe feature, counter-feeding Agree). In Wolof, however, the choice between the total orders in (158) is not made for the whole derivation; rather, it is made *anew every time the selected head H is assigned an edge feature* that needs to be ordered relative to other operation-inducing features on H. For the lowest complementizer, the system might choose the total order where the edge feature is ordered before the probe feature, resulting in class Agree; the next higher intermediate C head might follow the order in which the edge feature is ordered after the probe feature, resulting in default Agree, etc. The topmost C head will always show full class Agree due to the fix order [•F•] ≻ [∗F:□∗]. In the example just sketched, we would get the pattern in (159-c). Thus, in contrast to Ewe, a total order in Wolof is computed *more than once*.

    Let me now come to the alternation of patterns III and IV, as documented in Wolof *an*-chains. This optionality follows from the following underspecified ordering statement:

(160)   *Partial ordering for pattern III / IV alternation*:
        [∗F:□∗] ≻ [•F•] | [•EF•]

This partial ordering can be resolved into the total orderings in (161); again, the edge feature is unordered to the other features. In contrast to the pattern I / II alternation, the final movement step is ordered after Agree:

(161)   *Total ordering: pattern III / IV alternation*:
        a.   [∗F:□∗] ≻ [•F•], [•EF•]   *intermediate steps apply after Agree*
        b.   [•EF•] ≻ [∗F:□∗] ≻ [•F•]   *intermediate steps apply before Agree*

If the first order is established, we get pattern IV because both types of movement come too late to feed upward Agree. Under the second order, pattern III emerges because intermediate movement steps apply early enough to feed Agree, whereas final ones still come too late to do so. Let me illustrate the patterns in Wolof *an*-chains, repeated from section 2.3.3.1. Recall that the central difference between *u*- and *an*-chains is that in the latter (i)

the wh-phrase is overt, and (ii) that the complementizer in whose specifier the wh-phrase lands obligatorily shows default class agreement (Agree is counter-fed by the final movement step). In (162-a) we see pattern III: All intermediate complementizers show class agreement with the wh-phrase. In (162-b), all intermediate complementizers exhibit default class agreement, resulting in pattern IV. Interestingly, that is not the entire variation: As in *u*-chains, intermediate C-heads can actually show default or full class agreement with the wh-phrase in any possible combination, i.e. one intermediate C head exhibits default agreement whereas the other shows full agreement etc, cf. (162-c) and (162-d). (162-e) illustrates that the C-head immediately following the overt wh-phrase can never show class agreement.

(162) *Patterns of complementizer agreement in Wolof an-chains (Torrence 2012: 1173f.):*

    a.   [$_{CP}$ K-an$_k$ **l**-a-ñu   wax [$_{CP}$ **k**-u jigéén j-i   foog [$_{CP}$ **k**-u ma
          CL-an EXPL-a-3PL say   CL-u woman CL-DEF.PROX think   CL-u 1SG
          dóór ___$_k$]]]
          hit

    b.   [$_{CP}$ K-an$_k$ **l**-a-ñu   wax [$_{CP}$ **l**-a   jigéén j-i   foog [$_{CP}$ **l**-a
          CL-an EXPL-a-3PL say   EXPL-a woman CL-DEF.PROX think   EXPL-a
          ma dóór ___$_k$]]]
          1SG hit

    c.   [$_{CP}$ K-an$_k$ **l**-a-ñu   wax [$_{CP}$ **l**-a   jigéén j-i   foog [$_{CP}$ **k**-u
          CL-an EXPL-a-3PL say   EXPL-a woman CL-DEF.PROX think   CL-u
          ma dóór ___$_k$]]]
          1SG hit

    d.   [$_{CP}$ K-an$_k$ **l**-a-ñu   wax [$_{CP}$ **k**-u jigéén j-i   foog [$_{CP}$ **l**-a
          CL-an EXPL-a-3PL say   CL-u woman CL-DEF.PROX think   EXPL-a
          ma dóór ___$_k$]]]
          1SG hit

    e.   *[$_{CP}$ K-an$_k$ **k**-u-ñu   wax [$_{CP}$ **k**-u jigéén j-i   foog [$_{CP}$ **k**-u ma
          CL-an CL-u-3PL say   CL-u woman CL-DEF.PROX think   CL-u 1SG
          dóór ___$_k$]]]
          hit
          'Who did they say that the woman thinks that I hit?'

As before, this follows if the total orders in (161) are fixed anew every time an edge feature on a head needs to be ordered with respect to a probe feature.[79]

---

[79]Note that there is no single ordering statement for Wolof that holds across both *u*- and *an*-chains. The reason is that the fixed orders of [•F•] and [∗F:□∗] in the partial orderings in (157) and (160) are the opposite of one another. This alternation cannot simply be derived by underspecification of the order of these features because agreement between the moved XP and the head that projects its final landing site is not free; rather, it depends on the construction: In *u*-chains, the final movement step must feed Agree and in *an*-chains, it must not. Thus, we do need two (partial) ordering statements for Wolof. One way to implement this is the following: There are two different C heads in Wolof, one used in *u*-chains (selecting a zero operator) and one used in *an*-chains (selecting an overt operator). There is only one general ordering statement in the grammar that determines the order of operation-inducing features in the numeration, let's say the one in (157). Therefore, it also applies to the C head used in *u*-chains. The C-head used in *an*-chains, however, has a lexically specified order of [•F•] and [∗F:□∗], namely the one in (160). Since it is lexical, it overwrites the general ordering statement that would require the opposite order (an *Elsewhere Condition* effect). There is independent evidence that orderings can be lexically determined: In the Bantu languages Kinande and Lubukusu, long-distance movement of the subject from an embedded clause bleeds

Finally, let me consider the alternation between pattern I and IV. In these patterns, final and intermediate movement steps pattern alike in that they both apply either before or after Agree. Hence, they are unordered relative to the probe feature:

(163)   *Partial ordering for pattern I / IV alternation*:
         { [•F•], [•EF•] } | [∗F∗]

The total orderings that result from (163) are given in (164). Note that in order for the two structure-building features not to get split up by the probe feature, they must form a unit.

(164)   *Total ordering: pattern I / IV alternation*:
         a.   [•F•], [•EF•] ≻ [∗F:□∗]   *final steps apply before Agree*
         b.   [∗F:□∗] ≻ [•F•], [•EF•]   *final steps apply after Agree*

To summarize, optionality of patterns arises through partial ordering: It's not the operations that are optional, it is the timing of their application.

2.5.3.3   Multiple reflexes of movement

Some of the languages we have looked at exhibit multiple reflexes of movement: A single instance of Ā-movement can have several reflexes which follow different patterns. In Chamorro and Indonesian, a pattern I and a pattern II reflex co-occur under Ā-movement. Such multiple reflexes are a challenge for the theories presented in section 2.4.5 which are based on enriched representations: In the quasi-operators approach, traces are binders in pattern I languages but not in pattern II languages; under Haïk's (1990) approach, features are deleted on traces in pattern II languages, but not in pattern I languages. In languages with multiple reflexes of a single movement operation, both conditions would have to hold at the same time, which is paradoxical.

The ordering approach provides a simple solution for multiple reflexes: As long as the two reflexes are the result of two different Agree relations (different features are involved), the triggering probe feature can be ordered differently relative to the Merge triggering features. Look at the abstract ordering statement in (165):

(165)   *A pattern I and a pattern II reflex co-occur*:
         [•F•] ≻ [∗F:□∗] ≻ [•EF•] ≻ [∗L:□∗]

Given this order and upward Agree, the reflex that results form Agree in feature [F] will have a pattern II distribution: Early final movement steps feed it, but intermediate movement steps come too late to do so. In contrast, the reflex that is the result of Agree in feature [L] will have a pattern I distribution: Both movement types apply before Agree in [L] and thus, both feed Agree. This is the order needed for Chamorro and Indonesian. Indeed, the two reflexes in Chamorro involve different features: Wh-agreement on the verb (pattern I) tracks the grammatical function of the moved element (or of the clause it is extracted from, respectively). Following the literature, I take this to be the result of case

---

subject-verb-agreement in the embedded clause (for such bleeding interactions, which involve downward Agree, see section 3.2.1). But this only happens when a particular complementizer is used. With a different complementizer, there is no bleeding under movement. Since the choice of the reflex pattern (bleeding throughout or no bleeding at all) depends on the complementizer, the order must depend on particular lexical items in these languages. See Diercks (2010: 188ff.) on Lubukusu and Schneider-Zioga (2000) on Kinande. I thank Jason Zentz (p.c.) for pointing out to me the lexically-driven alternation of patterns and the relevant data.

agreement. Complementizer agreement in Chamorro (pattern II) involves the category feature of XP. The ordering statement in Chamoroo would thus be the following:

(166)   *Ordering statement in Chamorro (patterns I and II co-occur)*:
          [•F•] ≻ [∗CAT::□∗] ≻ [•EF•] ≻ [∗CASE::□∗]

Of course, there is no single head that bears all of the features in (166). In Chamorro, it is the C head which bears the category probe and one of the two internal Merge triggers (dependeing on whether it projects a final or an intermediate landing sites); a head in the extended projection of the verb (v or T) bears the case probe and an edge feature (Ā-movement will never be final in SpecvP or SpecTP). Due to the order in (166), we get a pattern I reflex in case on v / T, and a pattern II reflex in category on C. As for Indonesian, the two reflexes on v and C, respectively, are morpholgically invariant, so we cannot tell which features are involved in the Agree relations and hence, whether the reflexes involve the same or different features. But at least, the features could be different (as required in the analysis); the morphology simply does not encode it.

The logic of the derivation of multiple reflexes is exactly the same as for the opaque patterns II and III: The interaction of Merge and Agree is opaque because the two types of Merge operations are interleaved with an Agree relation. In (166), the same holds for Agree: Agree results in two different reflexes because there are indeed two distinct Agree relations that apply at different points relative to a Merge operation (here, the final movement step triggered by [•F•]).

The ordering approach to multiple reflexes actually predicts two mixed patterns: In addition to the the co-occurrence of patterns I and II, we can also get patterns I and III simultaneously. The orders that produce these multiple reflexes are given in (167).[80]

(167)   *Orders producing multiple reflexes*:
          a.   [•F•] ≻ [∗F:□∗] ≻ [•EF•] ≻ [∗L:□∗]: PI/PII
          b.   [•EF•] ≻ [∗F:□∗] ≻ [•F•] ≻ [∗L:□∗]: PI/PIII

In Kitharaka we find evidence for a mixture of PI and PIII: We have already seen in section 2.3.3.2 that Ā-movement triggers the occurrence of a preverbal focus marker in all clauses except for the clause where the wh-phrase lands (PIII). In addition, Kitharaka has a pattern I reflex: The same movement operations that leads to preverbal focus marking (wh-movement, relativization, focus movement) also trigger allomorphy. The present tense marker *ri* is obligatorily replaced by the form *ku* in all clauses that are crossed by movement, including the one where the wh-phrase lands (PI). (168-a) illustrates the default tense marker that occurs if there is no overt movement (here, with wh-in-situ), and (168-b) shows the special form *ku* under overt movement:

(168)   *Tense marker allomorphy in Kitharaka (Muriungi 2005: 52)*:
          a.   [CP U-**ri**-thugania [CP ati   John a-**ri**-ring-a uu   ]]
                    2SG-PRES-think   that John SM-beat-FV who
               'Who do you think that John is beating?'

---

[80]There are actually 24 possible permutations of the four operation-inducing features. However, under most orders the two probe features apply at the same point relative to the Merge operations and hence, both Agree relations result in the same pattern. In addition, there are a number of orders where the Agree relations do apply at different points relative to Merge, but where one Agree operation applies before all Merge operations. As a consequence, we get pattern IV plus another pattern. But since pattern IV means that there is no reflex at all, we will not be able to see that two patterns are mixed. The only orders that produce a visible mixture of patterns are those in (167).

    b.   [$_{CP}$ I-mbi$_k$ u-**ku**-thugania [$_{CP}$ ati   John **n**-a-**ku**-ring-a   ___$_k$]]
           FOC-what 2SG-PRES-think   that John FOC-SM-PRES-beat-FV
           'What do you think that John is beating?'

What is excluded under the present approach is that PII and PIII co-occur. PII requires the order final movement > Agree > intermediate movement, whereas PIII requires intermediate movement > Agree > final movement. The order of the final and the intermediate movement trigger relative to each other is the reverse in PII and PIII; [•F•] > [•EF•] and [•EF•] > [•F•] cannot hold at the same time. If PII and PIII co-occur in a language, only one order can result from the general ordering statement, the other must be the consequence of a lexically specified order on a particular lexical item. We have encountered such a case in Wolof where the complementizer can show PI (in *u*-chains) or PIII (in *an*-chains), see footnote 79 for discussion.

2.5.3.4 Advantages of the ordering approach

In this section I compare the ordering approach to opacity with approaches that are based on enriched representations. Recall that previous approaches either made reference to levels of representation and / or used enriched representations plus constraints on empty elements to derive opacity. In Chung (1998), for example, pattern II is derived by applying agreement at S-structure, and pattern I requires reference to traces in intermediate landing sites. In the quasi-operators approach, different types of empty elements need to be postulated: some are binders (pattern I) and some are not (pattern II). Approaches that apply operations and constraints to empty elements in intermediate positions require deletion or copying of features before agreement can access the empty elements. None of these assumptions is necessary in the present approach. It neither requires empty elements like traces in intermediate positions, nor operations applying to them, nor is there any reference to levels of representations. This is a good result because there is no designated S-structure in Minimalism which we could refer to, and trace theory has been abandoned. The only ingredients the present approach needs are (i) a distinction between (the triggers of) final and intermediate movement steps and (ii) ordering of operation-inducing features. Assumption (i) is indispensable in every approach that tries to account for the cross-linguistic variation. The existence of patterns II and III provides *empirical* evidence for this split of movement types. Previous approaches require a number of assumptions *in addition* to this one. Assumption (ii) has been argued for independently in Minimalism on the basis of transparent interactions of Merge and Agree triggered by a single head (cf. section 1.2.3); these cannot be derived by simultaneous application. What I add is simply that not only Agree and Merge are ordered relative to each other, but rather that a more fine-grained approach is needed that distinguishes between different types of Merge.[81]

---

[81] See Weisser (2014) for an analysis of extraction from clause-chains where it is also crucial that final and intermediate movement steps apply at different points when triggered by the same head (although the operations do not interact with Agree): intermediate steps apply before final steps. There is one analysis in the literature that also distinguishes between final and intermediate movement steps based on their different behaviour with respect to an Agree relation triggered by the same head: Haddad and Wurmbrand (2013) analyse full vs. partial agreement in Standard Arabic control and raising constructions. Usually, preverbal subjects trigger full agreement in ϕ-features with the verb, whereas postverbal subjects trigger only partial agreement in gender. This is derived as follows: T has the EPP property that can be satisfied either by movement of the subject or the v+V-complex to SpecTP. By assumption, v has only a gender feature acquired through Agree with the subject DP base-merged in SpecvP; v cannot agree in number with the subject DP.

Furthermore, some of the previous approaches had problems in accounting for variation: Pattern III has barely been discussed and it would be difficult to integrate it into existing approaches, especially into the quasi-operators approach. In the present approach, variation is accounted for by reordering of operation-inducing features. Languages do not have to differ in their inventories of empty elements as in the quasi-operators approach, or in whether operations apply to empty elements or not.

Moreover, as discussed at length in section 2.4.5, it is not trivial in Minimalism to make a distinction between occurrences of a moved XP in final and intermediate positions, which is required in all approaches. In GB, there was an inherent distinction between the two types of positions: Intermediate positions are occupied by traces and the final landing site by the antecedent. However, under the copy theory of movement and under the multi-dominance approach, there is no such inherent distinction; the occurrences are identical – though, there may be ways to reintroduce the distinction. Things get even worse if one assumes that movement does not leave behind anything in intermediate positions. None of the previous approaches can even be formulated given this assumption. Crucially, the present approach is fully compatible with all three Minimalist concepts of movement because it is only important *when* a phrase is moved, not *how* movement is implemented. Look at the feeding interaction in (153) again. To get feeding,

---

If the subject moves to SpecTP to check T's EPP, it feeds ϕ-agreement with T (in a Spec-head relation). If, however, the v+V-complex moves to T to check the EPP, it can only trigger agreement in gender, having no number feature. The interesting constructions are control structures. These exhibit opacity: The subject surfaces postverbally in the embedded clause, with partial agreement on the embedded verb, as expected; the matrix verb, however, shows full agreement with the embedded subject although the subject does not precede the matrix verb. Haddad and Wurmbrand (2013) provide evidence that the subject does move to the matrix clause via the embedded SpecTP position; in matrix SpecTP it triggers full agreement with the T-head. However, the moved subject DP is phonetically realized only in the embedded clause. But why is there no full agreement on the embedded verb if the subject DP moves through SpecTP? The answer Haddad and Wurmbrand give is that subject movement to SpecTP comes *too late* to trigger full agreement: First, v+V moves to the embedded SpecTP to check the EPP and v agrees with the complex in SpecTP, giving rise to partial agreement. The subject DP moves to an outer specifier of the embedded TP *afterwards*. This movement (which they call"edge movement" because it uses SpecTP as an intermediate landing site on its way to the matrix clause) differs from pure EPP-movement in that edge movement applies late and thus counter-feeds full Agree with v. Haddad and Wurmbrand (2013) do not use the counter-feeding terminology and they do not make this assumption explicit, but what is required to make the analysis work is ordering of operation-inducing features on the embedded T along the following lines: EPP-driven and edge feature-driven movement apply at different points of the derivation relative to Agree (the EPP is represented as [•D•]):

(i)   T { [•D•] ≻ [∗ϕ:□∗] ≻ [•EF•] }

EPP-driven movement applies before T triggers Agree and thus movement feeds Agree; edge movement applies after Agree and thus counter-feeds it. In current terminology, EPP-driven movement is a final movement step and edge-feature-driven "edge movement" is an intermediate movement step. Thus, their analysis also presupposes that there is a split between these two types of Merge-triggers. Note that Haddad and Wurmbrand's (2013) analysis could be reinterpreted as follows: Both EPP-driven movement and edge movement apply before Agree, with EPP-movement applying before edge movement. Given the *Extension Condition*, the EPP-moved v+V-complex ends up in the inner specifier of TP and the edge-moved DP ends up in the outer specifier of TP. If there is a minimality requirement on Agree, as in the definition in (137), T will Agree with the element in its inner specifier. The result would be the same as under the order in (i), i.e. partial agreement. I do not know which of the two analyses the authors have in mind, this is not made explicit. So, the data presented by Haddad and Wurmbrand provide evidence that two types of Merge need to be distinguished, though this conclusion is not necessarily based on the configuration I look at, viz. that we need the distinction because the two types of Merge are interleaved with Agree: Merge$_1$ ≻ Agree ≻ Merge$_2$. In their analysis, both types of Merge could apply before Agree as well, which would , however, not be sufficient to derive the data I investigate.

it is only important that at the point of the derivation where H triggers Agree, there is an occurrence of XP in SpecHP. What this occurrence looks like is completely irrelevant, i.e. whether it is a copy of XP, or a chain-link to a single occurrence of XP (multi-dominance), or whether it is the only occurrence of XP that moves on later without leaving behind anything. Especially under the latter approach, we get opacity: If the single occurrence of XP moves on to a higher position, it is not clear on the surface how H could Agree with the XP given that XP is not in the right structural position to trigger Agree with H and that there is not even an occurrence of XP left in SpecHP that encodes that XP has been there at a certain point. The present approach has no problem to derive the patterns even under this radical hypothesis about movement because the approach is strictly derivational.

From a broader perspective, the difference between the present and previous approaches comes from a shift in the responsibility for the success of Agree. In previous approaches, the success heavily depends on the properties attributed to occurrences in final and intermediate positions (whether they are binders or not, whether operations apply to them or not, etc.). Therefore, the nature of the occurrence (copy, multiply dominated XP) matters when the analysis is transferred to Minimalism. This problem vanishes if we take the burden of responsibility from the occurrences and put it onto the head that triggers the movement and agreement operations, as in the present approach.[82] There are no more manipulations of the occurrences, only the timing of operations on the triggering head is varied.

A further advantage of the ordering approach to opacity is that it can easily handle multiple patterns of reflexes as well as optionality between patterns. As for the first, we have seen in section 2.3 that in several languages different patterns co-occur under a single instance of Ā-movement. Such mixed reflexes are a challenge for approaches that apply constraints / operations on empty elements in intermediate positions: For example, if a language has patterns I and II and we use the quasi-operators approach, traces left by one and the same movement operation must be binders and must not be binders at the same time; similarly, features on traces must be deleted and must not be deleted in Haïk's (1990) approach, which is paradoxical. Under the present approach, different reflexes arise because distinct Agree relations are involved, and these are ordered differently relative to the Merge triggering features; consequently, a single movement operation can give rise to multiple reflex patterns. As for optionality, Ā-movement can trigger different patterns under a single instance of Ā-movement. However, these patterns do not co-occur, but alternate freely: Either one or the other pattern arises. I have argued that it is not satisfactory to simply say that the Move or Agree operations involved in the generation of a reflex are optional in such languages. In the present approach, optionality

---

[82]In this respect, the present analysis is similar in spirit to Nevins' (2007) analysis of PCC effects. The PCC is a constraint on possible combinations of weak arguments depending on their person features. Previous approaches impose restrictions on the properties of arguments to derive (different versions of) the PCC. For example, they assume that 3rd person is not a person, viz. it is the absence of a person feature. Furthermore, it is stipulated in some approaches that the feature content of the arguments involved differ within the same language: 3rd person is a feature on indirect objects, but it is the absence of person on direct objects. Nevins (2007) shows that the first assumption is empirically not tenable; the second is undesirable from a conceptual point of view. He develops an alternative analysis where the PCC effect is not controlled by the properties of the arguments, which vary across languages, but rather by the equipment and properties of probe features on the head that triggers agreement with these arguments. The combination of parameters along which the feature on a head can vary derives the attested versions of the PCC. Under this analysis, the representation of person on arguments can be held constant across languages with different versions of the PCC, which was not possible in previous approaches. As in the present approach, it is the features on functional heads that differ and produce variation; languages do not differ in the nature of the XPs that take part in Agree.

arises through partial ordering. It is not the operations which are optional, they always apply; rather, their timing of application is not fixed. Hence, the ordering approach provides a simple and elegant solution for the initially challenging fact that several patterns can co-occur or alternate within the same language.

Finally, let me add that the present approach provides the first uniform analysis of all three patterns of reflexes of movement (or four, if we include pattern IV).

2.5.4 Against a morphological account of variation

At this point I would like to discuss a potential alternative account of patterns I-IV that is not syntactic but rather morphological in nature and does not involve reordering of operation-inducing features. It has been proposed to me by Gereon Müller and Coppe van Urk (p.c.). They suggest to consider the following analytical option: Patterns II and III are actually instances of pattern I, i.e. there is always successful Agree (resulting in valuation) between H and an XP moved to SpecHP in the syntax, whether SpecHP is a final or an intermediate landing site; thus, movement to SpecHP always feeds Agree and therefore the order of operations on H does not vary, it is always Merge before Agree. The difference between the patterns arises in the postsyntactic morphological component as a consequence of the specification of VIs. Suppose VIs that realize features on H (which H acquired under Agree with XP) can be sensitive to type of feature XP has discharged on H when it moved to SpecHP, viz. an edge feature or a final feature (presupposing that features discharged in syntax are still visible in the morphological component, at least as context features for vocabulary insertion). Hence, some VIs that realize features of H can only be inserted if H bears a discharged final feature [•f•] or a discharged edge feature [•Edge•]. For example, pattern II would arise if there is an overt VI that realizes features of H in the context of [•f•] and a zero VI that realizes the same features in the context of [•Edge•] (or the second VI does not bear any context features at all, it could be the default realization of the respective features on H). As for pattern III, there would be an overt VI inserted into H in the context of [•Edge•] and a zero VI for H in the context of [•f•]. Pattern I arises if the VI is underspecified for the feature that triggers Merge on H, i.e. it can be inserted in every head H, regardless of whether H projects a final or an intermediate landing site. Pattern IV arises (i) either in the same way if there is a default marker (as in Wolof), or (ii) if there is (a) either no convenient VI in the first place (the features copied onto H under Agree with XP are not realized by a VI at all), or (b) a zero VI underspecified for the movement trigger on H. Under option (ii), we don't see any exponent on the surface.

Note first that this approach crucially relies on the assumption that final and intermediate movement steps have different triggers, just as the present approach; hence, both approaches provide an argument for this distinction. The alternative approach also requires ordering of operation-inducing features in the syntax because in order to have feeding of Agree by movement to SpecHP (which, by assumptions, happens in all languages, regardless of the surface reflex pattern). In this sense, the approaches are very similar in their basic assumptions. What the alternative account does not need is reordering of operation-inducing features on a head in the syntax in order to derive the different patterns.

There are differences, however. From a conceptual point of view, the patterns are a morphological coincidence under the alternative account; they are the result of idiosyncratic marker specifications, in contrast to the ordering approach. In addition, under the alternative approach we would expect languages where the distinct VIs inserted into final and intermediate heads H are overt, i.e. a pattern I reflex where the head H that projects

the final landing site has a different overt exponent than the heads H that project an intermediate landing site. Coppe van Urk (p.c.) informs me that this is actually the case in Dinka for complementizer agreement.[83] This is the only language I know of where this seems to happen; in the vast majority of languages the exponents in final and non-final are the same in pattern I languages. Under the alternative approach we would expect to see this more often, however. A related issue is the following: We would expect the VIs to be sensitive to other features of the head H they realize, not just the type of feature on H that triggers Merge of XP in the syntax. For example, we would expect to find VIs realizing features on a C head (copied onto C under Agree with the XP in SpecCP) to be sensitive to mood, if we assume that mood is encoded on C; or VIs spelling out features on T that are sensitive to tense, etc. But this does not seem to be the case in the respective languages. This might of course be a coincidence, there might be such languages, and the ordering approach would not have any difficulty in accounting for them by contextual allomorphy; but under the morphological account we would expect this to happen much more regularly. Of course, these are rather weak objections; both approaches can easily cope with the data so we cannot really distinguish between them at this point.

There is, however, an empirical argument against the alternative morphological analysis: It fails to account for all attested patterns of optionality. Let me elaborate on this point. We have seen in section 2.5.3.2 that in some languages the reflex alternates between pattern I and pattern II, i.e. the overt exponent obligatorily occurs in the final clause but only optionally in non-final clauses. Let us assume for the sake of illustration that the reflex of movement is morphological and additive in nature; the morpheme *-a* realizes a valued feature [F:V] on the C head; C acquires the value under upward Agree with an XP Ā-moved to SpecCP. Under the morphological account of variation, pattern I is derived if the exponent is underspecified for the type of feature on C that triggers Merge, cf. (169-a). It can thus be inserted into every C head, regardless of whether it projects a final or an intermediate landing site for XP.

(169)   *VIs, morphological account of optionality (option 1)*:

    a.   /a/ $\leftrightarrow$ [F:V] / ___ C

    b.   /Ø/ $\leftrightarrow$ [F:V] / ___ C$_{[\cancel{\text{EF}}]}$

In order to derive pattern II, *-a* has to be blocked in non-final clauses. To achieve this, we need to postulate a zero allomorph that realizes the same features in the context of a discharged edge feature on C, cf. (169-b). The problem is that the zero marker is more specific than the overt exponent in (169-a), so the zero marker should *always* block *-a* in non-final clauses, and pattern I cannot be derived. To postulate an additional marker for final C heads (syncretic with the default marker in (169-a)), does not help either; this marker and the zero marker are more specific than the VI in (169-a) and would block it, too, making it impossible to derive pattern I under the morphological approach to patterns of reflexes of movement. The only technical way out would be to restrict the marker *-a* to final C heads and to have two equally specific VIs for non-final C heads, one of them

---

a zero VI and the other one syncretic to the exponent for final Cs, cf. (170):

(170)　　*VIs, morphological account of optionality (option 2)*

　　　　a.　/a_1/ ↔ [F:V] / ___ C~~[◁#◁]~~

　　　　b.　/Ø/ ↔ [F:V] / ___ C~~[◁#◁]~~

　　　　c.　/a_2/ ↔ [F:V] / ___ C~~[◁#◁]~~

The VI in (170-a) will be inserted into C heads projecting a final landing site. Since the VIs in (170-b) and (170-c) are equally specific, one could assume that the choice of an exponent for intermediate C heads is optional. Either the zero VI is chosen or the overt VI /a_2/ and this is what causes optionality in the patterns: In the former case, we get pattern II because the exponent *-a* surfaces solely on the final complementizer; in the latter we get pattern I because *-a* occurs on all C heads along the path of movement. However, under this account, the fact that we find the same exponent *-a* in all clauses (pattern I) is a pure coincidence; it happens so that the exponent for the final and the one for intermediate C heads are syncretic, but this syncretism is accidental. It is not the result of underspecification of a VI. The two VIs in (170-a) and (170-c) could could very well be phonologically non-identical. But we have seen that in the languages with pattern I in section 2.3.1, we always find the same exponent in final and intermediate clauses. That is, pattern I is not derived in a systematic way under the alternative, purely morphological approach to variation – in contrast to the ordering approach where pattern I is never the result of accidental homophony, not even if a language exhibits optionality between patterns I and II. Optionality simply follows from partial ordering of operations in the syntax in a systematic way.

　　　But even if one accepted the necessity of accidental homophony in the morphological account of variation, which makes it technically possible to account for optionality between patterns I and II, this analysis still fails to account for more intricate patterns of optionality as e.g. in Wolof *u*-chains: In this construction, there is not only optionality between patterns I and II, but rather free variation between the overt and the zero exponent in non-final clauses (cf. (159)). In the morphological account, one can only distinguish between VIs that realize features on a head that projects a final landing site and VIs that realize features on a head that projects an intermediate landing site, but more fine-grained distinctions between the intermediate heads are not possible. Either there is an exponent on all intermediate heads or there is no exponent on any intermediate head, no other options can be encoded in the VIs. Under the (partial) ordering approach to variation such intricate patterns are actually expected. I conclude that the syntactic ordering approach is superior to the morphological account of variation in patterns of reflexes of movement. Optionality thus provides a strong argument for the necessity of reordering of operation-inducing features (which is in turn highly relevant for the discussion of intrinsic vs. extrinsic ordering, see section 5.2).

2.6   Analyses of the data

In the present approach, all reflexes of movement are the result of a successful Agree relation between a head H and an XP in its specifier. Agree leads to valuation of a feature on H, viz. it involves *addition* of feature values. What we would thus expect is that all reflexes of movement involve addition: The value on H acquired through Agree is realized postsyntactically by an exponent. However, we have seen that a reflex of movement can also manifest itself through deletion or replacement of exponents; furthermore, there are syntactic reflexes that neither involve addition nor deletion, but rather word order changes. The aim of this section is to show that the Agree-based approach can capture addition as well as non-addition reflexes. I will illustrate the analysis on the basis of a few examples for each class of reflexes. The case studies will also exemplify how asymmetries in the occurrence of the reflex (argument-adjunct, subject-object) can be derived.

2.6.1   Addition

The analysis of reflexes of movement that result in addition of an exponent is straightforward: Movement of an XP to SpecHP feeds Agree between H and XP. The value H acquired through Agree is realized by a vocabulary item in the postsyntactic component. Without movement, H does not receive a value and hence, there is no exponent for this value.

2.6.1.1   *no*-marking in Duala

In Duala, the marker *no* appears after the finite verb (the main verb or an auxiliary) if a constituent has undergone Ā-movement. *no*-marking is a pattern II reflex: It can only occur in the clause that hosts the operator on the surface, but not in lower clauses crossed by movement. Furthermore, there is a subject/non-subject split: *no*-marking is only triggered if an object or an adjunct is extracted, but not if the subject undergoes (local) Ā-movement.

I analyse the addition of *no* as follows: Since *no* is a reflex of operator movement to SpecCP, it is the result of an upward Agree relation between C and an operator moved to SpecCP. The feature involved in Agree is thus an operator feature [OP] that can have two values: [±OP]. All operators bear the positively valued feature, non-operators bear the negatively valued feature.[84] C contains the corresponding probe feature [∗OP:□∗] that seeks for a value under upward Agree. If the XP in SpecCP is an operator, we get *no*-marking on the verb. Therefore, *no* must realize the feature [OP:+OP] ([OP:–OP] could be realized by a zero marker; alternatively, there is no marker that realizes this feature):

(171)   *VI in Duala (first version)*:
         /no/ ↔ [OP:+OP]

In order for C to be able to Agree with the operator, Ā-movement must apply before Agree (feeding). To derive pattern II, the following order of operations is required:

(172)   *Ordering statement in Duala (first version)*:
         [•F•] ≻ [∗OP:□∗] ≻ [•EF•]

---

[84]Different types of operators may be distinguished by additional features such as [WH] for wh-elements, [REL] for relative oeprators, etc.

Hence, only final movement steps feed Agree and can thus trigger *no*-insertion; intermediate movement steps to SpecCP come too late to feed Agree.

I assume that valuation takes place on C; however, *no* does not surface in C. Rather, it attaches to the finite verb in T which is either an auxiliary or the main verb in case there is no auxiliary; the C head is realized separately in embedded clauses. We also know that T does not move to C in clauses with Ā-movement because the subject in SpecTP intervenes between the Ā-moved constituent and the finite verb if a non-subject undergoes movement, cf. the data in section 2.3.2.2. And crucially, obligatory movement of the subject to SpecTP does not trigger *no*-marking, viz. this is clearly a reflex of movement to SpecCP. Hence, the information that an operator has been moved to SpecCP must somehow be handed down from C to T. One way to implement this is to make use of Lowering as proposed in Embick and Noyer (2001). Lowering is a postsyntactic operation that applies before vocabulary insertion and is sensitive to hierarchical structure (viz., it applies before linearization). Lowering takes a head ad adjoins it to the head of its complement, cf. (173):

(173)   *Lowering of X^0 to Y^0 (Embick and Noyer 2001: 561)*:
         $[_{XP} X^0 \ldots [_{YP} \ldots Y^0 \ldots]] \rightarrow [_{XP} \ldots [_{YP} \ldots [_{Y^0} Y^0 + X^0] \ldots]]$

Crucially, the two heads do not have to be adjacent on the surface, material may intervene between them. This is exactly what we need for Duala: C lowers to T, the head of its complement, across the subject in SpecTP. However, I would like to suggest that lowering only applies to the feature [OP:+OP] of C (plus the [CAT:V]-feature on C, see below), and not to the whole C head. The reason is that in embedded clauses, the complementizer precedes the subject, viz., it is realized in a different position than the operator feature. Once the feature [OP:+OP] is lowered to T, it is realized by the VI in (175). Afterwards, it is linearized at the right edge of the element in T (the finite verb), cf Biloa (1993).

Let me now turn to the subject/non-subject asymmetry. Why do only Ā-moved objects and adjuncts trigger *no*-marking? Since the reflex seems to depend on the grammatical function of the extracted XP, one might think that the case value of the extracted XP plays a role. However, there are two reasons why case cannot be the decisive factor: (i) Objects and adjuncts are not a natural class with respect to case because adjuncts do not bear case at all; they could only be referred to by a negative statement such as only non-nominative XPs trigger *no*-marking. (ii) However, this does not help either: Recall that in contrast to local subject extraction, long-distance movement of the subject of an embedded clause does trigger *no*-insertion in the clause where the subject lands. But since the element bears nominative even in the matrix clause, it should not be able to cause *no*-marking under the aforementioned generalization.

For this reason, I would like to propose a different solution: The crucial factor is whether the extracted element has moved through SpecvP on its way to SpecCP. Objects and adjuncts need to do so (they are base-generated in the VP); subjects do not move through SpecvP. They are base-generated in SpecvP, but this involves external Merge instead of internal Merge. So first, we need a distinction between external and internal Merge triggers. I will argue in section 4.1 for this split of Merge types. External Merge in SpecvP is triggered by the feature [•D•] and internal Merge is triggered by the edge feature [•EF•]. Second, the XP must track that it was in SpecvP at an earlier stage of the derivation. To achieve this, I assume that there is an upward Agree relation between v and an XP in SpecvP that results in copying of the category value from the head to the XP; thus, v bears a category probe [∗CAT:V∗] and every XP bears a corresponding unvalued category feature

(in addition to its inherent category feature). Under the order of operations in (174), it follows that only internal Merge to SpecvP feeds category-Agree with v (due to [•F•], [•EF•] > [∗CAT:V∗]); external Merge comes too late to feed Agree (due to [∗CAT:V∗] > [•D•]).

(174)   *Ordering statement in Duala (final version)*:
$\quad$ [•F•] > [∗OP:□∗], [∗CAT:□∗] > [•EF•] > [∗CAT:V∗] > [•D•]

The probe feature [∗CAT:V∗] is different from the probe features we have seen so far because it does not seek for a value; rather, it already bers an inherent value. It is discharged if it copies its value onto another element with a corresponding unvalued feature. Such a relation is independently necessary for case assignment where the goal does not bear a case value and is assigned case under Agree with a head that bears a value. Hence, probes can in principle bear values.[85]

As a consequence of the order in (174), a VP-internal element that undergoes Ā-movement will bear the category feature [CAT:V]; the subject DP does not, however. Finally, we need to make the realization of the operator feature in C sensitive to the category feature assigned in the vP. Suppose that C not only bears an [∗OP:□∗]-probe but also a category probe [∗CAT:□∗], viz. it agrees with the XP in SpecCP in its operator status and its category. If [∗CAT:□∗] applies late, at some point after both final and intermediate movement steps, Ā-movement always feeds category Agree with C (cf. the ordering in (174)). If a VP-internal argument is merged in SpecCP, it values the category probe on C to [CAT:V]; if the subject of the clause moves to SpecCP, it cannot value the category probe on C to [CAT:V] because it did not get this value in the vP. The VI *no* is sensitive to the presence of [CAT:V]:

(175)   *VI in Duala (final version)*:
$\quad$ /no/ ↔ [OP:+OP] / ___ [CAT:V]

As a result, only an Ā-moved VP-internal constituent triggers *no*-marking; a locally Ā-moved subject values the operator probe on C to [OP:+OP], but the context for *no*-insertion is not met because the subject cannot copy the value v to the category probe on C.[86]

Let me go through the derivation of local wh-movement of the internal argument of a transitive verb. In what follows, traces in trees indicate previous positions of an XP; but they are only mnemonic devices and have no theoretical status.

We start at the stage of the derivation where v merges with VP. The wh-phrase is at the edge of VP because V has been assigned an edge feature (*Phase Balance* triggers

---

[85]The Agree relation in category between the XP in SpecvP and v is used to encode a part of the derivational history on the moved item, which is relevant at a later stage of the derivation for the occurrence of a reflex of Ā-movement. This analysis of the subject/non-subject split is inspired by Müller (2013*a*; 2014) who develops a category-valuation analysis to model (temporary) memorizing of previous derivational steps (though for an empirically completely different domain). Müller assumes that the movement-related feature on an XP that checks the final movement trigger (e.g. the wh-feature, the rel-feature, etc.) lacks a value. When the XP moves to the specifier of a phase head, e.g. v, v and the XP Agree with the result that the movement-related feature on the XP gets a copy of the category of the phase head, cf. (i).

(i)   a.   unvalued movement-related feature on a wh-moved XP: [∗WH:□∗]
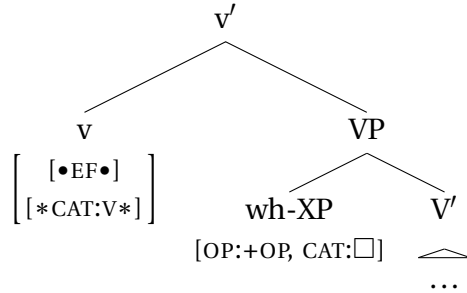$\quad$ b.   valued movement-related feature on a wh-moved XP after movement to SpecvP: [WH:V]

In this way, the category label on the XP tracks through which projections it has moved. I adopt this category-valuation model in my analysis of the subject/non-subject split in *no*-marking.

[86]Under local subject extraction, the category probe on C will be valued nevertheless, by the inherent category of the subject, but the value will be different from v and this is sufficient to prevent *no*-insertion.
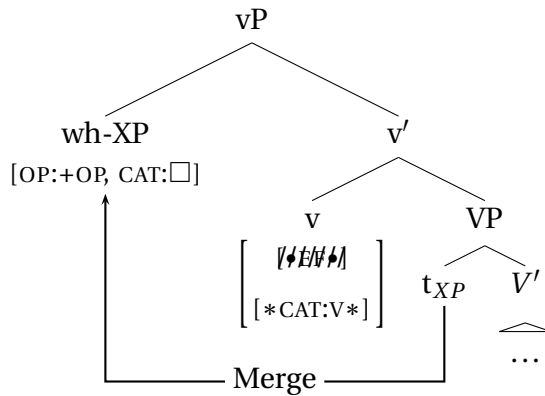
edge feature insertion because there is a C head in the numeration that bears [•WH•]). In the vP, early edge feature-driven movement of the wh-phrase to SpecvP feeds category assignment to the wh-phrase, cf. (176); v also bears an edge feature due to [•WH•] in the numeration. (I leave out the [•D•]-feature on v that will trigger external Merge of DP$_{ext}$ after all other operations.)
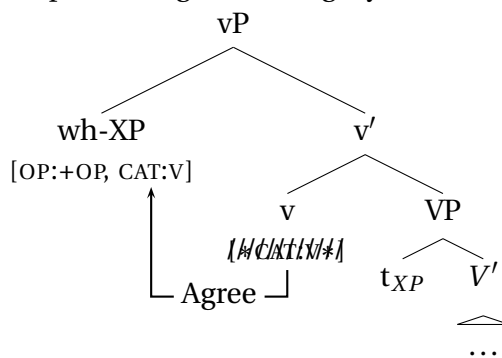
(176)   *Operations in the vP*:

    a.   v merges with VP

```
                    v′
                  /    \
                 v      VP
          ⎡ [•EF•]  ⎤  /   \
          ⎣[∗CAT:V∗]⎦ wh-XP   V′
                  [OP:+OP, CAT:□]  △
                                   …
```

    b.   Step 1: XP moves to SpecvP, [•EF•] is discharged

```
                      vP
                   /     \
               wh-XP      v′
          [OP:+OP, CAT:□]  /   \
             ↑            v      VP
             |      ⎡ [̶•̶E̶F̶•̶] ⎤  /   \
             |      ⎣[∗CAT:V∗]⎦ t_XP   V′
             |                         △
             |_____ Merge _____|   …
```

    c.   Step 2: v assigns its category to the wh-phrase, [∗CAT:V∗] is discharged

```
                      vP
                   /     \
               wh-XP      v′
          [OP:+OP, CAT:V]  /   \
             ↑            v      VP
             |      [̶∗̶C̶A̶T̶:̶V̶∗̶]  /   \
             |_ Agree _|     t_XP   V′
                                    △
                                    …
```

Afterwards, T merges with vP and triggers movement of the wh-phrase to its specifier (T bears an edge feature), cf. (177-a). When C is merged, it triggers internal Merge of the wh-XP first, cf. (177-b). Afterwards, C initiates upward Agree and can be valued by the wh-phrase in SpecCP, cf. (177-c-d) (the two Agree relations apply sequentially, but I illustrate them in one step). In the postsyntactic component, the valued features lower from C to T, where the VI in (175) is inserted in a subsequent step.

(177)   *Operations in the CP*:

    a.   C merges with TP
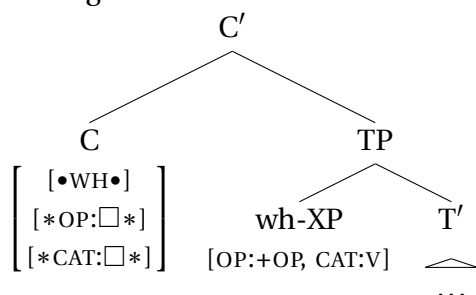
    b.   Step 1: XP moves to SpecCP, [•WH•] is discharged
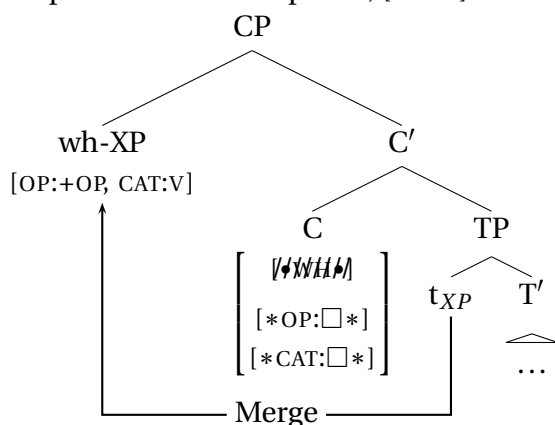
    c.   Step 2: C Agrees with the wh-phrase
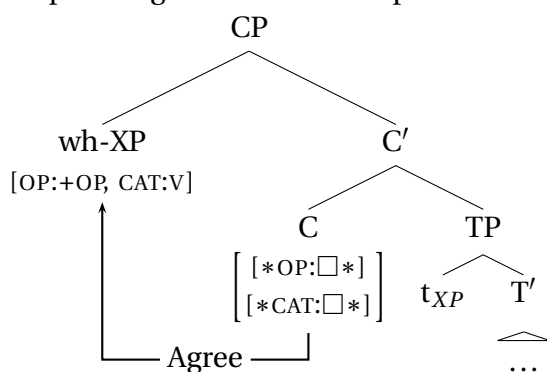
d. Result: probes are valued

```
                        CP
                   ╱          ╲
             wh-XP              C′
        [OP:+OP, CAT:V]      ╱      ╲
                           C          TP
                    [OP:+OP], [CAT:V]  ╱  ╲
                                    t_XP   T′
                                          ╱╲
                                          …
```
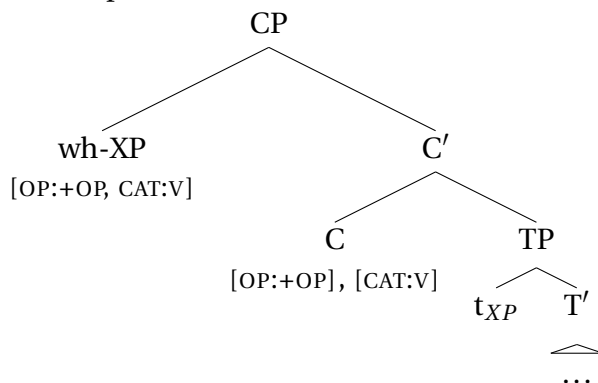
e. Postsyntactic component
Lowering of { [OP:+OP], [CAT:V] } from C to T

f. Insertion of *no* in T

This analysis of the subject/non-subject split correctly predicts that a long-distance moved subject triggers *no*-marking in the clause where the subject surfaces: Due to the PIC, the subject must move through every XP on the way to its final landing site; hence, it must target SpecvP of the higher clause (which embeds the clause it originates from). Since this involves internal Merge, movement applies early enough to feed category Agree with the higher v-head. Therefore, the subject is able to value the category probe on the C-head that projects its final landing site in case of long-distance extraction, and consequently, *no* will be inserted.

Furthermore, it predicts that the sole argument of an unergative verb does not trigger *no*-marking under extraction, whereas extraction of the sole argument of an unaccusative verb does. I repeat an example of extraction from an intransitive verb I presented in section 2.3.2.2:

(178) *Ā-movement of the sole argument of an intransitive verb in Duala*:

nja$_k$ ___$_k$ nu   pqi
who   3SG.REL come
'Who came?'

no-marking does not occur in this example, but I do not have independent evidence for whether the verb is unergative or unaccusative. I have to leave the verification of the prediction to future research.[87]

2.6.1.2 Complementizer agreement in Wolof *u*-chains

In Wolof *u*-chains, a complementizer agrees in definiteness and class with an operator that is Ā-moved to its specifier, e.g. wh-moved; this holds for all arguments and adjuncts. Since wh-phrases are indefinite, the complementizer surfaces in its indefinite form *u*; *u*

---

[87]If it turns out that the extraction of the sole argument of an unaccusative verb does not trigger *no*-marking, I would have to assume that Agree in category is related to the transitivity of the predicate. One way would be to postulate a feature co-occurrence restriction (FCR, Gazdar et al. 1985) according to which only a v that selects an external argument is assigned a category probe feature:
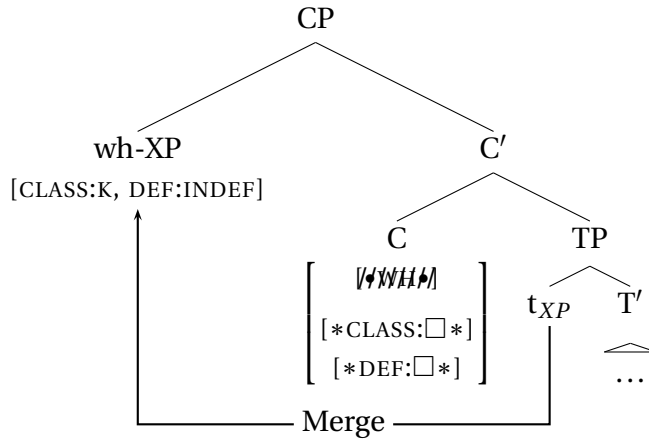
(i)   *FCR for v in Duala*: [•D•] ⊃ [∗CAT:V∗]

has a prefix that indicates the class of the wh-phrase. Hence, I assume that C bears a class-probe [∗CLASS:□∗] and a definiteness-probe [∗DEF:□∗] that are valued by the inherent class and definiteness feature of an XP under upward Agree. The reflex of movement can exhibit pattern I or II, viz. the final movement step must feed Agree, intermediate steps can feed Agree. This follows from the partial order of features in (179-a), which can be resolved into the total orders in (179-b-c):

(179)   *Ordering statement in Wolof (first version)*:

    a.   [•F•] ≻ [∗CLASS:□∗], [∗DEF:□∗] | [•EF•]   *partial order*

    b.   [•F•], [•EF•] ≻ [∗CLASS:□∗], [∗DEF:□∗]   *pattern I order*

    c.   [•F•] ≻ [∗CLASS:□∗], [∗DEF:□∗] ≻ [•EF•]   *pattern II order*

In case of pattern II, an intermediate movement step of an operator to SpecCP comes too late to feed class and definiteness Agree. In this case, we see that the default class marker *l* is inserted and the complementizer surfaces in its definite distal form *a* instead of the indefinite form *u*. As for class agreement, I assume that the probe is assigned a default value, call it [CLASS:X], if Agree does not result in valuation; it is realized as *l*, cf. the VIs in (180). Following the practice in the literature, the classes in Wolof are labeled by the letter corresponding to the phoneme that is characteristic for the class; human singular nouns, for instance, get the class prefix *k*, locative expressions bear the class prefix *f* (cf. McLaughlin 1997 and the references cited in Torrence 2012).

(180)   *VIs that realize class on C (first version)*:

    a.   /k-/ ↔ [CL:K] / ___ C

    b.   /f-/ ↔ [CL:F] / ___ C

    c.   …

    d.   /l-/ ↔ [CL:X] / ___ C

As for the definiteness probe, I assume that it is deleted by default if Agree does not result in valuation; *a* is the default realization of the C head that is inserted if no other definiteness feature is present on C, cf. the VIs in (181) (the definite proximate marker *i* is listed for completeness).[88]

(181)   *VIs that realize C*:

    a.   /u/ ↔ C / ___ [DEF:INDEF]

    b.   /a/ ↔ C

    c.   /i/ ↔ C / ___ [DEF:INDEF,PROX]

The following derivation illustrates feeding of Agree by the final movement step of a wh-phrase of class *k* (question for a person, 'who?'). We start at the point of the derivation where C merges with TP, cf. (182-a). The wh-phrase is located in SpecTP; it successive-cyclically moves to this position because the presence of [•WH•] in the numeration triggers edge-feature insertion on all lower heads. First, C triggers Merge of the wh-phrase to SpecCP, cf. (182-b). Afterwards, it triggers upwards Agree in class and definiteness that leads to valuation of both probes, cf. (182-c-f); note that a valued probe feature is removed from the stack of operation-inducing features because it does not trigger an operation anymore; valued features are stored on a different stack (hence, the class feature is no longer on top of the definiteness probe in (182-e)). In (182), class Agree applies before definiteness Agree, but the order of these two Agree relations does not matter, they could also apply in the reverse order.

---

[88] See also Martinović (2013: 33, 39) for the claim that *a* is the default realization of the C head.

(182)   *Class and definiteness Agree in Wolof, final movement step:*

     a.   C merges with TP

$$
\begin{array}{c}
\text{C}' \\
\diagdown \\
\text{C} \qquad\qquad \text{TP} \\
\left[\begin{array}{c} [\bullet\text{WH}\bullet] \\ [*\text{CLASS}:\square*] \\ [*\text{DEF}:\square*] \end{array}\right] \qquad \text{wh-XP} \qquad \text{T}' \\
[\text{CLASS:K, DEF:INDEF}] \qquad \ldots
\end{array}
$$

     b.   Step 1: XP moves to SpecCP, [•WH•] is discharged

$$
\begin{array}{c}
\text{CP} \\
\text{wh-XP} \qquad \text{C}' \\
[\text{CLASS:K, DEF:INDEF}] \qquad \text{C} \qquad \text{TP} \\
\left[\begin{array}{c} [\bullet\!\!/\!\!\text{WH}\!\!/\!\!\bullet] \\ [*\text{CLASS}:\square*] \\ [*\text{DEF}:\square*] \end{array}\right] \quad \text{t}_{XP} \quad \text{T}' \\
\text{Merge} \qquad \ldots
\end{array}
$$

     c.   Step 2: class-Agree triggered by C

$$
\begin{array}{c}
\text{CP} \qquad\qquad\qquad] \\
\text{wh-XP} \qquad \text{C}' \\
[\text{CLASS:K DEF:INDEF}] \qquad \text{C} \qquad \text{TP} \\
\left[\begin{array}{c} [*\text{CLASS}:\square*] \\ [*\text{DEF}:\square*] \end{array}\right] \quad \text{t}_{XP} \quad \text{T}' \\
\text{Agree} \qquad \ldots
\end{array}
$$

     d.   Result: class probe valued

$$
\begin{array}{c}
\text{CP} \\
\text{wh-XP} \qquad \text{C}' \\
[\text{CLASS:K, DEF:INDEF}] \qquad \text{C} \qquad \text{TP} \\
\left[\begin{array}{c} [\text{CLASS:K}] \\ [*\text{DEF}:\square*] \end{array}\right] \quad \text{t}_{XP} \quad \text{T}' \\
\ldots
\end{array}
$$

e.   Step 3: definiteness-Agree triggered by C

CP
wh-XP
[CLASS:K, DEF:INDEF]
C′
C
[CLASS:K], [∗DEF:□∗]
TP
Agree
t_{XP} T′
...

f.   Result: definiteness probe valued

CP
wh-XP
[CLASS:K, DEF:INDEF]
C′
C
[CLASS:K]
[DEF:INDEF]
TP
t_{XP} T′
...

g.   Insertion of the VIs /k/ and /u/ for [CLASS:K] and [DEF:INDEF], respectively

In case the pattern II order of operation-inducing features is chosen, an intermediate movement step of the wh-phrase comes too late to trigger class- and definiteness Agree with C.

(183)   *Intermediate movement step to SpecCP*:

a.   C merges with TP

C′
C
$\begin{bmatrix} [\ast\text{CLASS:}\square\ast] \\ [\ast\text{DEF:}\square\ast] \\ [\bullet\text{EF}\bullet] \end{bmatrix}$
TP
wh-XP
[CLASS:K, DEF:INDEF]
T′
...
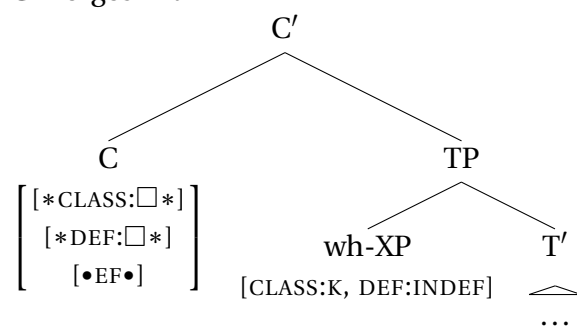
b.   Step 1: C initiates class and definiteness Agree but valuation fails

C′
Agree
C
$\begin{bmatrix} [\ast\text{CLASS:}\square\ast] \\ [\ast\text{DEF:}\square\ast] \\ [\bullet\text{EF}\bullet] \end{bmatrix}$
TP
wh-XP
[CLASS:K, DEF:INDEF]
T′
...
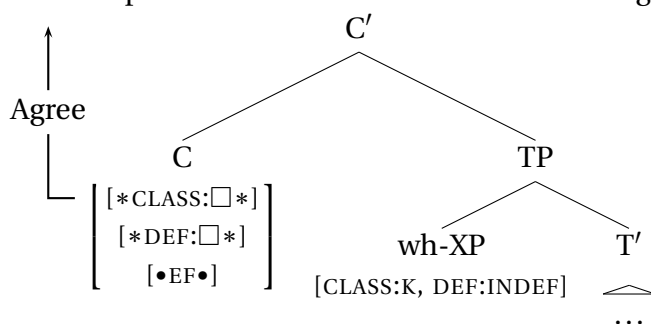
c. Result: default deletion of [∗DEF:□∗] and default valuation of [∗CLASS:□∗]

```
                          C′
                ┌─────────┴─────────┐
                C                   TP
      ⎡ [CLASS:X]        ⎤    ┌──────┴──────┐
      ⎢ [∗D̶E̶F̶:□∗]       ⎥  wh-XP          T′
      ⎣ [•EF•]          ⎦  [CLASS:K, DEF:INDEF]  △
                                              …
```

d. Step 2: movement of the wh-phrase
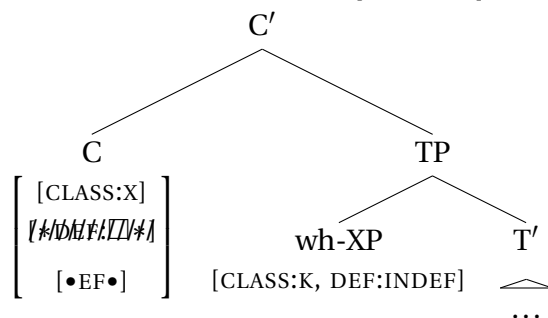
```
                        CP
              ┌──────────┴──────────┐
            wh-XP                    C′
      [CLASS:K, DEF:INDEF]    ┌──────┴──────┐
            ↑                 C             TP
            │         [CLASS:X], [•EF•]     │
            │                               T′
            │                               △
            │                          t_XP …
            └──────── Merge ───────────┘
```
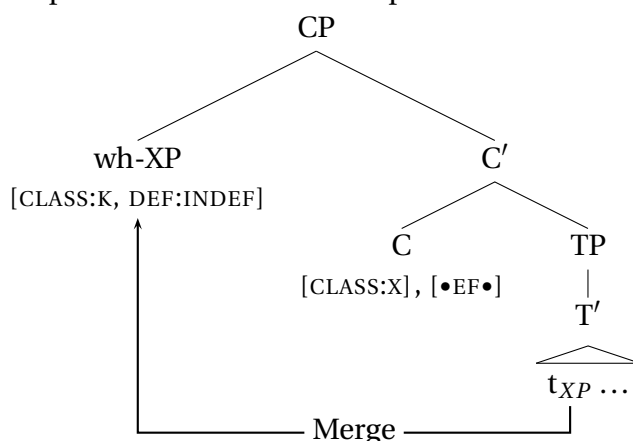
e. Insertion of the VIs /l/ and /a/ for [CLASS:X] and C (without a definiteness value), respectively

Wolof exhibits a subject/non-subject split not mentioned so far: In case of failure of class and definiteness Agree (as with intermediate movement steps under the pattern II order), the default class marker *l* only occurs if the object or an adjunct is Ā-moved, but it cannot occur in the clause from which a subject is extracted. In the latter case, only *a* surfaces, cf. the embedded clause in (184). However, a long-distance moved subject behaves like an extracted object in that *l* surfaces on the complementizer in higher clauses, where the subject is not base-merged. This is illustrated in (184-b) for long subject extraction in an *an*-chain (in *u*-chains, *l* cannot surface in the matrix clause in the first place because there is full class agreement on the complementizer that projects the final landing site of the wh-phrase; in *an*-chains, however, we find pattern III or IV, therefore the complementizer that triggers the final movement step cannot agree in class and definiteness, viz. it can in principle bear the default class prefix; hence, I provide an *an*-chain example here).

(184)   *Subject extraction and default class marking under long-subject extraction(Martinović 2013: 9,22)*:

a. [_CP_ Ø_k_ k-u   Musaa foog [_CP_ ___k mu **a** lekk gato bi   ]]
         Q   CL-*u* Musaa think   3SG *a* eat cake DEF.SG
    'Who does Musa think ate the cake?'

b. [_CP_ K-an_k_ **l-a**   Aali xam   [_CP_ ni   ___k mu **a** gis xale bi   ]]
         CL-an EXPL-*a* Aali know   that   3SG *a* see child DEF.SG
    'Who does ali know saw the child?'

Hence, *l-* has the same distribution as *no* in Duala. I propose the same analysis for this split: v assigns its category value to an XP that moves to SpecvP. C agrees with this cate-

gory value (in addition to class and definiteness Agree). The VI that leads to *l*-insertion is sensitive to the presence of the category [CAT:V] on C:

(185)   a.   Ordering statement for pattern II (final version):

[•F•] ≻ [∗CLASS:□∗], [∗DEF:□∗] ≻ [•EF•] ≻ [∗CAT:□∗], [∗CAT∗] ≻ [•D•]

b.   Default class VI (final version):

/l-/ ↔ [CL:X] / ___ C[CAT:V]

As a consequence of this ordering, only non-subjects that undergo an intermediate movement step to SpecCP will trigger insertion of the default class marker: The intermediate movement step counter-feeds class Agree, resulting in default valuation, but feeds category Agree. Therefore, the context for *l*-insertion is given on C. Subjects that make a stop-over in the minimal SpecCP also counter-feed class Agree but do not feed Agree in category v because they have not been assigned the value in the vP: External Merge in SpecvP comes too late to feed category valuation; only movement to SpecvP applies early enough to have this effect. But if a subject undergoes long extraction, it will move through SpecvP in all clauses in which it is not base-generated (due to the PIC); therefore, it will be assigned the category v and consequently, it can trigger *l*-insertion on the C-heads of the higher clauses.

2.6.1.3   ke-stranding in Dinka

In Dinka, the movement of a plural XP through SpecvP triggers the occurrence of the plural morpheme *ke* at the left edge of the vP (*ke*-stranding). This is a pattern III reflex: *ke* can only appear if the XP uses SpecvP as an intermediate movement step, but not if it remains in SpecvP (moved there from within the VP). Furthermore, *ke*-stranding exhibits a subject/object split: A plural external argument, which is base-merged in SpecvP, never triggers *ke*. To derive these facts, I assume that v bears a number probe [∗#:□∗] that seeks for a goal in SpecvP (upward probing). *ke* is the realization of the number value plural:

(186)   *VI in Dinka*:

/ke/ ↔ [#:PL]

The pattern III distribution and the subject/object asymmetry follow from the order of operations in (187):

(187)   *Ordering statement in Dinka (to be refined)*:

[•EF•] ≻ [∗#:□∗] ≻ [•F•], [•D•]

Intermediate movement steps apply before Agree and can thus feed Agree. Final movement steps to SpecvP and external Merge in SpecvP apply after v initiated number Agree; therefore, they counter-feed Agree (the split between internal and external Merge triggers was also needed for Wolof and Indonesian). As a result, only internal argument can ever feed number Agree. If the moving XP is plural and makes an intermediate movement step to SpecvP, it values the probe to [#:PL], which is realized postsyntactically by the VI in (186). The emergence of *ke*-stranding is illustrated in the derivation in (188): The plural direct object of a transitive verb undergoes wh-movement to SpecCP and makes a stop-over in SpecvP. The derivation starts when v merges with VP; the wh-phrase is at the edge of the VP.

(188)   *Intermediate movement step to SpecvP*:

    a.   v merges with VP

$$
\begin{array}{c}
v' \\
\diagup \qquad \diagdown \\
v \qquad\qquad VP \\
\begin{bmatrix} [\bullet EF\bullet] \\ [*\#:\square *] \end{bmatrix} \quad \text{wh-XP} \quad V' \\
\qquad\qquad [\#:\text{PL}] \quad \ldots
\end{array}
$$

    b.   Step 1: XP moves to SpecvP, $[\bullet EF\bullet]$ is discharged

$$
\begin{array}{c}
vP \\
\diagup \qquad\qquad \diagdown \\
\text{wh-XP} \qquad\qquad v' \\
[\#:\text{PL}] \qquad\quad \diagup \qquad \diagdown \\
\qquad\qquad v \qquad\qquad VP \\
\qquad\quad \begin{bmatrix} [\cancel{\bullet EF \bullet}] \\ [*\#:\square *] \end{bmatrix} \quad t_{XP} \quad V' \\
\qquad\qquad\qquad\qquad \ldots \\
\text{Merge}
\end{array}
$$
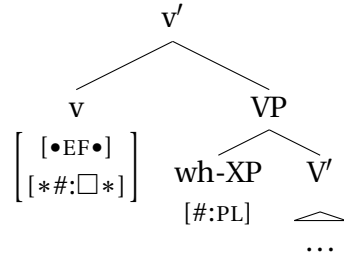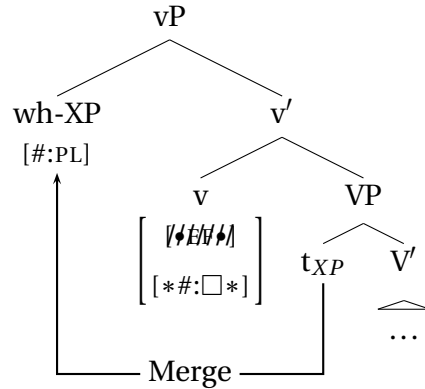
    c.   Step 2: number-Agree initiated by v

$$
\begin{array}{c}
vP \\
\diagup \qquad\qquad \diagdown \\
\text{wh-XP} \qquad\qquad v' \\
[\#:\text{PL}] \qquad\quad \diagup \qquad \diagdown \\
\qquad\qquad v \qquad\qquad VP \\
\qquad\quad [*\#:\square *] \quad t_{XP} \quad V' \\
\qquad\qquad\qquad\qquad \ldots \\
\text{Agree}
\end{array}
$$
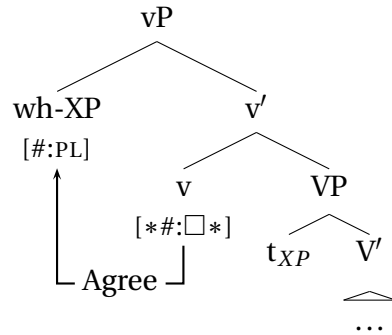
    d.   Result: number probe valued

$$
\begin{array}{c}
vP \\
\diagup \qquad\qquad \diagdown \\
\text{wh-XP} \qquad\qquad v' \\
[\#:\text{PL}] \qquad\quad \diagup \qquad \diagdown \\
\qquad\qquad v \qquad\qquad VP \\
\qquad\quad [\#:\text{PL}] \quad t_{XP} \quad V' \\
\qquad\qquad\qquad\qquad \ldots
\end{array}
$$
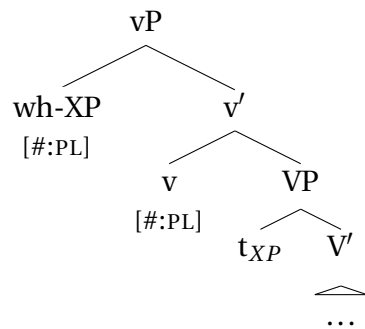
    e.   Postsyntactic component: insertion of the VI *ke* into v[89]

---

[89]If a singular XP is extracted, it values the probe to $[\#:\text{SG}]$. Either there is a zero VI for this number value or no matching VI at all; in any case, there will be no overt marker if a singular XP makes a stop-over in SpecvP.

Afterwards, the external argument is merged in SpecvP and the wh-phrase is successive-cyclically moved until it reaches its final landing site SpecCP where it checks the [•WH•] on C.

I follow van Urk and Richards (2013: 24) in assuming that *ke* is a clitic. It does not fill SpecvP; rather it marks the left edge of the vP, i.e. it cliticizes to the leftmost element of vP (which is the non-finite verb in the examples (101-a) and (101-c)).[90]

Let me now turn to the nature of the final movement trigger [•F•]. Recall that SpecvP (and SpecCP) must always be filled in Dinka, unless there is movement through these positions (cf. section 2.3.3.3). Hence, C and v inherently bear an EPP-like feature which triggers movement of an XP to SpecCP and SpecvP, respectively (although this generalization is not surface-true under movement). SpecCP can be filled by any XP, but SpecvP is more restricted: It can only host nominals but not PPs. van Urk and Richards (2013: 20f.) show hat this is not simply an argument-adjunct asymmetry: PPs cannot occupy SpecvP, regardless of whether they are adjuncts (of an intransitive verb) or arguments (locative argument of verbs like 'send'). Therefore, I assume that the EPP-feature on v (the feature [•F•] in (187)) is a [•D•]-feature that only attracts DPs (van Urk and Richards 2013 use a case feature to implement the difference: only DPs are active for case).[91] Note that in ditransitives, both objects are nominal, the indirect object is not a PP. The refined ordering statement looks as in (189); to distinguish the EPP from the external Merge trigger, the latter is identified by a θ-index, indicating that its discharge involves θ-role assignment (the feature labels do not matter, the only important thing is that the EPP and the external Merge trigger are different types of Merge features):

(189)   *Ordering statement in Dinka*:
          [•EF•] ≻ [∗#:□∗] ≻ [•D•], [•D$_\theta$•]

Since the [•D•]-feature is an inherent feature of v (filling of SpecvP is not optional), a question arises for the present analysis: If an XP makes an intermediate movement step to SpecvP, why is there an edge feature on v? There already is a trigger for movement to SpecvP, viz. the EPP. Recall that the purpose of edge feature insertion on a phase head H is to trigger movement of an element E from the complement domain of H to its edge in order to keep E (and its feature) accessible for heads that will be merged later: According to the PIC in (190-a), only the specifier and the head of a phase are accessible from outside. In the present example, the relevant feature is the wh-feature on the internal argument that must be accessible for the [•WH•]-feature on C. But since v inherently bears the EPP-feature which triggers movement of the internal argument to SpecvP anyway, one might think that an edge feature is not needed to get the wh-phrase to the edge of SpecvP. Hence, from a surface-oriented perspective, edge feature insertion should be superflu-

---

[90]It is not sufficient to say that *ke* cliticizes to the non-finite verb because, as van Urk and Richards (2013: 24, fn.17) emphasize, it must precede an XP in SpecvP in case that position is filled. (i) is ungrammatical with *ke* after the direct object in SpecvP:

(i)   *Position of ke in the vP (van Urk and Richards 2013: 22)*:

        [Ye `ɛ̀ɛ̀i   k´o̱   ]$_k$ cénné   nyank´ái ké wanm´a̱th tu̱ɔ̀ɔc ___$_k$
         Q villages which PERF.OBL sister   PL brother   send
        'Which villages did my sister send my brother to?'

[91]Actually, it is not only DPs that can fill SpecvP, but also CP can do so according to van Urk and Richards (2013: 26ff.). So, the EPP-feature is something more abstract than D, it is a feature that DPs and CPs have in common, for example [+N(ominal)]. This can be achieved if category features are decomposed into abstract binary features (cf. Chomsky 1965), e.g. D [+N, +D]; C [+N, –D], P [–N, –D], etc.

ous: The XP moving to SpecvP discharges the EPP-feature. It is crucial for my analysis, however, that a plural XP moving through SpecvP checks an edge feature and not the EPP. If it checked the EPP, movement could not feed number Agree because EPP-driven movement applies too late (we know this because XPs that have SpecvP as their final landing site, triggered by the EPP on v, cannot cause *ke*-stranding, cf. (102-b)). Importantly, the version of *Phase Balance* I proposed in section 2.5.2.3 is formulated in such a way that it actually predicts edge feature insertion in this configuration, as required for the present ordering approach. The definition is repeated below:

(190) *Phase Balance, applies in the numeration*:
   An edge feature [•EF•] is inserted on the selected phase head H for every feature [•F•] on a head Y in the numeration if:
   a. Y ≠ H,
   b. there is no accessible matching feature [F].

(191) *Accessibility*:
   A feature [F] is accessible if it is part of the workspace and not selected (workspace: lexical items in the numeration, previously generated trees that are unconnected to the current phrase marker).

When v is selected for Merge, there is another head in the numeration with a structure-building feature, viz. C. It bears [•WH•] in our example (or an EPP-feature if there is no wh-movement). C and v are distinct and there is no accessible matching feature [WH] left in the numeration (assuming there is only a single wh-phrase to begin with); the relevant feature is already part of the current phrase marker: DP_{int}, which bears [WH] in (188), is merged in the VP. Therefore, an edge feature is inserted on v, although v has an EPP-feature that triggers Merge to SpecvP. This is possible because *Phase Balance* applies in the numeration, before v has started to discharge any of its inherent operation-inducing features. If the system could wait until a later stage of the derivation where XP will be in SpecHP because of the inherent EPP-feature, there would be no need to insert an edge feature anymore. But since an apparently superfluous edge feature is inserted in Dinka, the system must be blind to what happens at later stages. This provides strong evidence for a strictly derivational system where *Phase Balance* is checked very early (and does not only apply at the XP-level, as in Heck and Müller 2000). We will see that the same situation arises in other languages as well, see section 3.2.2 on Hungarian and section 3.3 on Icelandic.

Moreover, there is independent evidence that an XP which uses SpecvP as an intermediate landing site does not discharge the EPP-feature (but rather an edge feature). Under certain circumstances, the remaining EPP-feature can be checked by a different phrase when XP moves through SpecvP: If a VP-internal PP undergoes movement to SpecCP through SpecvP (for which the possibility of *ke*-stranding provides independent evidence, cf. example (i) in footnote 90), there must be a DP that fills SpecvP, cf. (192).

(192) *PP-extraction (van Urk and Richards 2013: 22)*:

   Yétenô$_k$ cíi   yín thòk v̤ɔ̤ɔc   ___$_k$
   where   PERF.NS you goat buy.TRANS
   'Where did you buy a goat?'

The PP checks the edge feature because it moves through SpecvP; we would thus expect that another DP can check the EPP, which is the case in (192). Hence, we have strong evi-

dence that movement through SpecvP does not discharge the EPP but a different feature; therefore, the ordering analysis, which presupposes a split between final and intermediate movement triggers, can be upheld. Note finally that the EPP cannot be discharged by movement of a DP if another DP undergoes movement through SpecvP; the examples are repeated below:

(193)   *Obligatory empty SpecvP under DP_{int}-extraction   (van Urk and Richards 2013: 15)*:

    a.   Yeŋà$_k$ cíi   mòc ☐ yiɛ́n ___$_k$ kitàp
          who   PERF.NS man   give   book
          'Who did the man give the book to?'
    b. *Yeŋà$_k$ cíi   mòc kitàp$_n$ yiɛ́n ___$_k$ ___$_n$
          who   PERF.NS man book   give
          'Who did the man give the book to?'

I propose the following solution to the absence of EPP-movement under DP-extraction (recall that the EPP is a D-feature [•D•] because it can only attract nominals): The EPP on v is such that it can also be discharged under matching with an XP in SpecvP, viz. it does not trigger DP-movement if there already is a DP in SpecvP at the point of the derivation where the EPP is checked (put differently, the EPP-requirement says that there must be a D-element in SpecvP, if there is none, it triggers movement, otherwise it is trivially fulfilled). We know independently from *ke*-stranding that the order of operations is [•EF•] > [∗#:☐∗] > [•D•]; hence, when a DP makes a stop-over in SpecvP, discharging v's edge feature, it is present in SpecvP when the turn of the EPP has come. Therefore, it is discharged under matching. When a PP occupies SpecvP, however, the EPP cannot be discharged because the PP and the EPP do not match in category. Hence, the EPP triggers DP-movement as in (192).[92]

I will not provide detailed analyses for the addition reflexes in Chamorro (complementizer agreement), Indonesian (focus marking), and Kitharaka (focus marking). They do not exhibit any splits, so the derivations can be straightforwardly deduced from the ones presented so far (see the summary of the analyses in table 2.3 on page 181 for the ordering statements).

---

[92]The solution proposed by van Urk and Richards (2013) is similar. They also assume that there are two movement triggering features on v: One is a feature that triggers the movement step (of a wh-phrase) to SpecvP ([UWH]); the other feature is the EPP which they identify as a case feature [UCASE]. If a PP undergoes wh-movement, it checks the movement trigger but cannot check the [UCASE]-feature because it is inactive for case; hence, another DP checks [UCASE]. If, however, a DP wh-moves, it can check both features because it is active for case assignment. That it *must* check both features whenever possible is ensured by the transderivational constraint Multitasking.

(i)   *Multitasking (van Urk and Richards 2013: 24)*:
       At every step in a derivation, if two operations A and B are possible, and the features checked by A
       are a superset of those checked by B, the grammar prefers A.

The present solution to the difference between DP and PP extraction with respect to the availability of additional EPP-movement does not require such a constraint; if discharge is possible under matching, the EPP will automatically be discharged when a DP undergoes Ā-extraction because of the order of operations is intermediate movement step before final movement step (independently motivated by the distribution of *ke*): The edge-feature moved DP is in SpecvP at the point of the derivation when the EPP is checked. I thank Fabian Heck for pointing out to me that under my ordering approach the transderivational constraint is not needed.

2.6.2 Deletion

In some languages, the reflex of movement results in deletion of exponents that would be present without movement. Deletion is analysed as follows: H Agrees with an XP in its specifier. The value v that H is thereby assigned provides the context for a postsyntactic impoverishment rule. This rule causes deletion of the feature [G] on H in the context of v, viz. valuation of the feature [F] on H feeds impoverishment of [G] on the same head. As a consequence, the exponent that realizes the feature [G] can no longer be inserted into H. The valued feature [F:V] is also not realized. The result is that there are fewer exponents than there would be without movement. Crucially, deletion is not literally erasue; rather, the exponent is not inserted in the first place, due to previous impoverishment.

2.6.2.1 Downstep deletion in Kikuyu

In Kikuyu, the downstep (a floating tone) associated with the verb is deleted on all verbs crossed by Ā-movement (pattern I). According to Clements (1984*a*;*b*) the downstep only occurs on verbs in affirmative declarative sentences. Based on this observation, Lahne (2008*b*: 114) concludes that the downstep is the morphophonological realization of the polarity feature with the value affirmative [POL:AFF], cf. the VI in (194). I follow her in this assumption.

(194)   *Vocabulary item*:
         /$^!$/ ↔ [POL:AFF]

Furthermore, I assume that polarity is encoded on the C head of every clause. I analyse deletion as the result of a postsyntactic impoverishment rule that deletes the feature [POL:+AFF] on C in the context of a positively valued operator-feature on C. C receives a value under upward Agree with an XP moved to SpecCP. Thus, C bears an [∗OP:□∗]-probe feature. If the XP moved to SpecCP is an operator, it values this feature to [OP:+OP]. The presence of [OP:+OP] on C feeds the postsyntactic impoverishment rule in (195) because [OP:+OP] provides the context for the rule. The result of impoverishment is that the feature [POL:+AFF] on C is deleted.

(195)   *Impoverishment rule on C*:
         [POL:+AFF] → Ø / ___ [OP:+OP]

As a consequence, the downstep VI in (194) can no longer be inserted, viz. we get downstep deletion under operator movement. If, however, the XP in SpecCP is not an operator and values the operator probe to [OP:–OP], the deletion rule in (195) is not fed and the downstep is not deleted.[93] Movement to SpecCP can only feed Agree by the [∗OP:□∗]-probe if movement precedes upward Agree. Since downstep deletion has a pattern I distribution, both final and intermediate movement steps must apply before Agree to value the probe on C, cf. (196):

(196)   *Ordering statement in Kikuyu*:
         [•F•], [•EF•] ≻ [∗OP:□∗]

The derivation is illustrated in (197) for the final movement step of a wh-phrase to SpecCP (the inherent polarity feature of C is not an operation-inducing feature, therefore it is outside of the stack). The derivation starts at the point where C merges with TP. The wh-

---

[93]See section 2.6.3.3 on details of Lahne's (2008*b*) account of deletion reflexes of movement.

phrase is at the edge of the TP; it has been successive-cyclically moved there because edge features have been inserted on all lower heads (due to $C_{[\bullet WH\bullet]}$ in the numeration).

(197)   *Final movement step of an operator in an affirmative clause*:

    a. C merges with TP

$$
\begin{array}{c}
\text{C}' \\
\diagup\diagdown \\
\text{C}_{[\text{POL:+AFF}]} \qquad \text{TP} \\
\begin{bmatrix} [\bullet\text{WH}\bullet] \\ [*\text{OP:}\square*] \end{bmatrix} \quad \text{wh-XP} \quad \text{T}' \\
[\text{OP:+OP}] \quad \dots
\end{array}
$$

    b. Step 1: XP moves to SpecCP, $[\bullet\text{WH}\bullet]$ discharged

$$
\begin{array}{c}
\text{CP} \\
\text{wh-XP} \qquad \text{C}' \\
[\text{OP:+OP}] \\
\text{C}_{[\text{POL:+AFF}]} \quad \text{TP} \\
\begin{bmatrix} \cancel{[\bullet\text{WH}\bullet]} \\ [*\text{OP:}\square*] \end{bmatrix} \quad \text{t}_{XP} \quad \text{T}' \\
\text{— Merge —} \quad \dots
\end{array}
$$

    c. Step 2: Agree triggered by C

$$
\begin{array}{c}
\text{CP} \\
\text{wh-XP} \qquad \text{C}' \\
[\text{OP:+OP}] \\
\text{C}_{[\text{POL:+AFF}]} \quad \text{TP} \\
[*\text{OP:}\square*] \quad \text{t}_{XP} \quad \text{T}' \\
\text{— Agree —} \quad \dots
\end{array}
$$

    d. Result: probe valued

$$
\begin{array}{c}
\text{CP} \\
\text{wh-XP} \qquad \text{C}' \\
[\text{OP:+OP}] \\
\text{C}_{[\text{POL:+AFF}]} \quad \text{TP} \\
[\text{OP:+OP}] \quad \text{t}_{XP} \quad \text{T}' \\
\dots
\end{array}
$$

    e. Postsyntactic component: polarity feature deleted by (195)

        $C_{[\cancel{\text{POL:+AFF}}],[\text{OP:+OP}]}$

    f. Insertion of the VI in (194) fails

The derivation proceeds in exactly the same way for intermediate movement steps; the only difference would be that the structure-building feature on C is [•EF•] instead of [•WH•]. Note finally that it is not important for the analysis where the polarity feature is located. It could also be on a functional head below C, e.g. on T (see e.g. Zanuttini 1997). The deletion rule would then apply to this head. The only difference would be that movement to the specifier of that head would always be an intermediate movement step (assuming that operator movement targets the structurally higher position SpecCP). As long as intermediate movement steps apply before Agree, movement feeds upward Agree with the [∗OP:□∗]-probe on the lower head.

2.6.2.2 *men*-deletion in Indonesian

In Indonesian, the verbal prefix *men* must be deleted when a VP-internal argument is extracted. Indonesian exhibits an asymmetry: Deletion neither occurs with adjunct extraction nor with local subject extraction, but only with object extraction. If deletion is triggered, it affects all clauses along the path of movement (pattern I). Following the literature on *men*-deletion (cf. Sato 2008 for an overview), I assume that this verbal prefix is a voice marker that encodes active voice, cf. the vocabulary item in (198):[94]

(198)   Vocabulary item:
          /men/ ↔ [VOICE:ACT]

The first question is where the voice feature is located, i.e. on which head deletion takes place. Semantically, voice should be located on a low functional head in the extended projection of the verb such as v. Indeed, there is further evidence for the hypothesis that voice is encoded very low in the structure: *men* always attaches to the main verb and not to the complementizer or auxiliaries, cf. (199).

(199)   *Location of the active voice prefix (Macdonald and Dardjowidjojo 1967: 153)*:
          a. John telah mem-beli buku itu   di kedai semalam
              John has   TRANS-buy book DEF at shop yesterday
              'John has bought the book at the shop yesterday.'
          b. Saya akan mem-beli-kan   Siti buku itu
              I   will TRANS-buy-DITRANS Siti book DEM
              'I am going to buy that book for Siti.'

There is also no verb raising to C under Ā-movement: We have seen in section 2.3.1.4 that in case of non-subject extraction the Ā-moved constituent precedes the subject, which in turn precedes the finite verb. Further evidence for a low position of the voice feature comes from the fact that *men*-deletion is not only triggered by Ā-movement, but also by A-movement, cf. the example in (200) from Malay (which has the same deletion process as Indonesian). See Chung (1994) for evidence that this is really an instance of A-movement and not Ā-topicalization; one argument for this view is that the moved DP can be a controlled PRO.

---

[94]The prefix was glossed as a 'TRANS' in section 2.3.1.4 because I adopted the examples from Saddy (1991) who identified it as a transitivity marker. But since *men* can also occur with inherently intransitive verbs and derived intransitive verbs such as passivized verbs (see below for examples), it is clearly not a marker of the feature [+TRANSITIVE].

(200)   *Men-deletion under A-movement in Malay (Cole and Hermon 1998: 232)*:

Ali$_k$ saya (*men)-cubit ___$_k$

Ali I   TRANS-pinch

'I pinched Ali.'

If *men*-deletion was a reflex of movement on the C-head, it should not be triggered by A-movement to SpecTP. If voice is encoded on v, however, the reflex is triggered even with A-movement because the VP-internal argument must make a stop-over in SpecvP on its way to SpecTP (since every phrase is a phase). I conclude that *men* realizes active voice on v; the main verb picks up the inflection by V-to-v-movement. Deletion of *men* is thus a reflex of movement through SpecvP.

I analyse *men*-deletion as the result of an impoverishment rule: When an XP moves to SpecvP, v Agrees with the XP in its case feature (see the discussion below for why case is the relevent feature). Indonesian does not have overt case markers but I assume that arguments bear abstract case. To be able to enter into case-Agree with case-marked XPs, v bears a case probe [∗CASE:□∗] that probes upwards. If it is valued, the postsyntactic impoverishement rule in (201) is fed:

(201)   *Impoverishment rule on v*:
         [VOICE:ACT] → Ø / ___ [CASE:α]
         (α is a variable over case values)

It deletes the feature [VOICE:ACT] on v in the context of a valued case feature on the same head. As a consequence, the VI *men* in (198) cannot be inserted on v, and hence, we get "deletion" under movement (although there is no literal deletion).[95]

In order to enter into upward Agree with the case-probe on v, the XP must move to SpecvP before v probes. Since *men*-deletion is triggered by Ā-movement that terminates in SpecCP and A-movement that ends in SpecTP, all movement steps through SpecvP are intermediate steps. They feed Agree under the order in (202) (the trigger for final movement steps [•F•] is mentioned for completeness in the ordering statement, although it does not play a role on v):

(202)   *Ordering statement*:
         [•F•], [•EF•] ≻ [∗CASE:□∗]

Let me go through a derivation which involves wh-movement of the internal argument of a transitive verb in active voice to the minimal SpecCP position. Being an internal argument, the wh-phrase is base-merged as the complement of V and moves to SpecVP to check the edge-feature on V, which is inserted because there is a C-head with the feature [•WH•] in the numeration (cf. the definition of *Phase Balance* in (148)). Afterwards,

---

[95]Recall that the occurrence of *men* is optional in clauses without extraction. The optional absence of the prefix in such clauses cannot be due to the impoverishment rule in (201) because this rule is not fed by a previous movement operation to SpecvP. To account for the optionality, I assume that there is a zero VI in addition to /men/ that realizes the same features:

(i)   Active voice VIs in Indonesian:
       /men/ ↔ [VOICE:ACT] /Ø/ ↔ [VOICE:ACT]

Since the markers are equally specific, both can be inserted into v with the active voice feature, see Hein (2008) for such an account of optionality of VIs (possible if the Subset Principle is slightly reformulated in such a way that it allows for insertion of *one of the most specific VIs,* see Hein 2008: 61f.). Under extraction, this feature is deleted and hence, none of the two VIs matches the impoverished feature set of v.

v merges with VP. I assume that internal arguments are assigned case by v, and external arguments are assigned case by T. In the present example, v assigns case to the wh-phrase in SpecVP (under downward Agree because case is also assigned if the DP does not move at all). Case assignment is not indicated in (203), we start at the point where DP_{int} already bears the case value, cf. (203-a). v has also been assigned an edge feature in the numeration because of the presence of the [•WH•]-feature on C in the numeration. Due to the order of operations in (202), v first attracts the wh-phrase to SpecvP, cf. (203-b). Aferwards, v probes upwards and finds the wh-phrase with a case value in its specifier. As a result, the case probe on v is valued, cf. (203-c-d). The presence of a valued case feature on v feeds the impoverishment rule in (201) which deletes the feature [VOICE:ACT]. Consequently, the VI *men* cannot be inserted into v in the postsyntactic component.

(203)   *Derivation: final movement of an operator in an affirmative clause*:

    a.   v merges with VP
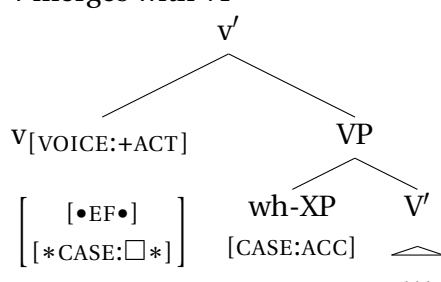
    b.   Step 1: XP moves to SpecvP, [•EF•] is discharged
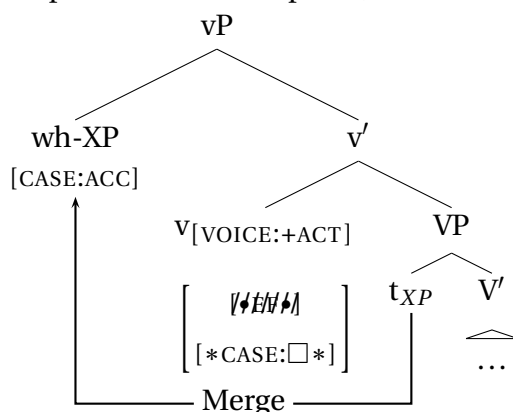
    c.   Step 2: case-Agree initiated by v

d. Result: case probe valued

$$
\begin{array}{c}
\text{vP} \\
\diagdown \\
\text{wh-XP} \qquad \text{v}' \\
[\text{CASE:ACC}] \\
\text{v}_{[\text{VOICE:}+\text{ACT}]} \quad \text{VP} \\
[\text{CASE:ACC}] \quad \text{t}_{XP} \quad \text{V}' \\
\ldots
\end{array}
$$
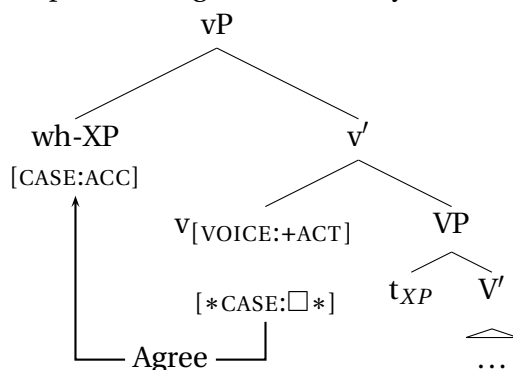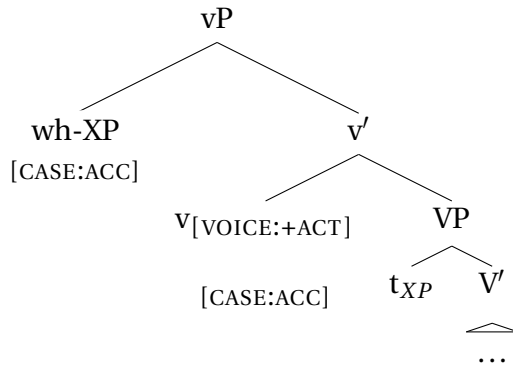
e. Postsyntactic component: active voice feature deleted by (201)

V$_{[\overline{\text{VOICE:+ACT}}],[\text{CASE:ACC}]}$

f. Insertion of the VI in (198) fails

I left out Merge of the external argument in the derivation. Assume for the sake of concreteness that external Merge is triggered after case-Agree, viz. DP$_{ext}$ is base-merged in the outer specifier of vP. But nothing hinges on this decision, external Merge does not interfere with case-Agree and impoverishment, as will become clear in the following discussion.

Why do I assume that case is the feature in which v Agrees with an XP in its specifier? In this way, we can derive the asymmetry in the occurrence of the reflex of movement: Recall that *men*-deletion never occurs under extraction of a transitive subject or of an adjunct, but only if a transitive object is extracted.[96] Subjects and adjuncts form a natural class in that they do not bear a case value within vP. Adjuncts do not have case at all and subjects are assigned case at a later stage in the derivation by T. VP-internal arguments, however, are assigned case by v. Assuming that case assignment applies very early (so that it cannot be bled by movement of an internal argument), the internal argument bears a case value when it moves to SpecvP. Consequently, only movement of an object can ever feed case-Agree with v, which is a precondition for impoverishment of the active voice feature. I assume that the case probe [∗CASE:□∗] on v is deleted by default if it does not find a matching feature in SpecvP.[97]

This case-based account of the object/non-object split in Indonesian predicts that movement of the sole argument of an unaccusative verb – which starts out as an internal argument (*Unaccusative Hypothesis*, cf. Perlmutter 1978; Burzio 1986) – cannot trigger *men*-deletion. The reason is the following: In many accounts of ergative and accusative encoding systems, v assigns the morphologically marked case (ergative or accusative) and T assigns the morphologically unmarked case (absolutive or nominative), cf. e.g. Murasugi (1992); Müller (2009) and Bittner and Hale (1996) for a related proposal. In intransitive contexts, one of the heads must not be a case assigner; we know that the sole argument in ergative and accusative systems bears the unmarked case, hence, only T must be active in intransitive contexts.[98] But if the sole argument of an unaccusative verb is assigned case by T, it does not bear a case value in the vP, and therefore, movement of

---

[96]With a ditransitive verb, extraction of either of the objects leads to the deletion of *men*, see Sato (2012: 50f.) for examples.

[97]I thank Martin Salzmann for the suggestion that the object/non-object split follows from case-Agree.

[98]In the accounts of alignment patterns by Chomsky (1993); Bobaljik (1993); Laka (1993); Řezáč (2003), this only holds for the nominative-accusative pattern; in languages with an ergative-absolutive pattern the sole argument is assigned case by v, however.

this argument to SpecvP cannot feed case Agree with v; the argument is assigned case at a later stage of the derivation. This prediction is indeed borne out, see the example in (204) where the argument of an unaccusative verb is A-moved and the voice prefix can occur (for reasons of consistency I continue to gloss *men* as a transitivity marker as in section 2.3.1.4, although it is a voice marker). It is clear that A-extraction has applied in this example because the argument surfaces in preverbal position; in contrast to the sole argument of unergatives, it can also surface in its basic post-verbal position, see Soh and Nomoto (2009: 44ff.):[99]

(204)   *Extraction of the sole argument of unaccusatives (Standard Indonesian; Sato 2012: 45, based on Soh and Nomoto 2009)*:

[Tarif listrik   ]$_k$ (me-)turun ___$_k$
  price electricity   TRANS-fall
'The electricity price fell.'

Thus, extraction of an unaccusative subject patterns with extraction of a transitive subject rather than with extraction of a transitive object.[100]

   In addition, the present analysis correctly predicts another asymmetry: Local subject extraction does not cause *men*-deletion, but long subject extraction triggers deletion in all clauses affected by movement above the clause where the subject DP is base-merged, cf. (205).

(205)   *Long subject extraction in Indonesian*:

[$_{CP}$ Siapa$_k$ yang Bill (*men)-beri [$_{CP}$ Tom (*men)-harap [$_{CP}$ ___$_k$ (men)-cintai
    who   FOC Bill TRANS-think   Tom TRANS-expect   TRANS-love
Fred]]]
Fred
'Who does Bill think Tom expects loves Fred?'

---

[99]There is independent evidence e.g. from causativization that the verb 'fall' in Indonesian is unaccusative, see Sato (2012: 44) and references cited there.

[100]Previous approaches to the split in *men*-deletion wronlgy predict that extraction of the sole argument of an unaccusative verb must trigger deletion. Aldridge (2008) and Sato (2012) suggest that the argument/adjunct asymmetry follows if *men*-deletion is sensitive to the category of the element moved through SpecvP: Only elements of category D can trigger deletion; since adjuncts are either adverbs of category Adv or PPs, they do not cause deletion, only arguments can do so (the subject/object asymmetry has a different explanation which need not concern us here). Sato (2012) explicitly states that this predicts *men*-deletion for extraction of the sole argument of an intransitive verb, because it is also of category D. He proposes the following solution, which I take to be unsatisfactory: Unaccusative v is not a phase head (cf. Chomsky 2001) and therefore, movement of the sole argument does not make a stop-over in SpecvP when it is extracted (see also Aldridge 2008 for this assumption). However, it has repeatedly been argued against the hypothesis that unaccusative v is not a phase head because the vP it projects displays the same properties as transitive vP (with respect to reconstruction, linearization, cf. Fox 2000; Legate 2003; Sauerland 2003; Richards 2004; Epstein 2007). In addition, Sato (2012) needs to assume that there are actually two homophonous *men*-VIs: In his account, *men* spells out the feature D on v which triggers movement of the internal argument to SpecvP (a kind of edge feature). This feature is deleted when movement applies, accounting for *men*-deletion. Unaccusative v, which is not a phase head by assumption, does not have a corresponding movement-trigger, and consequently, *men* cannot be deleted by movement of the sole VP-internal argument. But now it is unclear how *men* can surface with unaccusative verbs in the first place if it spells out the D-feature. Sato (2012: 47) claims that the *men* we see in this context is different from the one we see on non-unaccusative verbs without extraction from VP: The unaccusative *men* spells out the feature [+progressive] because it "contributes a progressive viewpoint aspect". Hence, the syncretism is not accounted for. In the present analysis, the syncretism is resolved, and it is correctly predicted that deletion does not apply when the sole argument of an unaccusative verb is extracted. I take this to be a desirable outcome. For a discussion of alternative accounts of *men*-deletion, which I will not review here, see Sato (2008: ch.2).

This follows because the subject is assigned case by T in the clause where it originates and the PIC enforces that it makes a stop-over in SpecvP of structurally higher clauses when it undergoes cross-clausal movement. Since the subject already bears a case value if it stops in SpecvP of higher clauses, it will feed case Agree with v and therefore impoverishment.

2.6.3   Replacement

In some languages Ā-movement has a morphological reflex that replaces an exponent that is present without movement. Replacement can have different sources in the present approach where movement initially leads to addition of a feature (valuation under Agree): (i) Blocking by contextual allomorphy: The feature [F] valued on a head H under Agree with a moving XP provides the context for the insertion of VI α on an adjacent terminal node G; α is identical to another VI β in its feature make-up except for the fact that β does not have the context feature [F]. Whenever [F] is present on H, the more specific VI α blocks β-insertion on G. (ii) Blocking without invoking the context: The VI β realizes features of the head H. The VI α realizes the same features as β plus the feature [F] which is valued on the same head H under Agree with a moving XP. Thus, the more specific α blocks β in case movement feeds Agree. (iii) Deletion: Valuation of [F] on H feeds an impoverishment rule that deletes another feature [E]. [E] is realized by the VI α; if deletion applies, α can no longer be inserted, but another VI β which is less specific than α and which does not contain [E] will be realized; β thus replaces α. I will illustrate each strategy with an example.

2.6.3.1   Pronoun choice in Ewe

In Ewe, the form of the 3rd person subject pronoun *é* changes to *wò* if an XP (argument or adjunct) undergoes Ā-movement (wh-movement, focus movement, relativization). Collins (1993: 41f.) argues that the subject is located in SpecTP in Ewe because it precedes auxiliaries, negation, and aspectual heads:

(206)   *Position of the subject in Ewe (Collins 1993: 40, 161)*:

    a. me fo   Kofi
         I   hit Kofi
         'I hit Kofi'

    b. me le   Kofi fo
         I   PROG Kofi hit
         'I am hitting Kofi'

    c. bé mì mé gă   dzó o
         that 1PL NEG SUBJ leave PRT
         'that I don't leave'

Movement to SpecTP does not trigger the change in the pronoun form, only Ā-movement to SpecCP does (see e.g. (79-a) with the 3rd person subject pronoun in its default form in SpecTP). Thus, replacement of the subject pronoun is a reflex in the TP domain that is due to movement to SpecCP. I assume that the reflex is due to an Agree relation between C and an operator in its specifier. C bears an [∗OP:□∗]-feature that probes upwards. To feed Agree, movement to SpecCP must apply before Agree. Replacement of the subject pronoun can follow pattern I or pattern II in Ewe. Hence, final movement steps must feed and thus precede Agree, intermediate movement steps can feed Agree, but do not have to. As shown in section 2.5.3.2, this follows if the ordering statement is underspecified, cf. the partial order in (207-a) and the resulting total orders in (207-b-c):

(207)   *Ordering statement in Ewe*:

    a.   [•F•] > [∗OP:□∗] | [•EF•]   *partial ordering*

    b.   [•F•], [•EF•] > [∗OP:□∗]   *pattern I order*

    c.   [•F•] > [∗OP:□∗] > [•EF•]   *pattern II order*

The change in the form of the 3rd person subject pronoun follows if the realization of the pronoun is sensitive to the presence or absence of the positively valued operator feature [OP:+OP] on the adjacent head C, i.e. the subject pronoun shows contextual allomorphy, cf. the VIs in (208):

(208)   *VIs for the 3rd person subject pronoun*:

    a.   /wò/ ↔ [D, 3SG, NOM] / ___ C [OP:+OP]

    b.   /é/ ↔ [D, 3SG, NOM]

If C bears [OP:+OP], due to Agree with an operator in SpecCP, the more specific VI *wò* is inserted in the D head in SpecTP; otherwise the default form for 3rd person singular subject *é* is chosen. This is a legitimate instance of outward-sensitive allomorphy according to the conditions on contextual allomorphy in (i) and (ii) from the literature (cf. e.g. Bobaljik 2000; Carstairs-McCarthy 2001b; Adger et al. 2003; Embick 2010): (i) If insertion applies cyclically from bottom to top and X undergoes insertion before Y, X can only be sensitive to the morphosyntactic features of Y but not to the phonological features of Y because Y has not yet undergone vocabulary insertion, hence, it lacks phonological features. This condition is met in Ewe because the pronoun in SpecTP is realized before C and is sensitive to a morphosyntactic feature on C. (ii) X can only condition allomorphy of Y if X and Y are linearly adjacent. This is also the case in Ewe: The subject pronoun always immediateley follows the complementizer or the Ā-moved XP (in the latter case, the C-head is zero).[101] In addition, this condition explains why only the subject pronoun can change and not, for instance, an object pronoun: The C-head is too far away from the object pronoun to condition allomorphy.

    Let me go through a sample derivation for a pattern II reflex of long wh-movement of the internal argument of a transitive verb across one clause-boundary. In the embedded clause, the wh-phrase moves successive-cyclically through the edge of every XP because the feature [•WH•] on the root C-head in the numeration triggers edge feature insertion on all lower heads. We start at the point of the derivation where the C head of the embedded clause merges with the TP. The 3rd person subject DP is moved to the inner specifier of T; the wh-phrase is in the outer specifier, cf. (209-a). C bears an edge feature and an operator-probe feature. Due to the pattern II order in (207-c), C first triggers upward Agree. However, it does not find a goal because there is no operator in SpecCP yet. The probe feature is deleted by default, cf. (209-b-c). Afterwards, C triggers the intermediate movement step of the wh-phrase to SpecCP, which comes too late to feed Agree, cf. (209-d). Consequently, the context for the insertion of the more specific pronoun *wò* is not met, only the default form can be inserted.

---

[101]In embedded questions, the Ā-moved XP precedes the subject pronoun but it can follow the complementizer. This suggests a CP-recursion structure: The Ā-moved phrase lands in the specifier of a lower CP whose head is not phonologically realized, but it is adjacent to T and thus local enough to trigger allomorphy. There is another CP-shell above this CP; the head of the additional CP-shell is overtly realized.

(209)   *Wh-movement in the embedded clause*:

    a. C merges with TP

$$
\begin{array}{c}
\text{C}' \\
\diagdown \diagup \\
\text{C} \qquad\qquad \text{TP} \\
\begin{bmatrix} [*\text{OP}:\square *] \\ [\bullet\text{EF}\bullet] \end{bmatrix} \qquad \text{wh-XP}_{[\text{OP}:+\text{OP}]}, \text{DP}_{[3\text{SG}]} \cdots
\end{array}
$$

    b. Step 1: Agree is triggered but valuation fails

❶ Agree

$$
\begin{array}{c}
\text{C}' \\
\text{C} \qquad\qquad \text{TP} \\
\begin{bmatrix} [*\text{OP}:\square *] \\ [\bullet\text{EF}\bullet] \end{bmatrix} \qquad \text{wh-XP}_{[\text{OP}:+\text{OP}]}, \text{DP}_{[3\text{SG}]} \cdots
\end{array}
$$

    c. Result: default deletion of the probe

$$
\begin{array}{c}
\text{C}' \\
\text{C} \qquad\qquad \text{TP} \\
\begin{bmatrix} \cancel{[*\text{OP}:\square *]} \\ [\bullet\text{EF}\bullet] \end{bmatrix} \qquad \text{wh-XP}_{[\text{OP}:+\text{OP}]}, \text{DP}_{[3\text{SG}]} \cdots
\end{array}
$$

    d. Step 2: movement of the wh-phrase, [•EF•] is discharged

$$
\begin{array}{c}
\text{CP} \\
\text{wh-XP} \qquad\qquad \text{C}' \\
{[\text{OP}:+\text{OP}]} \qquad \text{C} \qquad \text{TP} \\
\qquad \cancel{[\bullet\text{EF}\bullet]} \qquad t_{XP}, \text{DP}_{[3\text{SG}]} \cdots
\end{array}
$$

❷ Merge

    e. Postsynstactic insertion: DP$_{[3\text{SG}]}$ realized by *é*

The wh-phrase moves on to the matrix clause from phrase to phrase until the matrix C head is merged, this is the stage shown in (210-a). C bears a [•WH•]-feature and an operator-probe, but this time, it triggers Merge of the wh-phrase first because it is a final movement step, cf. (210-b). Afterwards, the probe on C initiates Agree and finds the wh-phrase (which bears [OP:+OP]) in SpecCP. Consequently, the probe is valued to [+OP], cf. (210-c-d). If the matrix clause has a 3rd person subject pronoun, the presence of this feature triggers insertion of the more specific VI *wò*.

(210) *Wh-movement in the matrix clause*:

a. C merges with TP
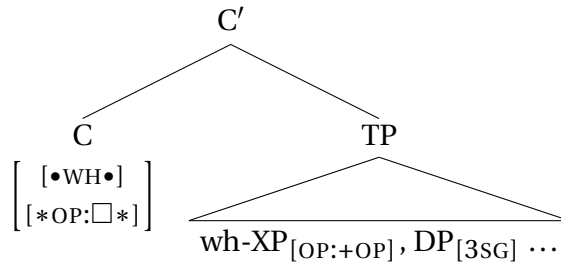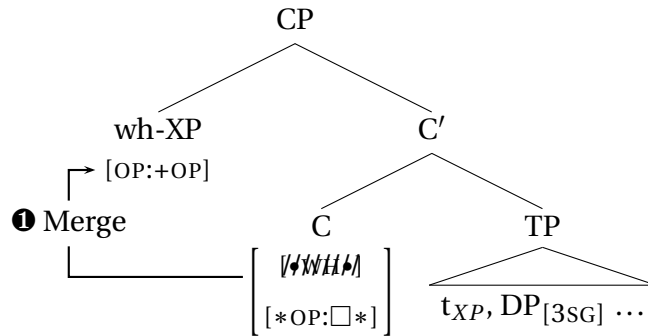
```
                        C′
              ┌─────────┴─────────┐
              C                   TP
        ⎡ [•WH•]    ⎤        ┌─────┴─────┐
        ⎣ [∗OP:□∗] ⎦    wh-XP[OP:+OP], DP[3SG] …
```

b. Step 1: movement of the wh-phrase, [•WH•] is discharged

```
                           CP
                 ┌─────────┴─────────┐
              wh-XP                  C′
            [OP:+OP]           ┌─────┴─────┐
      ┌─►                      C           TP
   ❶ Merge              ⎡ |/|WH|/|| ⎤   ┌───┴───┐
      └──────────────── ⎣ [∗OP:□∗]  ⎦  t_XP, DP[3SG] …
```

c. Step 2: Agree triggered by C

```
                           CP
                 ┌─────────┴─────────┐
              wh-XP                  C′
            [OP:+OP]           ┌─────┴─────┐
              ▲                C           TP
              │           [∗OP:□∗]   ┌─────┴─────┐
      └─ ❷ Agree ─┘          t_XP, DP[3SG] …
```

d. Result: probe valued

```
                           CP
                 ┌─────────┴─────────┐
              wh-XP                  C′
            [OP:+OP]           ┌─────┴─────┐
                               C           TP
                           [OP:+OP]   ┌─────┴─────┐
                                   t_XP, DP[3SG] …
```
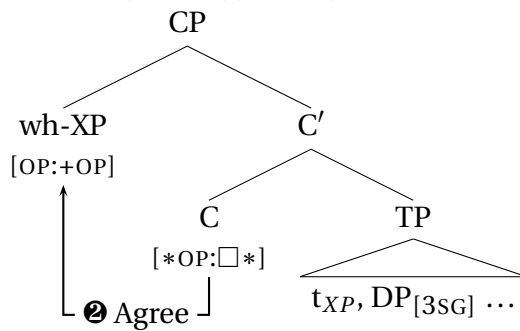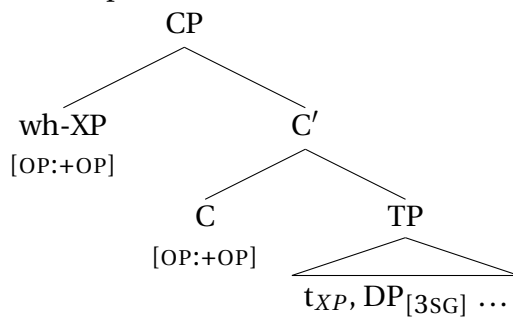
e. Postsynstactic insertion: DP[3SG] realized by *wò*

Hence, there is a reflex of wh-movement solely in the clause where the wh-phrase lands, viz. we have derived a pattern II reflex.

2.6.3.2 Complementizer selection in Irish

In Irish, the complementizer *go* is replaced by a special form if an operator (argument or adjunct) undergoes Ā-movement to SpecCP. If the operator is moved to SpecCP, the complementizer surfaces as *aL*. That movement is involved is evident from the fact that the dependency is island-sensitive and exhibits cross-over effects. *aL*-marking is a pattern I reflex: The change occurs in all clauses crossed by movement. If, however, the operator is base-merged in SpecCP, as is evident from the fact that the dependency is not island-sensitive, does not exhibit cross-over effects and requires a resumptive pronoun in the theta-position, the complementizer adjacent to the operator surfaces as *aN*; all other complementizers between the resumptive pronoun and the complementizer that projects the final landing site of the operator surface in the default form *go*. As for Ewe, I assume that C-heads bear an upward-probing [∗OP:□∗]-feature. Given the ordering statement in (211), final and intermediate movement steps to SpecCP apply early enough to feed Agree with the probe on C (I will come to the feature [•X•] below, it can be ignored for the moment):

(211)   *Ordering statement in Irish*:
        [•X•], [•F•], [•EF•] > [∗OP:□∗]

If the XP merged in SpecCP is an operator, bearing the feature [OP:+OP], it values the probe on C to [OP:+OP]. This is illustrated for the final movement step of a wh-phrase in (212). The derivation starts at the point where C is merged with TP with the wh-phrase in SpecTP.

(212)   *Operator movement feeds Agree in Irish*:
        a.   C merges with TP
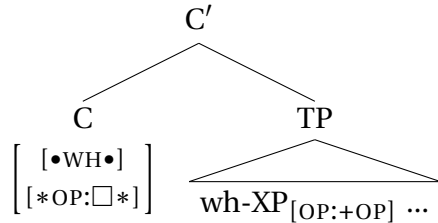
        b.   Step 1: movement of the wh-phrase, [•WH•] is discharged
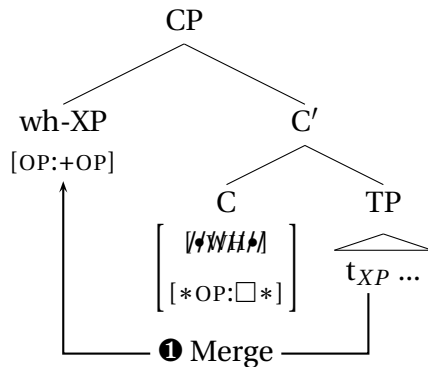
        c.   Step 2: C triggers Agree
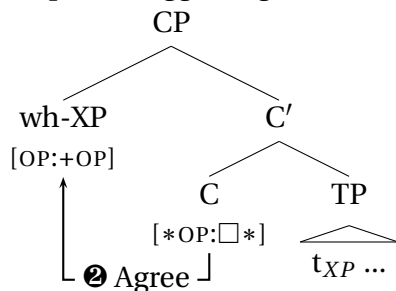
d.   Result: probe valued
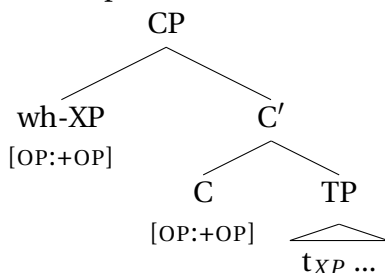
The derivation proceeds in the same way if the operator undergoes an intermediate movement step to SpecCP, the only difference is that C bears an edge feature instead of [•WH•]. But since intermediate movement steps also apply before Agree (cf. (211)) they also feed Agree with C.

I propose that the spell-out of the complementizer is sensitive to the feature [OP:+OP], cf. (213). The default form of the complementizer is *go*, but if C bears [OP:+OP], the more specific VIs *aL* or *aN* are inserted:

(213)   *VIs in Irish*:

a.   /go/ ↔ [C]
b.   /aL/ ↔ [C { [OP:+OP] }]
c.   /aN/ ↔ [C { [•X•], [OP:+OP] }]

What is the difference between *aL* and *aN*? Following McCloskey (2002), I assume that the form of the complementizer depends on the features on the C head that trigger Merge.[102] There are three different C heads in Irish: (i) C without a Merge trigger, (ii) C with an external Merge trigger, (iii) C with an internal Merge trigger (bearing either [•F•] or [•EF•]). C head (i) will be realized as *go* because the operator probe cannot be valued – there is no operator movement to SpecCP that could feed upward Agree. I assume that the probe is deleted by default in this case. C-head (ii) triggers external Merge of an operator (resulting in an *aN*-chain). I will argue in chapter 4 that external and internal Merge must be distinguished because they can apply at different points of the derivation, just like final and intermediate movement steps. In the present system, this means that they must have distinct triggers. Let us call the external Merge trigger [•X•] for the present purposes (it cannot simply be a [•D•]-feature because the operator can be of any category). If external Merge applies before Agree, as all other Merge operations (cf. the order in (211)), it feeds operator Agree. The complementizer will be realized as *aN* because the VI in (213-c) is sensitive to the fact that C has a discharged external Merge trigger. If a type (iii) com-

---

[102]McCloskey (2002) convincingly argues that the difference between *aL* and *aN* cannot be due to a property of the operator itself, as previously proposed in McCloskey 1990. Evidence comes from mixed chains where different types of complementizers appear in a chain (the phenomenon is also known as *chain hybridization*, see e.g. Huybregts 2009; Assmann et al. 2010). An abstract example of a mixed chain is given in (i) (*pro* is a resumptive pronoun):

(i)   *Mixed chains in Irish (McCloskey 2002: 198)*:
        [_CP_ aL … [_CP_ aN … pro]]

aN surfaces in the embedded clause and *aL* in the matrix clause. *aN* shows that an operator is base-merged in the embedded SpecCP. But in the matrix clause, there is operator movement, as signalled by *aL*. Hence, there is the same operator in both the matrix and the embedded SpecCP position, but it triggers different forms of the complementizer. As McCloskey (2002: 199) concludes, the choice bewteen *aL* and *aN* can thus not be due to a property of the operator; rather, it is a consequence of the features on C, viz. whether C triggers internal or external Merge.

plementizer is merged, its probe-feature will also be valued to [OP:+OP] by the operator that has been internally merged in SpecCP in a previous step of the derivation. However, since it bears an internal Merge trigger that is distinct from the external Merge trigger, the VI *aN* cannot be inserted. The most specific matching VI is the one in (213-b), *aL*. We can thus derive the difference between *aL* and *aN* by making reference to the external / internal Merge distinction in the VIs that realize the complementizer. This presupposes that discharged features are only deactivated for the syntax, but are visible for postsyntactic vocabulary insertion. That *aL*-chains exhibit pattern I follows from the fact that they involve movement which applies successive-cyclically through the specifier of every CP (and other XPs); since final and intermediate movement steps both apply before Agree, the operator-probe on every C head along the path of movement will be valued and consequently spelled out as *aL*. Descriptively, *aN*-chains follow pattern II. However, this is not the result of a different timing of intermediate movement steps; rather, this follows since there is no successive-cyclic movement through intermediate SpecCP positions in the first place since the operator is base-merged in its surface position. Thus, it can only feed Agree with the adjacent C-head.

2.6.3.3 Wh-agreement in Chamorro

In Chamorro, extraction of an XP results in wh-agreement on all verbs along the path of movement (pattern I). The morphosyntactic changes and the context in which they occur are summarized in the table in (214).

(214)   *Wh-agreement exponents (Chung 1998: 236)*:

case/GF of extracted XP	exponent
nominative	- *um*-infix
	(if the predicate is realis and transitive)
objective	optional nominalization + -*in*-infix
	(infix occurs if the predicate is transitive)
oblique	nominalization + optional -*in*- infix
	(infix occurs if the predicate is unaccusative)

Crucially, wh-agreement replaces the regular subject-verb-agreement morphology on the verb. If the context for the occurrence of the wh-agreement morphology is not met, the there is no replacement, i.e. the regular agreement markers show up.

Let us first concentrate on local extraction. A widely adopted assumption in the literature on wh-agreement in Chamorro (and extraction morphology in related Austronesian languages) is that it is the result of case agreement between the head that exhibits subject-verb-agreement and the moving XP (see e.g. Chung 1982; 1991; 1998; Georgopoulos 1991; 1985; Watanabe 1996; Rackowski 2005; Rackowski and Richards 2005). I adopt this assumption and connect it with the approach to wh-agreement in Lahne (2008*b*).

Lahne (2008*b*) develops a very detailed and explicit analysis of wh-agreement morphology in Chamorro.[103] The basic idea is the following: Since wh-agreement replaces the regular subject-verb-agreement morphology on a head H, movement to SpecHP must

---

[103]Chung (1998) does not provide an explicit analysis of the morphology-syntax-interface in Chamorro. She says that the verb (or a head in the extended projection of the verb) agrees in case with the trace of the moved XP, but she does not provide morphological realization rules; hence, it remains unclear how the regular subject-verb-morphology is replaced and how the respective morphological changes arise (infixation vs. nominalization).

lead to deletion of the corresponding person or number features of the subject on H. Given a postsyntactic realizational morphology and the Specificity Principle, the consequence is that after extraction, the regular agreement morphology can no longer be inserted and a less specific marker is used, viz. *um* or *in*, which she identifies as voice markers. Lahne (2008*b*: 59f.) models deletion as follows: Intermediate movement steps are triggered by structure-building features [•F•]. However, there are no designated features (like the edge feature in the present account) that trigger movemen; rather, the edge property [• •] is added onto a feature that the head already bears. The edge property turns this feature into a structure-building feature. It triggers internal Merge and is discharged afterwards; as a consequence, the feature that was assigned the edge property is deleted. Thus, a head whose specifier is targeted by movement is impoverished in the syntax (= *probe impoverishment*).

Since I also use deletion to account for reflexes of movement (cf. section 2.6.2 on downstep deletion in Kikuyu and *men*-deletion in Indonesian), the analysis developed by Lahne (2008*b*) for Chamorro can easily be translated into the present system that uses addition. For concreteness, I will combine Lahne's (2008*b*) approach and the insight that wh-agreement involves case agreement: The XP moved to SpecHP enters into case-Agree with the head H (which initially leads to addition of feature values on H). The case value acquired by H feeds an impoverishment rule that deletes a different feature on the same head in the context of this case value. As a consequence, the VI which would have realized this feature can no longer be inserted postsyntactically. Only the less specific voice markers *um* and *in* match the impoverished feature set on H (nominalization also involves impoverishment, but deletion does not result in insertion of a different VI, see below). The features that are deleted by an impoverishment rule in my reanalysis are exactly those that receive the edge property in Lahne (2008*b*).

Let me illustrate the analysis on the basis of some case studies. If the external argument is extracted from a transitive realis predicate, the marker -*um*- is used instead of the regular subject-verb agreement morphology. According to Lahne (2008*b*: 67ff.), this is an effect on the T-head, i.e. it is triggered by movement to SpecTP. The reason for this assumption is the following: Chamorro is a VSO language. According to Chung and McCloskey (1987); Chung (1998), the verb-initial order is derived by V-tov-to-T-movement. Due to this movement, the ϕ-probe on v that initiates agreement with the subject, ends up on the T-head. Consequently, movement to SpecTP can influence the subject's ϕ-features on T.

When the external argument of a transitive verb is extracted – assume that it is wh-moved – it must make a stop-over in SpecTP. To trigger this intermediate movement step, Lahne proposes that the edge property [• •] targets the number feature in T in case the verb is transitive realis; if it is irrealis, the edge property targets another feature on T, viz. the voice feature. Hence, the number or the voice feature becomes a structure-building feature that triggers the intermediate movement step of DP_{ext} to SpecTP. The insertion rules are given in (215). Note that [ARG-ARG] is the representation for transitivity (two arguments are present in the vP); this information is encoded on v but is present on T due to head movement of v to T. Note that the context of the rules is sensitive to the category V; again, this category is visible on T due to head movement.

(215) *Edge property insertion rules for T (Lahne 2008*b*: 69)*:

    a.   [VOICE:−AG] > [•VOICE:−AG•] / ___ [CAT:V +IRR ARG-ARG]

    b.   [NUM] > [•NUM•] / ___ [CAT:V −IRR ARG-ARG]

As a result, the feature with the edge property is deleted on T after it has triggered movement. Some of the relevant VIs Lahne postulates are given in (216):

(216)   *VIs in Chamorro (Lahne 2008*b*: 64, 66)*:

     a.   /hu-/ $\leftrightarrow$ [+1 −2 +SG] / ___ [CAT:V]

     b.   /un-/ $\leftrightarrow$ [−1 +2 +SG] / ___ [CAT:V]

     c.   /ha-/ $\leftrightarrow$ [−1 −2 +SG −IRR] / ___ [CAT:V]

     d.   /-um-/ $\leftrightarrow$ [VOICE: +AG]

The regular subject-verb-agreement markers are all specified for person and number. Since number is deleted in the transitive realis under extraction, cf. (215-b), these markers cannot be inserted anymore. Instead, the less specific matching voice marker *-um-* is chosen. In this way, movement of the external argument of a transitive verb replaces subject-verb-agreement markers by special wh-agreement morphology (here, the voice marker). In the transitive irrealis, the voice feature [VOICE:−AG] is deleted on T. Hence, the regular agreement markers can still be inserted because they are more specific than the voice marker *-um-*.

    The reanalysis works as follows: An edge feature [•EF•] is inserted on T due to *Phase Balance* if there is a C head with a [•WH•]-feature in the numeration. DP_{ext} moves to T and T agrees upwards in case with the argument (which presupposes that DP_{ext} already bears a case value at this point; this may be so either because T assigns case before it attracts DP_{ext} or because DP_{ext} is assigned case by a lower head, e.g. v). If movement is to feed upward Agree, it needs to apply before Agree:

(217)   *Part of the ordering statement in Chamorro*:
     [•F•], [•EF•] $\succ$ [∗CASE:□∗]

As a result, T bears the case value nominative. Afterwards, the following impoverishment rule is fed, based on Lahne's rules in (216):

(218)   *Impoverishment rules for T in Chamorro*:

     a.   [VOICE:−AG] $\rightarrow$ Ø / ___ [CAT:V +IRR ARG-ARG CASE: NOM]

     b.   [NUM] $\rightarrow$ Ø / ___ [CAT:V −IRR ARG-ARG CASE: NOM]

If the verb is realis, the number feature is deleted; if it is irrealis, the voice feature is deleted. The consequences for vocabulary insertion are exactly as in Lahne's analysis: Deletion of the number feature blocks insertion of the regular agreement markers and leads to insertion of the less specific voice markers. If an element with a different grammatical function moves to SpecTP, it will value a different case on T. As a consequence, the impoverishment rules in (218) will not be fed, and the voice marker *-um-* is blocked if the verb is in realis mood.

    Let me now turn to nominalizations. Nominalization occurs obligatorily with the extraction of obliques and optionally with the extraction of the direct object of a transitive verb. Lahne (2008*b*: 79) concludes that what these contexts have in common is that the XP moves through SpecVP when it undergoes extraction (obliques are base-generated in the VP). She proposes that this movement causes deletion of the category feature V because the edge property is inserted on the category feature:

(219)   *Edge property insertion rules for V (Lahne 2008*b*: 80)*:
     [CAT:V] > [•CAT:V•]

Since a category value is required for labeling, the category N of the verbal argument is copied onto V a repair mechanism; the lexical verb thereby becomes a noun and and projects an NP. This NP is selected by n instead of v. Since n has different properties than v, this accounts for the case and agreement facts under nominalization (n assigns oblique case instead of accusative, ϕ-features on n are realized by the morphemes from the possessive paradigm).

The reanalysis in my system works as follows: V receives an an edge feature when a VP-internal XP is to be extracted (because there is a structure-building feature on a head in the numeration, e.g. [•WH•] on C). The XP moves to SpecVP and V initiates upwards case Agree. The case value on v feeds the following impoverishment rule:

(220)   *Impoverishment on V in Chamorro (based on Lahne 2008b: 80)*:
        [CAT:V] → Ø / ___ [CASE:α]
        (α is a variable over a case feature for which accusative and oblique form a natural class)

Note that this analysis requires that DP$_{int}$ of a transitive verb bears accusative when it is moved to SpecCP. Therefore, I assume that accusative is assigned by V before V triggers internal Merge. Note furthermore that the obliques that trigger wh-agreement also bear case in Chamorro (the oblique case), they are not PPs. This case must also be assigned early within VP. Deletion of the category feature then has the consequences described by Lahne (repair by N-insertion, projection of NP, selection by n instead of v).[104]

Since obliques are base-generated in the complement domain of VP (Lahne 2008b: 83), they must go through SpecVP when extracted, and hence, this will always lead to nominalization. But why is nominalization optional with the extraction of the internal argument of a transitive verb? Lahne (2008b: 82f.) proposes that this follows from the fact that the external argument of a transitive verb can be base-merged either in SpecvP or in SpecVP. If DP$_{ext}$ is base-merged in SpecVP, DP$_{int}$ is base-merged as the complement of V and hence, it must go through SpecVP (the edge of the phase head V) when it undergoes extraction; this causes nominalization. If, however, DP$_{ext}$ is base-merged in SpecvP, DP$_{int}$ is the only argument merged in the VP, and therefore, it is already at the edge of VP in its base position (the edge of a head X is defined as "the residue outside of X′; it comprises specifiers of X.", cf Lahne (2008b: 13)). Hence, movement to the edge of VP is not necessary. Consequently, deletion is not fed and nominalization does not arise. In my Agree-based reanalysis the inability of DP$_{int}$ to cause category impoverishment on V when it it the sole argument merged in the VP follows because upward Agree requires asymmetric c-command between the goal and the probe; but when V is base-merged as the sister of V and no further argument is merged with V, V and DP$_{int}$ c-command each other, hence, the case probe on V cannot be valued and impoverishment is not fed.

If DP$_{int}$ extraction triggers nominalization, it also causes the occurrence of the wh-agreement marker *-in-*. Lahne (2008b: 84ff.) analyzes this as a reflex on the functional

---

[104]This analysis of nominalization presupposes that impoverishment applies in the syntax and not postsyntactically, as I assumed for Kikuyu and Indonesian. This is because the head that selects the projection of the main verb in the syntax is sensitive to the category of this projection: v selects VP, but after deletion of the category feature and insertion of N, it is n that selects NP. What is crucial in the present system is that impoverishment applies *after* Agree (here, in case) and before vocabulary insertion. For this condition to be met, it is not necessary that impoverishment applies postsyntactically; it can also apply in the same component as Agree, i.e. in the syntax, as long as it is ordered after Agree (cf. Keine 2010). For most cases, it does not make a difference whether impoverishment applies in the syntax or in the post-syntactic component. Chamorro nominalization is a case where it needs to apply in the syntax.

head n. Since all heads are phase heads, extraction of DP_{int} to a higher projection must go through SpecnP. The edge property is inserted on the agent voice feature on n in case the verb, which has been turned into N under nominalization and which moves to n, is transitive ([CAT:N] is present on n due to n-to-N-movement):

(221)   *Edge property insertion rule for n (Lahne 2008*b*: 85)*:
         [VOICE:+AG] > [•VOICE:+AG•] / ___ [CAT:N ARG-ARG]

Hence, extraction of DP_{int} leads to deletion of the agent voice feature under nominalization. The feature [VOICE:+AG] serves as the context for the insertion of a highly specific zero marker, cf. (222-a). But if this feature is deleted under DP_{int}-extraction, the zero marker is not a subset of the features on n anymore and the less specific marker *-in-* is selected, cf. (222-b) (the voice feature [VOICE:–AG] is an inherent feature of V; it is visible on n due to head movement of V=N).

(222)   *More voice VIs in Chamorro (Lahne 2008*b*: 66)*:
         a.   /Ø/ ↔ [VOICE:–AG] / ___ [+AG]
         b.   /-in-/ ↔ [VOICE:–AG]

What we need to do to integrate this deletion analysis into the upward Agree approach is to replace (221) by the impoverishment rule in (223):

(223)   *Impoverishment rule for n in Chamorro (based on Lahne 2008*b*: 85)*:
         [VOICE:+AG] → Ø / ___ [CAT:N, ARG-ARG, CASE:ACC]

Consequently, the agent voice feature is deleted in a transitive context if the XP moved to SpecnP bears accusative (valued onto n by upward Agree with the DP in SpecvP).[105] Therefore, only the VI *in* can be inserted on n.

    The examples may suffice to illustrate the basic idea of the analysis, which heavily borrows the insights in Lahne (2008*b*). I will not go through other contexts. The reader is referred to Lahne (2008*b*: ch. 4.2) for further examples (also for the account of why the *in* can occur with the extraction of the sole argument of an unaccusative verb – this is due to different deletion rules than the one in (223)). It should be straightforward how her analysis can be translated into my system.[106]

    At this point, one might ask whether there is a difference between Lahne's and the present analysis. As pointed out by Zentz (2013), Lahne's (2008*b*) probe impoverishment account for reflexes of successive-cyclic movement is able to handle morphological reflexes that result in deletion (as in Kikuyu, Malay/Indonesian) or replacement (as in Chamorro, Irish) of the regular morphology; but reflexes that lead to *addition* of markers under extraction are problematic. The reason is that the special extraction morphology is always a retreat to the general case in Lahne's system: Movement involves deletion of features which leads to the insertion of a less specific marker:

---

[105]The restriction to accusative case correctly predicts that *in*-infixation does not arise under extraction of an oblique, although it results nominalization.

[106]Dukes (1992) notes that the wh-agreement morphology in Chamorro also occurs in environments where no Ā-extraction has taken place: The subject extraction marker *-um-* also occurs in infinitives and as a nominal actor marker; likewise, the object extraction marker *-in-* occurs in e.g. nominal passives. In contrast to the analysis proposed in a number of works by Chung, Lahne's analysis captures these syncretisms. She argues that there is also movement to the specifiers of V, n, and T, respectively, in the contexts where wh-agreement morphology shows up but where there is no Ā-movement to SpecCP. Hence, we find the same reflexes of movement. This result is retained under my reformulation of her analysis.

(224)   *Generalization (Lahne 2008b: 60)*:
          When a language shows different exponents in movement and non-movement
          contexts, then the marker appearing in the context of movement is less specific
          than the marker appearing in non-movement contexts (= retreat to the general
          case, emergence of the unmarked).

This generalization does not hold for languages where we find an exponent under extraction *in addition* to the regular non-extraction morphology. A way out for Lahne is to assume that in languages with addition reflexes, there is a highly specific zero VI in contexts without extraction. Extraction leads to impoverishment of a feature which the zero VI realizes and consequently, the zero VI is blocked. A less specific overt VI is inserted instead (cf. the analysis of *in*-infixation above). The question is whether the postulation of such zero VIs is plausible and can be motivated in the respective languages. In the present account of reflexes of movement, addition reflexes follow straightforwardly from the fact that movement (in case it applies early enough) feeds Agree, which in turn leads to *addition* of a value on the probing head. The value may have further consequences, for instance, it can feed an insertion or an impoverishment rule which in turn results in a different pattern on the surface (deletion, replacement, syntactic reflex). Hence, the present system is flexible enough to account for all reflex types.

   So far for the account of local extraction. Let me now come to long extraction in Chamorro: Recall that under long extraction, the extracted element determines the choice of the wh-agreement marker solely on the verb of the clause in which it is base-merged. Wh-agreement on all verbs in structurally higher clauses crossed by movement register the grammatical function of the clause from which the element is extracted. Hence, extraction of the subject of an embedded object clause yields nominative wh-agreement (-*um*-) in the embedded clause and accusative wh-agreement (-*in*-) on the matrix verb. Assuming that CPs bear case in Chamorro (Chung 1994; 1998), this pattern has been analyzed as a minimality effect that arises under downward Agree: The element that is to be extracted moves to the specifier of the minimal CP that dominates it because CP is a phase. Both the CP and the extracted element in its specifier bear case; if intervention is determined on the basis of domination (*A-over-A-Principle*, here necessarily relativized to case, see Chomsky 1973; Bresnan 1976 for a relativized version of this principle), the CP is closer to the case-probe in the matrix clause than the element in its specifier (cf. Rackowski and Richards 2005 for a similar phenomenon in Tagalog). Recall that Chamorro does not allow for wh-in-situ, so we cannot test directly test whether Agree applies upwards or downwards. However, I have been assuming so far that wh-agreement involves upward Agree. The reason for this is the evidence from the closely related language Palauan (Austronesian, Palauan islands and Guam). Palauan also exhibits wh-agreement under Ā-movement. The distribution of wh-agreement in Palauan is the same as in Chamorro: In case of long extraction, wh-agreement is determined by the moving XP only in the clause that contains the base position of XP; in all higher clauses, it is the grammatical function of the clause from which the XP is extracted that is registered on the verb (cf. Georgopoulos 1985). In Palauan, movement affects the voice morphology of the verb: Under subject extraction, the verb shows realis mood, and under non-subject extraction it has irrealis mood. That the change in the voice morphology is indeed tied to movement is clear from the fact that the dependencies that trigger wh-agreement are sensitive to adjunct islands and coordination islands, show strong cross-over effects and

license parasitic gaps (cf. Georgopoulos 1991: 111f., 115, 191ff.).[107]

The important difference between Chamorro and Palauan is that Palauan allows wh-words with matrix scope to stay in situ (and also for partial wh-movement). Crucially, there is *no* wh-agreement in this construction, cf. (225) (in the wh-ex-situ cases, these clauses contain an overt resumptive in the base position of the wh-phrase, but the wh-agreement morphology would be the same if there was a 'gap', viz. a zero resumptive):[108]

(225) *Wh-in-situ, partial and full movement in Palauan (Georgopoulos 1991: 93)*:

    a. ke-dilu   [el te-mengiil er   ngak [el mo   meruul er   a **ngerang**]]
       R.2SG-said   C R.3PL-wait PREP me   C R.FUT R.do   PREP   what

    b. ke-dilu   [el te-mengiil er   ngak [el ng-**ngera**$_i$ a bo   kuruul er   ngii$_i$]]
       R.2SG-said   C R.3PL-wait PREP me   C 3SG-what   IRR.FUT IRR.do PREP it

    c. ng-**ngera**$_i$ [a 'om-dilu   [el longiil   er   ngak [el bo   kuruul er
       3SG-what   IRR.2SG-said   C IRR.3PL-wait PREP me   C IRR.FUT IRR.do PREP
       ngii$_i$]]]
       it
       'What did you say that they're waiting for me to do?'

In (225-a), a non-subject is questioned. The wh-word is in situ. If Agree applied downward, we would expect irrealis morphology on the verb, but we see realis morphology, hence there is no wh-agreement. The same holds for the higher verbs: Both of them embed non-subject clauses and should thus show wh-agreement, i.e. irealis morphology, but they do not. (225-b) and (225-c) show that wh-movement, partial and full movement, triggers wh-agreement, manifested by irrealis morphology on all verbs crossed by overt movement of the object of the embedded clause. Following the practice in this thesis, the absence of a reflex of movement in in-situ constructions in Palauan means that Agree ap-

---

[107]It has been called into question whether the dependencies that trigger wh-agreement in Palauan actually involve movement because they are not sensitive to all islands. For some reason, Ā-movement in Palauan shows no sensitivity to complex NP islands, subject islands and wh-islands (Georgopoulos 1991: 80ff.). Let me begin with the first type: Chamorro also shows violations of subject islands, so this is not a crucial difference between the two languages. In fact, is has been observed that sentential subjects are transparent for extraction in a number of VSO languages (e.g. Irish, Niuean; see Chung 1983: 216). A suggestion by Chung (1983: 226f.) is that this is due to the fact that, in contrast to non VSO languages, the subject in VSO languages is properly governed by the predicate at S-structure; in ECP terms, extraction is only prohibited out of non-properly governed positions. Another suggestion is that they are not islands because they are merged VP-internally (see e.g. Müller 2010 for an analysis of the variable barrier status of VP-internal elements). Whatever the explanation of these effects might be, it is a more general effect that does not necessarily tell us that no movement is involved. As for wh-islands, these are weak islands in general, so it might not be surprising that extraction from them is easier in some languages. What is really striking is the possibility to extract out of relative clauses. Whatever the explanation for this effect is, I take the remaining evidence (sensitivity to adjunct / coordination islands, strong cross-over, licensing of parasitic gaps) to indicate that movement is involved.

[108]Ā-dependencies in Palauan can either have a gap or a resumptive pronoun in the base position of the extracted phrase. Interestingly, both construction behave alike with respect to tests for movement, i.e. both are subject to / exempt from the same island constraints, both show strong cross-over and license parasitic gaps, in both cases wh-agreement is triggered. This suggests that the structures are the identical and derived in the same way. Georgopoulos (1991) suggests that both contain pronouns in the base position, i.e. that the 'gaps' are indeed zero resumptive pronouns. Note that the presence of a resumptive pronoun and movement do not exclude each other in general. There have been a number of proposals that make the presence of a resumptive compatible with movement, see e.g. Demirdache (1991); Pesetsky (1998); Aoun et al. (2001); Boeckx (2003); Grohmann (2003); Bianchi (2004); Salzmann (2006). The core ideas have been that (i) the resumptive is the operator in situ that is only moved at LF, (ii) the resumptive is stranded by movement, or (iii) the resumptive is the morphological realization of the trace / copy left by movement.

plies upwards in this language. I assume that the same holds for Chamorro; and even if it did not, we would still face the following problem at least in Palauan: How can a head in the extended projection of the verb in CP_1 in (226) agree in case with CP_2 from which an element is extracted? Only the moving XP surfaces in its final landing site, not the whole CP_2 it is moved from. But when CP_2 does not move (overtly), the higher verb cannot enter into upward Agree with it because CP_2 is in its c-command domain.[109]

(226) $[_{CP_1}$ wh-XP$_k$ $[_{C'_1}$ C $[_{TP}$ ... V $[_{CP_2}$ ___$_k$ C'_2 C ... $[_{VP}$ V ___$_k$ $]]]]]$

So how can upward Agree be reconciled with the facts? Obviously, the CP must undergo movement to feed upward case Agree in Chamorro. Assume, following Rackowski and Richards (2005), that the CP becomes transparent for extraction after it entered into Agree with v. But why does it surface in its base position? Two solutions have been independently proposed in the literature on similar reflexes of movement: (i) CP is extraposed before spell-out (van Urk and Richards 2013); (ii) CP-movement is followed by remnant TP-movement (Noonan 2002). van Urk and Richards (2013) propose for the pattern I reflex in Dinka that under long-distance extraction the whole embedded CP moves, enters into Agree, thereby becomes transparent for subextraction, and is obligatorily extraposed after extraction of an XP from CP (a Duke-of-York derivation). Noonan (2002) argues for Irish that movement of CP_2 to SpecCP$_1$ is followed by leftward remnant movement of TP_2, which resurrects the order of constituents before movement of CP_2. I will not choose between the two options; one of these them is required to account for the wh-agreement pattern under the present upward Agree analysis. But let me mention that there is independent evidence for the availability of CP-extraposition in Chamorro: Chung (1991:88f.) shows that a verb cannot be separated from its argument by an intervening clause, although the postverbal order of arguments is rather free in Chamorro. Instead, the clause must be extraposed (this holds at least for transitive verbs).

2.6.4 Syntactic reflexes

Syntactic reflexes are characterized by word order changes. They are derived as follows: Agree between a head H and XP moved to SpecHP results in valuation of a feature on H. The valued feature provides the context for an insertion rule in syntax. This rule causes the addition of an operation-inducing feature, which in turn triggers a syntactic operation that manifests itself as movement. The insertion of the feature violates *Inclusiveness* (cf. (130)), but this seems unavoidable if word order changes are treated as the *result* of a previous syntactic operation. Syntactic reflexes are thus derived in a similar fashion as deletion reflexes: Agree between H and a moved XP feeds a rule that changes the feature content of H; either a feature on H is deleted postsyntactically or a feature is added in the syntax, with different consequences for PF-realization.

2.6.4.1 Inversion in Spanish

In Spanish, subject-verb-inversion applies in case of Ā-movement of an argument XP. Inversion exhibits pattern I: It occurs in all clauses crossed by overt movement. In a declarative clause, the subject surfaces in SpecTP and the finite verb is in T (the verb moves

---

[109]In return, proponents of the downward Agree analysis must explain why there is no wh-agreement in wh-in-situ constructions. The analysis proposed by Georgopoulos (1991:105) is not compatible with Minimalist assumptions because it relies on the notion of S-structure and refers to abstract elements left by movement.

successive-cyclically from V to T via v). Evidence for the high position of the verb comes from the fact that it precedes auxiliaries, clitics and vP-adjuncts. The subject precedes the finite verb but follows complementizers in embedded clauses. It is thus in SpecTP. Ā-moved constituents precede the subject, viz. they are in SpecCP. It is standardly agreed that inversion involves head movement from T to C; this movement results in the reverse order of the subject and the finite verb if a non-subject is extracted. Hence, inversion is a syntactic reflex of movement to SpecCP on the C head. Recall from section 2.3.1.6 that the inverted verb in embedded clauses follows the declarative complementizer in Spanish, i.e. there is a functional projection above the embedded CP that exhibits the reflex of Ā-movement. To integrate this fact, I assume that there is a CP-recursion structure in embedded clauses, cf. the example in (227) for extraction of the direct object of a transitive verb (verb movements not indicated). For ease of illustration, I distinguish the two CP-layers by an index:

(227)   *Structure of an embedded clause in Spanish (without Ā-extraction)*:
          $[_{CP_2}$ C$_2$ $[_{CP_1}$ C$_1$ $[_{TP}$ DP$_{ext_k}$ $[_{T'}$ T $[_{vP}$ ___$_k$ v $[_{VP}$ V DP$_{int}$]]]]]]

CP$_1$ is the CP-layer in which inversion applies as a consequence of Ā-movement to SpecCP$_1$. This CP is embedded in another CP, CP$_2$, whose head hosts the complementizer in embedded clauses.[110] I assume that CP$_2$ is absent in matrix clauses, but nothing hinges on this assumption with respect to inversion.

Inversion is triggered as follows: When an argument XP undergoes movement to SpecCP$_1$, it feeds upward Agree triggered by C. When C successfully agrees with the XP in SpecCP, the probe feature on C$_1$ is valued. The valued feature provides the context for a rule that leads to the insertion of an operation-inducing feature. This feature in turn triggers head movement of T to C. The first question is which feature is involved in Agree. I propose that it is a case feature because this derives the argument-adjunct asymmetry of inversion: Since argument XPs have a case value but adjuncts do not, only Ā-moved arguments can value the probe [∗CASE:□∗] on C$_1$ and thus trigger inversion. Since the case probe is on C, every argument can value the probe: An argument will bear a case value at the point of the derivation where C is merged; case is assigned by the structurally lower heads v and T, respectively.[111] If an adjunct moves to SpecCP, the case probe on C$_1$ cannot be valued since adjuncts do not bear case in the first place. Because the case-probe searches upwards, it can also not be valued if there is no Ā-movement at all: There is no XP in its probing domain SpecCP$_1$ in this context. Consequently, inversion cannot apply unless there is Ā-movement of an argument XP. I assume that a case probe that cannot find a goal undergoes default deletion in Spanish, viz. it does not get a default value. Crucially, Ā-movement to SpecCP$_1$ must apply before Agree in order to be able to feed Agree. Since inversion is a pattern I reflex, both final and intermediate movement steps must apply early, cf. (228):

---

[110]Alternatively, one can replace CP-recursion with a split CP à la Rizzi (1997) such that there are two different projections in the CP-domain. Belfast English also exhibits inversion under Ā-movement. But in contrast to Spanish, inversion is in complementary distribution with the complementizer in embedded clauses. Thus, there is no CP-recursion in Belfast English.

[111]This is in contrast to case-Agree in Indonesian: Due to the position of the case-probe on v, only internal arguments of a transitive verb will bear a case value when they move to SpecvP and be able to value the probe, but subjects do not, cf. section 2.6.2.2. Hence, the position of the case probe determines whether there is a subject-object-asymmetry in addition to an argument-adjunct-asymmetry with respect to reflexes of movement.

(228) *Part of the ordering statement*:
[•F•], [•EF•] ≻ [∗CASE:□∗]

If the case probe on C_1 is valued, it feeds the following insertion rule:

(229) *Insertion rule*:
Ø → [∗T:□∗] / ___ C_1 { [CASE:α] }
(α is a variable over case values)

The effect of this rule is that a probe feature is inserted on C_1. This probe feature will ultimately lead to head movement. How can a probe feature trigger movement (so far, internal Merge is triggered by structure-building features)? For concreteness, I adopt the approach to head movement by Roberts (2010). According to this analysis, head movement is reanalysed as Agree: A head H_1 agrees downwards in all features with another head H_2; afterwards, H_1 contains the features of H_2. Consequently, there are two identical occurrences of H_2 (H_2 itself and H_2 on the probing head H_1), just as there are several occurrences of a moved XP under the copy theory of movement. At PF, it is decided which occurrence is realized. If the topmost occurrence is spelled out, viz. the one on the probing head H_1, we get a movement effect: H_2 is realized in a structurally higher position than its base-merge position. In Spanish, the probe feature is inserted on C_1 and probes downward for the features on T, which contains the finite verb (at least the verbal root, inflectional features are realized postsyntactically). At PF, the topmost occurrence of T on C_1 is realized, resulting in spell-out of the finite verb in C.

The following derivation illustrates how inversion is triggered in an embedded clause that does not contain the final landing site of the moved phrase. Assume the internal argument of a transitive verb undergoes wh-movement. We start with the stage of the derivation at which C_1 merges with the TP, cf. (230-a). The wh-phrase has moved successive-cyclically through the edge of every phrase to SpecTP (triggered by edge features that are inserted because [•WH•] is still in the numeration). Since Spanish exhibits nominative-accusative alignment of case, DP_{int} bears accusative case which it has been assigned by v at an earlier stage of the derivation. C_1 also bears an edge feature, because it triggers an intermediate movement step. In addition, C_1 bears a case probe. Due to the order of operations in (228), the intermediate movement step of the wh-phrase is triggered first, cf. (230-b). Afterwards, C_1 initiates upward Agree and is valued by the accusative case on the wh-phrase, cf. (230-c-d). This valued feature feeds the insertion of the T-probe, cf. (230-e), which in turn triggers Agree with T, cf. (230-f). Afterwards, C_2 merges with CP_1 and the wh-phrase undergoes successive-cyclic movement to its final landing site. Finally, the higher occurrence of the T head on C_1 is realized in the postsyntactic component, resulting in inversion.

(230) *Wh-movement of the internal argument of a transitive verb, embedded clause*:
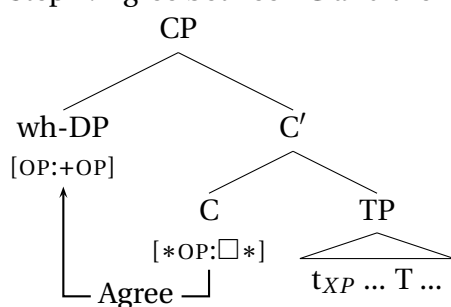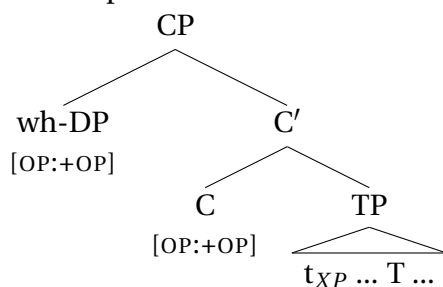a. C_1 merges with TP
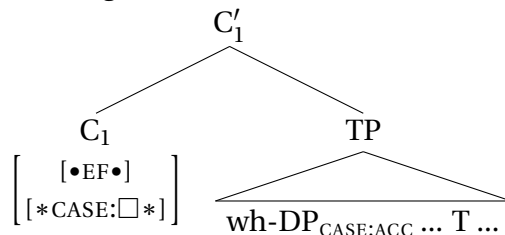
b.   Step 1: internal Merge of the wh-phrase, [•EF•] is discharged
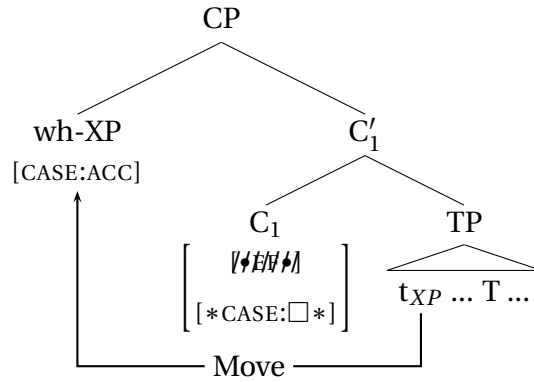
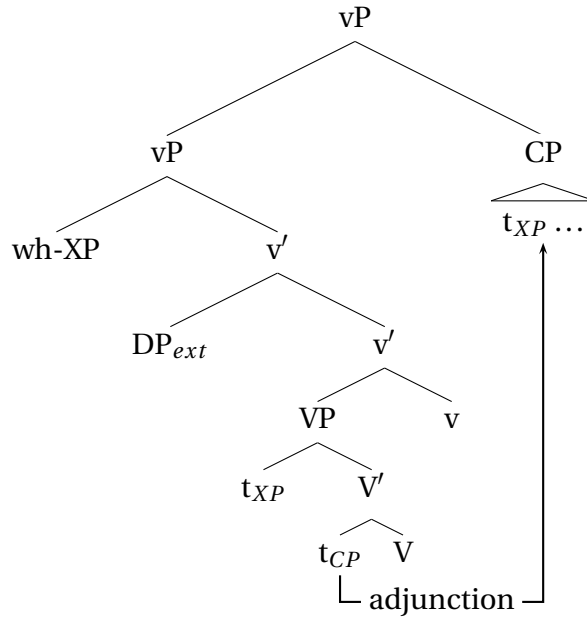c.   Step 2: case-Agree between C_1 and the wh-phrase
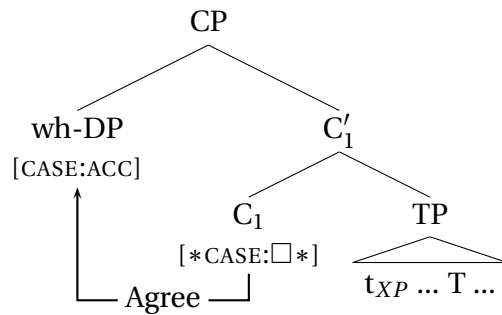
d.   Result: case probe valued
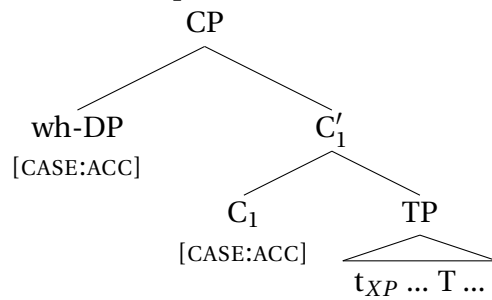
e.   [∗T:□∗] is inserted according to (229)
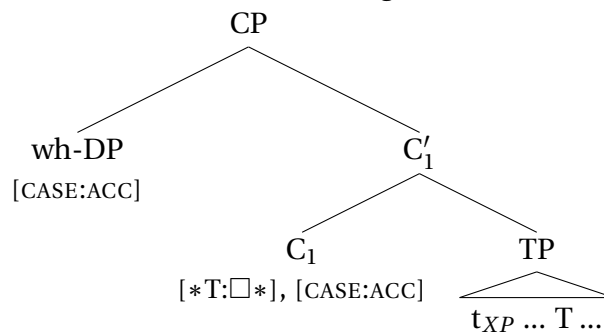
f.   Step 3: [∗T:□∗] triggers Agree with T

CP
wh-DP [CASE:ACC] C'_1
C_1 [∗T:□∗], [CASE:ACC]   TP t_{XP} ... T ...
Agree
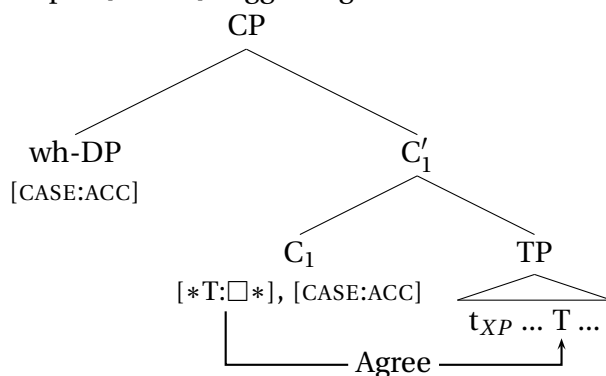
g.   Result: T-probe valued ([T] =is an abbreviation for all features of the T head)

CP
wh-DP [CASE:ACC] C'_1
C_1 [T], [CASE:ACC]   TP t_{XP} ... T ...
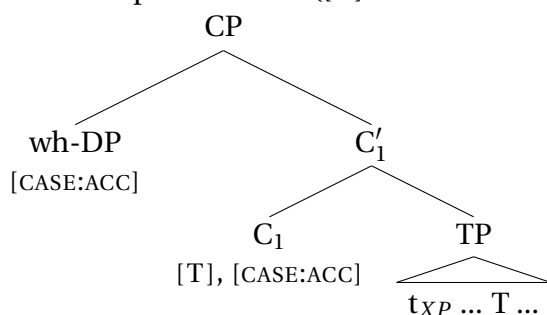
h.   Step 4: Merge of C_2 with CP_1
i.   PF: T realized in the position of C_1

The C_1-head that projects the final landing site bears a [•WH•]-feature instead of an edge feature. But since the [•WH•]-feature is also discharged before case-Agree (cf. (228)), the derivation is exactly the same as the one in (230).

    A few remarks about this analysis of inversion are in order: First of all, note that head-movement, reanalysed as Agree, is not counter-cyclic although it applies after movement to SpecCP. This is because I assume that every phrase constitutes a cyclic domain (see section 3.3.2 for evidence from feeding relations under downward Agree): The Agree relation between C_1 and T applies within CP_1, just like the previous movement operation of XP to $SpecCP_1$. Second, head movement necessarily involves downward Agree, but the initial Agree relation between C_1 and the Ā-moved XP (which triggers subsequent operations that ultimately lead to inversion) is upwards, as argued for extensively in section 2.3.1.6. Hence, I need to assume that the direction of Agree is not fixed for a language as a whole, but rather that it is a property of a single probe. The case probe on C_1 looks upwards but the T-probe searches downwards. I do not see any reason why the direction of Agree should be the same for all probes in a language; it may, but it does not have to. Moreover, I prefer the analysis of head-movement as downward Agree to alternative approaches to head-movement because it is the only one that is consistent with the assumptions of the present approach. If head-movement resulted in a head-adjunction structure [$_{C_1}$ T C_1] (cf. Baker 1988), it would probably be triggered by a structure-building feature on C that causes internal Merge of T. Recall that I assume that internal Merge triggers must be asymmetrically c-commanded by an element with a matching feature in order to be discharged (cf. section 2.5.2.3, footnote 74). But under head adjunction, the two heads c-command each other (this holds at least for one of the segments of the head that is the target of adjunction); consequently, the internal Merge trigger could not be discharged by head movement. To circumvent this problem, one might adopt the approach by Matushansky (2006) who assumes that a moved head targets the innermost

specifier of the head that attracts it (the heads then undergo a special merger opera-
tion in the postsyntactic component that results in a head adjunction structure). As a
result, the moved head asymmetrically c-commands the head that attracts it. However,
this approach violates the *Extension Condition*: Since inversion is a reflex of movement to
SpecCP, Ā-movement applies before head movement from T to C_1. But since head move-
ment must target the innermost specifier of C_1 in Matushansky's (2006) approach, the
moved head would have to tuck-in below the Ā-moved XP, contrary to what the *Extension
Condition* demands. The Agree-approach to head movement does not encounter these
problems.[112]

Note finally that movement of a relative operator is the only Ā-dependency that does
not trigger inversion in Spanish. In section 4.2 I will argue on the basis of such asymme-
tries between Ā-dependencies that we need to distinguish between different types of fi-
nal and intermediate Merge triggers, viz. between structure-building features that trigger
wh-movement, movement of a relative operator, etc. The designated features that trigger
movement of a relative operator would be ordered after the case probe in the ordering
statement for Spanish in (228); all other triggers of Ā-movement are ordered before the
case probe. As a consequence, movement of a relative operator comes too late to feed
case Agree (counter-feeding) and thus to trigger inversion.

2.6.4.2   Extraposition in German

In German, a CP that contains an intermediate landing site of a wh-phrase in SpecCP
obligatorily extraposes; but if the clause contains the final landing site, extraposition is
optional. Thus, obligatory extraposition is a pattern III reflex of movement. The core
idea of the analysis is the same as for inversion: Movement of a wh-phrase to SpecCP can
in principle feed upward Agree between C and the wh-phrase. The value on C acquired
through Agree causes the insertion of an operation-inducing feature that in turn triggers
extraposition. I assume that C bears an operator probe [∗OP:□∗]; wh-phrases bear the
corresponding valued feature [+OP]. Given upward probing and the order of operations in
(231), it follows that only intermediate movement steps feed Agree in the operator feature
and thus extraposition; final movement steps to SpecCP come too late to feed Agree.

(231)   *Ordering statement*:
        [•EF•] ≻ [∗OP:□∗] ≻ [•F•]

If the probe on C is valued to [OP:+OP], the rule in (232) is fed:

(232)   *Insertion rule*:
        ∅ → [▷CAT:v◁] / ___ C { [OP:+OP] }

It triggers insertion of the feature [▷CAT:v◁] on C. This feature triggers extraposition of
the CP, which I take to be right-adjunction to the next higher vP because it is possible in
German to topicalize the main verb together with an extraposed clause, suggesting that
they form a constituent:

---

[112]For alternative analyses of inversion see Rizzi and Roberts (1989) on French as well as Henry (1995) and
Lahne (2008*b*) on Belfast English.

(233) *Topicalization of main verb + extraposed CP in German:*

[$_{CP}$ [$_{vP}$ überzeugt [$_{CP}$ dass Fritz   Claudia   liebt]$_n$] $_k$ hat Peter
        convinced   that Fritz.NOM Claudia.ACC loves   has Peter.NOM
Maria   nicht __$_n$ __$_k$]
Maria.ACC not
'Peter didn't convince Maria that Fritz loves Claudia.'

The derivation for an intermediate movement step to SpecCP is given in (234). It starts at the point where C merges with TP; the wh-phrase is at the edge of TP.

(234) *An intermediate wh-movement step feeds extraposition:*

   a.   C merges with TP
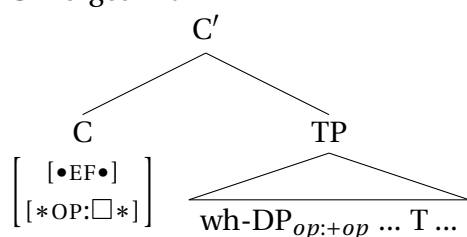
```
                C′
              /    \
            C        TP
      ⎡ [•EF•]   ⎤   /  \
      ⎣ [∗OP:□∗] ⎦  wh-DP_{op:+op} ... T ...
```

   b.   Step 1: internal Merge of the wh-phrase, [•EF•] is discharged

```
                CP
              /    \
          wh-XP      C′
        [OP:+OP]    /  \
           ↑       C     TP
           |   ⎡ [•E̸F̸•̸] ⎤   / \
           |   ⎣ [∗OP:□∗]⎦  t_{XP} ... T ...
           |_____ Move _____|
```

   c.   Step 2: Agree between C and the wh-phrase

```
                CP
              /    \
          wh-DP      C′
        [OP:+OP]    /  \
           ↑       C     TP
           |   [∗OP:□∗]  / \
           |          t_{XP} ... T ...
           |__ Agree __|
```

   d.   Result: probe valued

```
                CP
              /    \
          wh-DP      C′
        [OP:+OP]    /  \
                   C     TP
               [OP:+OP]  / \
                      t_{XP} ... T ...
```
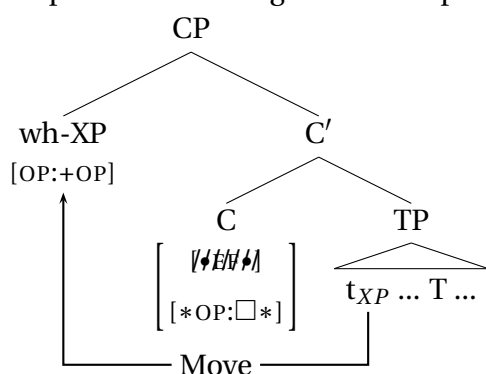
e.   [▷CAT:v◁] is inserted on C according to (232)

```
                        CP
              /                    \
          wh-DP                     C′
          [OP:+OP]            /            \
                             C              TP
                    [▷CAT:v◁], [CASE:ACC]   /‾‾\
                                          t_XP ... T ...
```
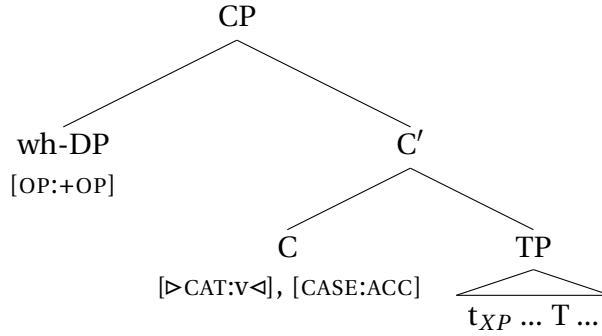
f.   [▷CAT:v◁] triggers extraposition of CP (adjunction to the dominating vP)

```
                          vP
               /                        \
             vP                          CP
         /        \                     /‾‾\
      wh-XP        v′                 t_XP ...
                /      \
            DP_ext      v′
                      /     \
                    VP        v
                  /    \
                t_XP    V′
                      /‾\
                   t_CP   V
                     └── adjunction ──┘
```

How exactly is extraposition triggered by [▷CAT:v◁]? Since the operation-inducing feature is located on the (head of the) phrase that undergoes rightward movement, extraposition is greedy movement: The CP is not attracted by a feature on a c-commanding head, but rather moves to be able to discharge the structure-building feature it bears (default deletion obviously not being an option). By assumption, this feature cannot be checked in situ, thus, extraposition is a Last Resort movement to prevent the crash of the derivation (due to the presence of a non-discharged operation-inducing feature). If [▷CAT:v◁] cannot be discharged in-situ, it must be subject to different locality conditions than the other operation-inducing features [•F•] and [∗F:□∗]. I propose that [▷F◁] can only be checked if it is in a *symmetric* c-command relation with the node that bears the matching feature (the other features require asymmetric c-command). Since the added feature contains the category v, it must symmetrically c-command vP. This is possible if (i) [▷CAT:v◁] projects to CP, (ii) CP adjoins to the structurally higher vP that dominates CP, and (iii) if it is sufficient for symmetric c-command that one segment of vP fulfills this condition (due to adjunction, vP consists of two segments; symmetric c-command holds between CP and the lower vP). Note that projection of features from the head to the phrase is generally required in Greed-based approaches to movement as in Bošković (2007*b*) which assume that movement to SpecHP applies in order to bring the operation-inducing feature in a position from which it c-commands a matching feature on the head H.

The derivation goes as follows after [▷cat:v◁] is inserted on C: [▷CAT:v◁] cannot be discharged in situ and projects to CP; since CP is not the sister of vP [▷CAT:v◁] can still not

be discharged. Structure-building continues until the v-head of the structurally higher clause is merged with the VP that embeds the CP bearing [▷CAT:v◁]. Due to the order of operations in (231), the next step in the v′-projection (after external Merge of the external argument) is the intermediate movement step of the wh-phrase (which has been moved to SpecVP before) to SpecvP. This early movement of the wh-phrase creates a remnant CP. Movement of the wh-phrase leads to projection of v′ to vP. The remnant CP can now undergo adjunction to this structurally higher vP. As a result, CP and (the lower segment of) vP c-command each other and [▷CAT:v◁] can be discharged.[113]

Rightward extraposition differs from leftward movement operations (triggered by [•F•])we have seen so far in several respects: [▷CAT:v◁]-discharge can be minimally delayed, whereas Agree and leftward internal Merge cannot (which is crucial to derive bleeding and counter-feeding), and extraposition is a last resort operation but Agree and Merge are not. However, these properties are due to the greedy nature of extraposition. Leftward movements are based on Attract rather than Greed in the present system. And as far as I can see, it is unavoidable to treat extraposition as greedy movement if it is analysed as a *reflex* of intermediate movement to the specifier of CP: The insertion of the extraposition trigger must happen in the CP in which wh-movement takes place; insertion on another head, e.g. in the numeration, would involve a non-local dependency.[114]

That extraposition does not behave like other internal Merge operations may not be an offensive assumption in light of the fact that empirically, rightward movement patterns differently from leftward movement in a number of ways: Leftward Ā-movement in German is not clause-bound (at least for many speakers), but extraposition is subject to the Right Roof Constraint (Ross 1967; Perlmutter and Soames 1979): It cannot leave the next higher CP, cf. (235-a) vs. (235-b). In addition, extraposition can take place out of islands, e.g. PP-islands, which leftward movement cannot (except for R-pronouns), cf. (235-c) vs. (235-d).

(235)   *Extraposition vs. wh-movement in German*:

    a.   [$_{CP}$ Wen$_k$ denkst du [$_{CP}$ dass Hans glaubt [$_{CP}$ dass Maria weiß [$_{CP}$
        who.ACC think you   that Hans believes   that Maria knows
        dass Peter ___$_k$ liebt]]]]
        that Peter   loves
        'Who do you think that Hans believes that Maria knows that Peter loves?'

---

[113]Recall that extraposition of CPs is optional in German if no wh-movement takes place in the CP. Therefore, I assume that the feature [▷CAT:v◁] can be freely inserted on every C head in the numeration. If it is present, the CP will be extraposed; if it is not present, it cannot be extraposed (as for the root C-head, the derivation can only converge if it does not bear this feature because there will be no higher vP where the root CP can adjoin to and [▷CAT:v◁] would remain unchecked). That extraposition becomes obligatory under (intermediate) wh-movement is a consequence of the insertion rule in (232). The insertion rule is redundant if the C-head through whose specifier a wh-phrase moves already bears a [▷CAT:v◁]-feature, because this feature can be freely inserted onto C in the numeration. For this case, we need a redundancy rule that deletes the inserted [▷CAT:v◁]-feature if C already bears such a feature.

[114]Locality of insertion also rules out an alternative analysis of extraposition that would be compatible with the assumptions about movement made so far in this thesis: The feature that triggers extraposition is inserted on the root C-head in the numeration. This feature could then be a usual structure-building feature that triggers internal Merge of the CP (attract-based). Its presence in the numeration would also enforce successive-cyclic CP-movement because it will be visible for *Phase Balance* in the numeration (but greedy movement can also apply successive-cyclically in order to fulfill the PIC, see Bošković 2007*a*). However, intermediate movement of a wh-phrase to SpecCP would have to trigger insertion of a structure-building feature on a head in the numeration for this analysis to work.

b. *[$_{CP}$ Maria   denkt [$_{CP}$ dass Hans   [$_{DP}$ die   Aufgabe ___$_k$]
    Maria.NOM thinks   that Hans.NOM   the.ACC task.ACC
    vergesen hat [$_{CP}$ weil   er   sehr beschäftigt war [$_{CP}$ die
    forgotten has   because he.NOM very busy   was   which.NOM
    sie   ihm   gegeben hat]$_k$]]]
    she.NOM him.DAT given   has
    'Maria thinks that Hans forgot to fulfill the task which she had given him
    because he was very busy.'

c. dass ich   [$_{PP}$ mit dem Mann ___$_k$] gesprochen habe [$_{CP}$ der
    that I.*nom*   with the man   spoken   have   who.NOM
    gegenüber   wohnt]$_k$
    across.the.street lives
    'that I talked to the.DAT man.DAT who lives across the street'

d. *Wem$_k$ hast du   [$_{PP}$ mit ___$_k$] gesprochen
    who.DAT have you.*nom*   with   talked
    'Whom did you talk to?'

Let me finally compare the present analysis to the approach proposed in Müller (1999). Müller also assumes that wh-movement applies successive-cyclically (at least) through every SpecCP position along the path of movement. All movement steps require checking of a feature. In its scope position, the wh-phrase with the feature [+WH] is checked against the matching [+WH] on the C head. C heads that project an intermediate landing site bear the feature [-WH]. Müller proposes that it is sufficient for checking if the features of the moving phrase and the head to whose specifier it moves match, their values do not have to match, however. In case a wh-phrase bearing [+WH] moves to the specifier of a [-WH] C head, they can enter into a legitimate checking relation because they have the same feature [WH], though with diverging values. This is what Müller calls *imperfect checking*; in the scope position where the wh-phrase and the C head also match in the feature value, there is *perfect checking*: the feature and the value of the C head and the moving phrase are identical. Imperfect checking has (among other things) the consequence that a the [-WH]-feature of the C head, which is not discharged under imperfect checking, becomes a strong feature. Strong features triggers movement operations in syntax (in contrast to weak features which can be checked at LF). More precisely, the [-WH]-feature that became strong due to imperfect checking triggers rightward movement of the CP, pied-pieped by the C head. In Müller's analysis, wh-movement to an intermediate SpecCP also leads to the insertion of a property that enforces movement of the CP. I adopt his basic idea, but since feature strength is no longer assumed to trigger movement in more recent versions of Minimalism, I replace it with the insertion of an operation-inducing feature. Note that the distinction between final and intermediate movement steps in Müller (1999) is also encoded in the trigger of the two types of movement ([+WH] on final vs. [-WH] on intermediate C heads), as in my analysis, although the timing of the movement operations plays no role.

2.6.5   Summary of the analyses

In this section I have shown how addition, deletion, replacement and syntactic reflexes of movement can be derived under the Agree-based ordering approach. Addition reflexes are straightforward: The value acquired by a head H under upward Agree is realized by a VI. As for the other reflex types, I suggested that valuation can feed a (post)syntactic rule, which in turn leads to deletion or insertion of an operation-inducing feature. As a result, we get word order changes or blocking of the VI that would be inserted without movement. The analyses are summarized in table 2.3 (for reasons of clarity, I only include the features required to produce the reflex pattern and leave out additional operation-inducing features that are necessary in some cases to derive asymmetries).

The various asymmetries we have encountered in the languages (argument/adjunct, subject/non-subject, etc.) are the result of (i) which feature is involved in Agree and (ii) the order of operations (external vs. internal Merge). One might prefer a different analysis of the splits, but this does not affect the crucial insight that the patterns of reflexes are the result of the order of the two types of internal Merge and Agree. The language-specific asymmetries are something that comes on top of the orderings and may therefore well be replaced with alternative analyses as long as the order of final and intermediate movement steps as well as Agree is preserved.

language	phenomenon	reflex type	pattern	reflex on	order of features	valuation feeds	asymmetry
Chamorro	wh-agr	morphol. replacem.	I	T, v, V	[•F•], [•EF•] ≻ [∗CASE:□∗]	impoverishment (in syntax)	–
	C-Agr	morphol. addition	II	C	[•F•] ≻ [∗CAT:□∗], [∗LOC:□∗] ≻ [•EF•]	–	–
Irish	C-selection: *aL*-chains	morphol. replacem.	I	C	[•F•], [•EF•] ≻ [∗OP:□∗]	–	–
Wolof	C-agr: *u*-chains	morphol. addition	I / II	C	[•F•] ≻ [∗CLASS:□∗], [∗DEF:□∗] | [•EF•]	–	SU / non-SU
	C-agr: *an*-chains	morphol. addition	III / IV	C	[•EF•] ≻ [∗CLASS:□∗], [∗DEF:□∗] | [•F•]	–	SU / non-SU
Indonesian	*men*-deletion	morphol. deletion	I	v	[•F•], [•EF•] ≻ [∗CASE:□∗]	impoverishment (postsyntactic)	OBJ / non-OBJ
	focus marking	morphol. addition	II	C	[•F•] ≻ [∗OP:□∗] ≻ [•EF•]	–	–
Kikuyu	downstep deletion	phonol. deletion	I	C	[•F•], [•EF•] ≻ [∗OP:□∗]	impoverishment (postsyntactic)	–
Spanish	inversion	syntactic	I	C	[•F•], [•EF•] ≻ [∗CASE:□∗]	insertion (in syntax) (result: head movem.)	ARG / ADJ
Duala	*no*-marking	morphol. addition	II	C (realized on T)	[•F•] ≻ [∗OP:□∗], [∗CAT:□∗] ≻ [•EF•]	–	SU / non-SU
Ewe	pronoun choice	morphol. replacem.	I / II	C (realized in SpecTP)	[•F•] ≻ [∗OP:□∗] | [•EF•]	–	–
Kitharaka	focus marking	morphol. addition	III	C	[•EF•] ≻ [∗FOC:□∗] ≻ [•F•]	–	–
Dinka	*ke*-stranding	morphol. addition	III	v	[•EF•] ≻ [∗#:□∗] ≻ [•D•]	–	–
German	extraposition	syntactic	III	C	[•EF•] ≻ [∗FOC:□∗] ≻ [•F•]	insertion (in syntax) (result: extraposition)	–

Table 2.3: Summary: Analyses of reflexes of movement under upward Agree

2.7   Interim conclusions

In this chapter I have investigated the interactions of Merge and upward Agree that emerge if a single head triggers both operations. The empirical basis for the discussion are languages with reflexes of Ā-movement that solely occur as the result of overt movement. I have shown that there are four different patterns of such reflexes: The reflex occurs in all clauses crossed by movement (PI), only in the final clause (PII), only in intermediate clauses (PIII) or in no clause at all (PIV). If a reflex of movement is the spell-out of a successful Agree relation between a head and the XP moved to its specifier, patterns II and III are opaque: Only some movement steps to SpecHP feed Agree with H, others do not (counter-feeding). Hence, these two patterns strongly suggest a split between final and intermediate movement steps. In pattern II languages only final movement steps feed Agree, and in pattern III languages only intermediate movement steps feed Agree. I have presented empirical arguments against the claim that patterns II and III are not opaque because there is no movement to or no Agree relation with the specifier of the heads that do not exhibit a reflex. The few existing analyses of the variation between patterns I and II that acknowledge the opacity of pattern II are based on enriched representations. I have argued that these analyses have both empirical and conceptual shortcomings. As for the empirical side, some analyses cannot account for pattern III, and none of them is able to capture languages in which patterns I and II co-occur under a single instance of Ā-movement. Conceptually, previous analyses are rather complex because they have to assume different types of empty elements or apply operations and constraints to empty elements; in addition Agree is counter-cyclic.

I have presented a derivational account to opacity that relies on the timing of operations: If a head triggers more than one operation, the operation-inducing features on the head are ordered and are discharged sequentially. Early movement feeds upward Agree, whereas late movement counter-feeds a previous Agree relation. To capture the opaque patterns II and III, final and intermediate movement steps must be distinguished; they need to apply at different points relative to Agree such that only one type of movement feeds Agree while the other counter-feeds Agree. If Merge was a uniform operation, it could either apply before or after Agree, but not both before and after Agree at the same time. I have implemented this difference by postulating designated triggers for each of the two internal Merge operations. The four patterns of reflexes of movement then simply follow from reordering of the internal Merge triggers and the feature that triggers Agree:

(236)   *Orderings of final vs. intermediate internal Merge and upward Agree*

	order of features	interactions		pattern
		final steps	intermediate steps	
a.	[•F•], [•EF•] ≻ [∗F:□∗]	feed Agree		P I
b.	[•F•] ≻ [∗F:□∗] ≻ [•EF•]	feed Agree	counter-feed Agree	P II
c.	[•EF•] ≻ [∗F:□∗] ≻ [•F•]	counter-feed Agree	feed Agree	P III
d.	[∗F:□∗] ≻ [•F•], [•EF•]	counter-feed Agree		P IV

The ordering approach provides a uniform analysis for all four patterns. In addition, it can easily account for languages that mix different patterns: When distinct Agree operations are interleaved with the two types of internal Merge operations, they can be (counter-)fed by different types of Merge. Optionality is derived through partial ordering. Since it is only important *when* movement applies, and not how it is implemented, the anal-

ysis is compatible with all Minimalist approaches to movement. Moreover, the ordering approach dispenses with the additional assumptions of approaches based on enriched representations: There is no reference to (different types of) empty elements and there are no operations that apply on them. Importantly, Agree applies in accordance with the *Strict Cycle Condition*.

Finally, I have presented detailed derivations for some of the reflexes. Thereby, I have illustrated that the present system, though based on the addition of feature values, is flexible enough to derive deletion, replacement, and syntactic reflexes as well because the feature value acquired by a head under upward Agree can feed other (post)syntactic operations that result in deletion or insertion of features.

Chapter 3

Interactions of internal Merge and downward Agree

In this chapter I argue that we also find opaque interaction of internal Merge and Agree triggered by a single head if Agree applies downwards, with the probe c-commanding the goal. Some movement steps bleed or feed downward Agree, whereas others counter-bleed or counter-feed it. Whether we find (counter-)feeding or (counter-)bleeding interactions depends on whether the element that undergoes movement is the potential goal of the downward Agree relation or not. I will present evidence for exactly the same split of movement types into final and intermediate movement steps we have already encountered in chapter 2 for upward Agree. Section 3.1 abstractly illustrates how movement can interact with internal Merge. In 3.2 I introduce two instances of a (counter-)bleeding interaction: the Anti-agreement effect and a movement-related case split in Hungarian. The former involves an interaction on C, and the latter on n, a DP-internal functional head. In section 3.3 I present a (counter-)feeding example from defective intervention in a dialect of Icelandic. Based on cross-linguistic variation in the removal of intervention effects, I argue that Icelandic B involves an opaque interaction of Merge and Agree on the T head. Section 3.4 summarizes the insights of this chapter.

3.1 (Counter-)bleeding under downward Agree: abstract analysis

In the last chapter I investigated the interaction of internal Merge and upward Agree: Movement of an XP to the specifier of HP can feed Agree between H and XP if movement applies before Agree. If movement applies after Agree, valuation fails because the XP is not in the position that H can access (SpecHP) when H probes, viz. we get counter-feeding. In this section I illustrate that we get the reverse pattern under downward Agree where the probe must c-command the goal: Late movement leads to successful valuation, whereas early movement does not. The difference between the languages examined in this chapter and those from the previous chapter is that in the former, Agree between a head H and an XP that originates in H's c-command domain is also possible if the XP is not moved to SpecHP at all. Following Baker (2008*b*) this implies that the probe looks downwards, seeking for a goal in its c-command domain.

Let me begin with an abstract derivation of a bleeding interaction: Internal Merge applies before downward Agree. As before, we look at a stage of the derivation where a head H is merged with WP; WP includes a DP and H triggers both (internal) Merge

and Agree (exemplified with ϕ-Agree), cf. (1-a). Assume that the DP is both the closest matching goal for the probe on H and that it is the phrase attracted to SpecHP. If Merge applies before Agree, H first attracts DP and thereby discharges its structure-building feature, cf. (1-b). Afterwards, it attempts to Agree with a goal in its c-command domain, but it does not find one (assuming that there is no other potential goal within WP), cf. (1-c). DP would have been a suitable goal, but due to prior movement, it is no longer in H's c-command domain when H probes. Thus, early movement of XP bleeds downward Agree between H and DP. Consequently, the probe is discharged by default. This derivation presupposes that the occurrence of DP in its base position (e.g. a copy) is not accessible for the probe on H; see the discussion below.

(1)   *Internal Merge bleeds Agree*:

   a.   Initial structure
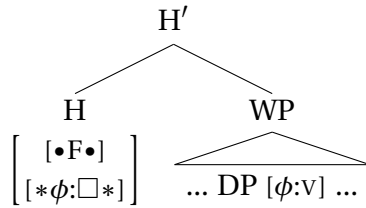
   b.   First step: Merge, [•F•] discharged
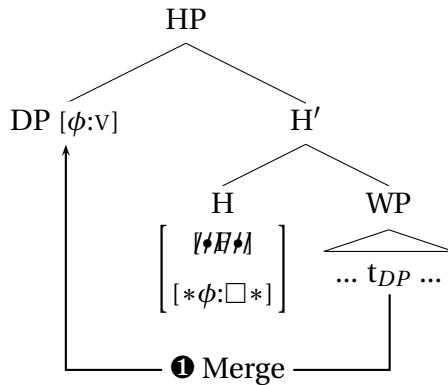
   c.   Second step: Agree, no goal found
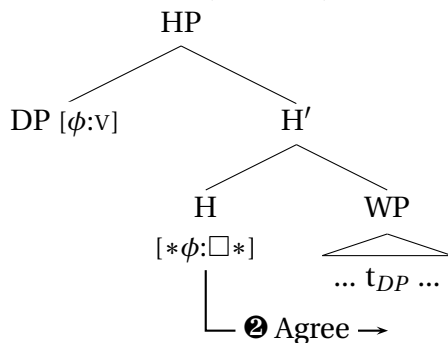
d. Last resort: probe discharged by default

$$HP$$

DP [ϕ:V] H$'$

H WP

~~[∗[ϕ:□]∗]~~

... t$_{DP}$ ...
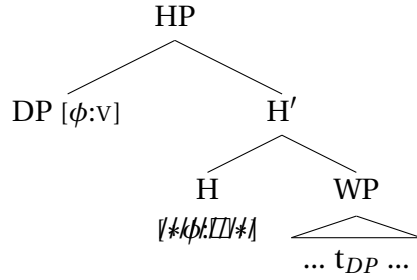
If the operations apply in the reverse order (Agree before Merge) we get counter-bleeding: First, H successfully Agrees with DP because it is still in the c-command domain of H, cf. (2)b-c. Afterwards, DP moves to SpecHP and thereby discharged [•F•] on H, cf. (2-d). On the surface, this is opaque: DP has valued the ϕ-probe on H although it is not in the structural position to be a goal; movement to SpecHP should bleed downward Agree but doesn't. Derivationally, this effect is derived because DP moves after H agreed with it. This counter-bleeding derivation is standardly assumed for Agree between T and the subject DP in English.

(2) *Internal Merge counter-bleeds Agree*:

a. Initial structure

$$H'$$

H WP

$\begin{bmatrix} [∗\phi:\square∗] \\ [\bullet F\bullet] \end{bmatrix}$ ... DP [ϕ:V] ...

b. First step: Agree

$$H'$$

H WP

$\begin{bmatrix} [∗\phi:\square∗] \\ [\bullet F\bullet] \end{bmatrix}$ ... DP [ϕ:V] ...
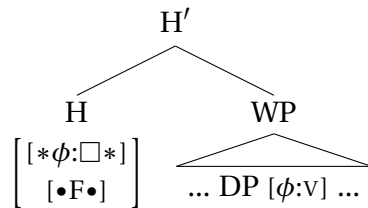
❶ Agree

c. Result: probe valued

$$H'$$

H WP

$\begin{bmatrix} [\phi:V] \\ [\bullet F\bullet] \end{bmatrix}$ ... DP [ϕ:V] ...
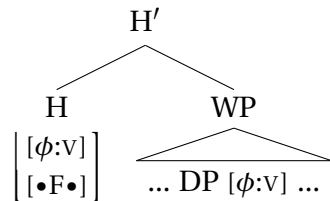
   d.   Second step: Merge, [•F•] discharged
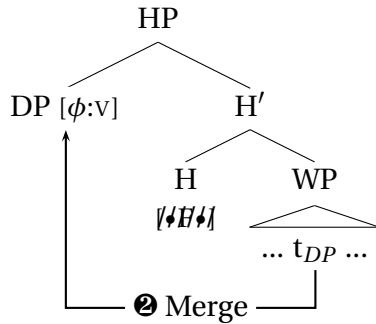
Crucially, if DP is not moved out of WP at all, H also agrees with it, just as in (2). This is trivial, since DP is in the c-command domain of H at every point of the derivation. Recall that in languages with upward Agree, no Agree relation can be established if DP does not move to SpecHP; that this is possible in the languages we will look at in this chapter suggests that Agree applies downwards.

     The bleeding derivation in (1) requires a comment. To obtain bleeding, it is crucial that the occurrence of DP within WP is not visible for the probe after DP has moved to SpecHP, i.e. the trace, copy or the lower chain link of DP must not count as a goal, otherwise H could Agree with the lower occurrence (see e.g. Chomsky 2000: 131, Anand and Nevins 2005: 16 for this conclusion; see Řezáč 2004: 51ff. for general discussion). The condition that traces (or their Minimalist counterparts) must be invisible for Agree is not a specific requirement of the present approach, it is needed in most existing analyses of such bleeding effects, especially for the data that will be discussed in section 3.3. The question is why this should be so. It follows trivially if movement does not leave behind anything (cf. Epstein and Seely 2002), an assumption the present approach to interactions is fully compatible with. If one adopts the copy theory or the multi-dominance approach, however, the problem is real. I have nothing new to add to this discussion here; I will simply assume it as a fact that copies / lower chain links (if existent) are invisible for Agree.[1]

     With this much background on the interaction of internal Merge and downward Agree, let me come to the data that suggest that there is a split between final and intermediate movement steps: One type of movement bleeds downward Agree whereas the other type counter-bleeds it.

---

[1]Here are two solutions proposed in the literature: van Koppen (2007) argues that the internal structure of lower copies is not accessible to Agree because lower copies are reduced copies. Only the features on the highest projection of the copy are accessible, but not features on nodes dominated by this projection (the internal structure of the copy is invisible). So, if the features that take part in Agree are not projected to the DP-level, the probe cannot access the features it searches for on the lower copy of DP. We will mainly be concerned with ϕ-feature Agree in this section, hence, there must not be ϕ-feature projection if van Koppen's (2007) approach was adopted. Note, however, that ϕ-feature projection *is* required for the phenomenon van Koppen investigates (first conjunct agreement), so languages would have to differ in which features can project. Another solution is put forward in Řezáč (2004: 52ff.). In a nutshell, he assumes that a functional shell is inserted on top of lower copies and this shell shields the features of the copy from being visible to an outside probe. For concreteness, he proposes that the D-head of lower copies is replaced by a D-head that yields definite descriptions (interpretation of copies as definite descriptions has been proposed independently, see e.g. Fox 2002). By assumption, the newly introduced determiner lacks ϕ-features. Thus, the DP projected by D is not a goal for ϕ-Agree anymore; the NP that D embeds bears these features but it is too deeply embedded to be visible. This solution is not fully compatible with the assumptions made in this thesis. For one, insertion of the D-shell violates *Inclusiveness* and second, it is counter-cyclic (the inserted D-head even projects).

3.2   (Counter-)bleeding

I will present two case studies of the interaction of downward Agree and internal Merge: the Anti-Agreement effect and movement-related case splits. The former involves bleeding of ϕ-Agree triggered by C and the latter bleeding of case-Agree triggered by a DP-internal functional head.

3.2.1   The Anti-agreement effect

3.2.1.1   Data: AAE in Berber

In a number of typologically diverse languages such as Berber, Breton, Welsh, Palauan, Yimas and Turkish, ϕ-Agree between the verb and the subject reduces to default agreement (3sg or absence of any marker) if the subject is Ā-moved (questioned, relativized, focussed) to SpecCP of the minimal clause containing it. This phenomenon is known as the Anti-agreement effect (AAE, cf. Ouhalla 1993; Richards 1997; Phillips 2001 for an overview). It is illustrated with wh-extraction data from Berber in (3). If the subject is not Ā-moved, the verb agrees with it in person, number and gender, cf. (3-a).[2] If, however, the subject is questioned, as in (3-b), the verb shows default agreement (3sg masculine, glossed as 'participle').[3] Full agreement is ungrammatical in this context, cf. (3-c). Interestingly, in some of the languages that exhibit the AAE, long-distance Ā-movement of the subject to a higher clause does not result in default agreement in the embedded clause in which the subject originates. Rather, there is full ϕ-agreement on the embedded verb (cf. (3-d)) as if the subject had not been extracted at all.

(3)   *Anti-agreement in Tarifit Berber under wh-movement (Ouhalla 1993: 479f., 487)*:

    a. **zri-n**   imhdarn Mohand
        saw-3PL students Mohand
        'The students saw Mohand.'   *no movement, full agr.*

    b. man   tamghart ay **yzrin**   Mohand
        which woman   C see.PARTCP Mohand
        'Which woman saw Mohand?'   *local SU-extraction, default agr.*

    c. *man   tamghart ay **t-zra**   Mohand
        which woman   C 3SG.FEM-saw Mohand
        'Which woman saw Mohand?'   *local SU-extraction, full agr.*

    d. man   tamghart ay nna-n   qa **t-zra**   Mohand
        which woman   C said-3PL that 3SG.FEM-saw Mohand
        'Which woman did they say saw Mohand?'   *long SU-extraction, full agr.*

---

[2]In the absence of Ā-movement, there is always full subject-verb-agreement, independent of the position of the subject (preverbal vs. postverbal), see the example in (i) from Tamazight Berber:

(i)   *Full agreement without Ā-movement in Tamazight Berber (Ouali 2008: 164)*:

    (thamttut) th-e3la   (thamttut) araw
    woman   3SG.FEM-see.PERF woman   boys
    'The woman saw the boys'

[3]The verb form in (3-b) is glossed as 'participle', but Ouhalla (1993) argues that this form contains the default 3rd person masculine form of agreement. In most other languages with the AAE, the verb form under local Ā-extraction is completely identical to 3rd person singular agreement.

In terms of rule interaction, the AAE can be described as follows: Local Ā-movement of the subject DP to SpecCP *bleeds* ϕ-Agree, whereas long Ā-movement has the opposite effect. Assuming that a subject DP that is to undergo long-distance extraction uses the minimal SpecCP position of the clause in which it originates as an intermediate landing site (for which we have seen massive empirical evidence in chapter 2), we have an instance of opacity: Under both local and long subject Ā-movement, there is a DP in the minimal SpecCP position, but sometimes movement to this position bleeds ϕ-Agree and sometimes it does not, i.e. long extraction counter-bleeds ϕ-Agree in the embedded clause.[4]

3.2.1.2 Analysis

The central difference between local and long-distance Ā-movement is the nature of the landing site SpecCP: Under local Ā-movement of the subject DP, SpecCP is the final landing site; movement to this position discharges the [•WH•]-feature on C (or a [•REL•]-feature that triggers operator movement in a relative clause, etc.). Long-distance movement of the subject DP uses the minimal SpecCP only as an intermediate landing site; this movement step discharges an edge feature [•EF•] on C. The C head that projects the intermediate landing site does not bear a [•WH•]-feature, the embedded sentence is a declarative sentence, cf. (3-c). The different effects of final and intermediate movement steps on ϕ-Agree follow if the features that trigger the respective movement types are ordered differently relative to the ϕ-probe. I assume that the ϕ-probe is located on C and looks downwards. That ϕ-Agree applies downwards in Berber is obvious: If the subject is not Ā-moved, there is full agreement on the verb, suggesting that the probe seeks for a goal in its c-command domain. The opacity effect is derived if final movement steps to SpecCP apply before Agree and intermediate movement steps apply after Agree, cf. (4). This corresponds to the pattern II order of chapter 2:

(4)   *Ordering statement in Tarifit Berber (first version)*:
       $[\bullet\text{WH}\bullet] \succ [*\phi{:}\square*] \succ [\bullet\text{EF}\bullet]$

The consequences of this ordering are as follows: If there is no subject extraction, as shown in (5), C only bears a ϕ-probe and no structure-building features that trigger Merge. When the probe searches for the closest goal in its c-command domain, it finds the subject DP, which is either in its postverbal vP-internal position (cf. (3-a)), or in its preverbal SpecTP position (cf. footnote 2; see Ouali 2011: ch.4 for arguments that preverbal subjects in Berber occupy SpecTP). Consequently, the probe on C is valued with the ϕ-features of this DP (the value is abstractly represented as 'v').

---

[4]The AAE is also prominent in Bantu languages. However, the AAE in Bantu differs in a number of properties from the AAE in other languages: The agreement morphology under subject extraction is not a default marker in Bantu (see Henderson 2013 for arguments); in addition the effect only occurs under extraction of a small class of (non-pronominal) subject DPs, viz. with singular DPs of noun class 1. In other AAE languages, the effect applies to all subject DPs, and if it results in 3sg marking, it only becomes visible when plural DPs are extracted (in contrast to Bantu languages where only singular DPs trigger AAE in the first place). Therefore, I do not claim that the present analysis also applies to AAE in Bantu; it may well be that the Bantu AAE has a different source. I will leave it for future research to investigate whether the AAE in Bantu can be subsumed under the ordering approach.

(5)　　*No Ā-movement of the subject DP*:

　　a.　φ-probe searches and finds a goal

$[_{CP}$ C $_{[*\phi:\square*]}$ $[_{TP}$ DP$_{[\phi:V]}$ $[_{T'}\ldots]]]$

└── Agree ──┘

　　b.　Result: φ-probe valued and discharged

$[_{CP}$ C $_{[\phi:V]}$ $[_{TP}$ DP$_{[\phi:V]}$ $[_{T'}\ldots]]]$

Under local subject extraction, C bears a φ-probe as well as the structure-building feature [•WH•], which triggers a final movement step. The structure-building feature is discharged first; it attracts the wh-subject to SpecCP, cf. (6-a) (for the sake of readability, I represent the order of operation-inducing features not by a stack but in the form of an ordering statement with X ≻ Y meaning X triggers an operation before Y). Afterwards, the probe initiates Agree, but since the closest goal, the subject DP, is no longer in C's c-command domain due to prior movement, Agree fails and the probe is valued by default, cf. (6-b-c). Hence, final movement of the subject DP to SpecCP applies too early and thereby bleeds φ-Agree.

(6)　　*Local Ā-extraction of the subject DP*:

　　a.　Step 1: movement of the subject to SpecCP, [•WH•]-feature discharged

$[_{CP}$ wh-DP$_{[\phi:V]}$ $[_{C'}$ C $_{[\cancel{\bullet WH\bullet}}$ ≻ $_{[*\phi:\square*]\}}$ $[_{TP}$ t$_{DP}$ $[_{T'}\ldots]]]]$

└────────── Move ──────┘

　　b.　Step 2: Agree, no goal found

$[_{CP}$ wh-DP$_{[\phi:V]}$ $[_{C'}$ C $_{[*\phi:\square*]}$ $[_{TP}$ t$_{DP}$ $[_{T'}\ldots]]]]$

└ *Agree ┘

　　c.　Result: default valuation of the probe

$[_{CP}$ wh-DP$_{[\phi:V]}$ $[_{C'}$ C $_{[\phi:3SG\ MASC]}$ $[_{TP}$ t$_{DP}$ $[_{T'}\ldots]]]]$

In contrast to that, long Ā-extraction of the subject that uses the minimal SpecCP as an intermediate landing site is triggered by an edge feature. This movement applies after Agree and therefore counter-bleeds Agree: First, the probe on the C head of the embedded clause initiates Agree and finds the subject DP as its closest goal, cf. (7-a-b). It is only afterwards that the subject DP moves to SpecCP, triggered by the edge feature on C, cf. (7-c). On the surface, the subject DP is no longer in the c-command domain of the embedded C, which gives rise to opacity; however, the subject was in the relevant domain at the point of φ-Agree.

(7)　　*Long Ā-extraction of the subject DP*:

　　a.　Step 1 (embedded clause): φ-Agree initiated by C

$[_{CP}$ C $_{[*\phi:\square*]\succ[\bullet EF\bullet]\}}$ $[_{TP}$ wh-DP$_{[\phi:V]}$ $[_{T'}\ldots]]]$

└────── Agree ──────┘

　　b.　Result: probe valued and discharged

$[_{CP}$ C $_{[\phi:V]\succ[\bullet EF\bullet]\}}$ $[_{TP}$ wh-DP$_{[\phi:V]}$ $[_{T'}\ldots]]]$

　　c.　Step 2 (embedded clause): movement to SpecCP, [•EF•]-feature discharged

$[_{CP}$ DP$_{wh}$ $[_{C'}$ C $_{[\bullet EF\bullet]}$ $[_{TP}$ t$_{DP}$ $[_{T'}\ldots]]]]$

└────── Move ──────┘

　　d.　Step 3: successive-cyclic movement of the wh-DP to its final position in the matrix clause

Note that the opaque AAE is just what we expect if final and intermediate movement steps can apply at different points of the derivation, as argued for extensively in the previous chapter, and if Agree applies *downwards* instead of upwards. Ouhalla (2005: 664, fn.5) discusses interesting variation between varieties of Berber with respect to the features involved in the AAE. In Tamazight Berber, for instance, person and gender agreement is bled under local extraction, but number agreement is not; in Ouargli Berber, only person agreement is bled under Ā-movement, number and gender agreement is preserved, however. This suggests that there is no single ϕ-probe; rather, person, number and gender are separate probes. Consequently, these probes can be ordered differently relative to Merge. The relevant orders are given in (8). The Agree relations that are initiated before final movement steps cannot be bled by movement because the subject DP is still in the c-command domain of C; it is moved in a subsequent step.

(8)   *Order of probe features relative to Merge triggers in varieties of Berber*:

   a.   Order in Tarifit Berber (final version):
        [•WH•] ≻ [∗π:□∗], [∗#:□∗], [∗GEN:□∗] ≻ [•EF•]

        final movement bleeds person / number / gender Agree

   b.   Order in Tamazight Berber:
        [∗#:□∗] ≻ [•WH•] ≻ [∗π:□∗], [∗GEN:□∗] ≻ [•EF•]

        final movement bleeds person / gender Agree but not number Agree

   c.   Order in Ouargli Berber:
        [∗#:□∗], [∗GEN:□∗] ≻ [•WH•] ≻ [∗π:□∗] ≻ [•EF•]

        final movement bleeds person Agree but not number / gender Agree

Note that when an object DP is extracted, we do not get the AAE because the probe on C will still find the subject DP in SpecTP to Agree with (Agree applies downward). The same holds for the matrix clause under long extraction of the embedded subject: There is full ϕ-Agree between the probe on C and the subject of the matrix clause which is in C's c-command domain.

   In what follows, I address some issues that arise for the present approach. First of all, the AAE in Berber (and other AAE languages) surfaces on the verbal complex in T and not on the complementizer, where I locate the ϕ-probe. But this assumption is unavoidable because the AAE is only triggered by movement to SpecCP. Thus, we find exactly the same situation that we already encountered in Duala and Ewe: Ā-Movement to SpecCP has a morphological reflex in the TP-domain. As in Duala, I assume that there is a relation between C and T by which the inflection on C gets to T, where it surfaces. There are various ways to achieve this: (i) Either the verb picks up the inflection in C by V-to-C movement, or (ii) the inflection is lowered to the T-V-complex in the morphological component.[5] For Berber, solution (ii) is to be preferred because the complementizer is realized as a separate constituent, it is not an affix to the verb. This assumption that the Agree relation that leads to the reflex is triggered by C although the reflex is found in the TP-domain has also been made for Ewe by Collins (1993), and in fact, for AAE languages: Ouali (2008) and Henderson (2013) argue that the AAE must involve ϕ-Agree between C and the extracted subject (in addition to Agree between T and the subject).

---

[5]Lowering of the inflection is reminiscent of Affix-Hopping (Chomsky 1957). Furthermore, it seems plausible given the concept of Feature Inheritance, i.e., ϕ-feature transfer from C to T, as proposed in Chomsky (2004); Richards (2007); the only difference is that I assume that this transfer may apply late, in the morphological component (cf. morphological merger in Embick and Noyer 2001) so that syntactic movement can interact with the features on C. See also Ouali (2008) on variation in the timing of Feature Inheritance.

Another issue that arises is the following: If the subject of a transitive verb is locally extracted, why does the ϕ-probe on C not Agree with the lower object DP? Since Agree applies downwards and the object is the closest goal in its c-command domain after the subject has been extracted, this should be a possibility. Thus, local extraction should bleed subject agreement and at the same time feed object agreement. However, this does not happen in any AAE language. A related problem occurs under long object extraction: The object uses the minimal SpecCP as a intermediate landing site, discharging an edge feature on C; since I assume that every phrase is a phase, the object must also make a stop-over in SpecTP. The following situation could arise under long-object extraction if the subject is preverbal (in SpecTP): The object lands in the outer specifier of the embedded T, preceding the subject in the inner specifier of T (since the subject undergoes a final movement step to SpecTP, it moves first and lands in the inner specifier). It is thus closer to the ϕ-probe on C than the subject. And since the object will undergo edge feature-driven movement to SpecCP (due to $[*\phi{:}\square*] > [\bullet\textsc{edge}\bullet]$), it is still in C's c-command domain when C probes. Consequently, we would expect object agreement on the C head in this configuration, which is not what we find. Several solutions to these problems come to mind. First, under local subject extraction, the object DP might be too far away from the probe on C; it could be inaccessible for C because it is in the complement domain of the phase vP dominated by CP. However, recall that I assume that Agree is not subject to the PIC (cf. section 2.5.2.1.2). In section 3.3 I will present evidence from Icelandic for this assumption. In Icelandic, the T head can establish a ϕ-Agree relation with a DP in an embedded infinitive, suggesting that Agree can span several XP boundaries (at least two: the matrix vP, VP; and probably more if auxiliaries, negation, etc. intervene); the distance between C and the object DP in Berber is not larger than the one that Agree can span in Icelandic. Therefore, the absence of object agreement under local object extraction cannot be due to a locality restriction on Agree. Moreover, the locality solution would not provide an answer why there is no C-object agreement in the embedded clause under long-distance object extraction; in this case, the object in the outer SpecTP is definitely accessible to C. Another solution for both problematic cases could be to invoke the *Activity Condition* (Chomsky 2001). According to this condition, the goal must be active to enter into an Agree relation with a probe; it is active for ϕ-Agree if it has an unvalued case feature. Since the object of a transitive verb is standardly assumed to be assigned case by v (at least in nominative accusative languages like Berber), it is inactive at the point where C probes and thus, it is not a possible goal for the ϕ-probe on C. There are, however, several problems with adopting the *Activity Condition* in the first place, see Nevins (2004) for an overview. In addition, Nevins (2004) and Bošković (2007*b*) argue that Activity can be dispensed with since its effects follow from independent factors. Let me name a few problematic phenomena for the *Activity Condition*: First of all, the existence of defective intervention is mysterious if we adopt Activity: Inherently case-marked DPs block an Agree relation across them although they should be inactive (see section 3.3 for an example). A related problem is complementizer agreement in a number of West Germanic languages: There is ϕ-Agree between C and the subject DP in its c-command domain. But by standard assumptions, the subject is already case-marked when C probes. In addition, I will show in section 3.2.2.3 that in the Uralic language Nenets there is ϕ-Agree between a DP-internal head and an already case-marked possessor. Likewise, I also assume that the Ā-moved DP in AAE languages is assigned case before it moves (by C or a lower functional head). If local Ā-movement would also bleed case assignment, the DP would end up without case. Given this assumption, ϕ-Agree must be possible between C and the case-marked subject DP if the latter is to undergo long-distance extraction or if it is not

moved at all. This would be impossible if the *Activity Condition* was generally valid. Thus, we need a different explanation for the absence of Agree between C and the object DP in AAE languages.

I propose the following solution: ϕ-Agree can only target goals with a certain case value, viz. nominative in Berber. Put differently, only DPs with a certain case value are visible for ϕ-Agree; DPs with a different case are invisible for a ϕ-probe. Thus, I make use of the concept of *agreement opaqueness* proposed by Keine (2010: 30) according to which "[t]he presence of certain language-specific case features on a DP renders the features of this DP incapable of acting as a goal for Agree."; Keine uses this, for instance, to account for the fact that ergative marked DPs cannot trigger ϕ-agreement on the verb in Hindi. See also Schütze (1997) and Řezáč (2008) on the non-transparency of DPs with inherent case for ϕ-Agree. If only nominative DPs are visible for Agree in Berber, the object DP, which is assigned accusative in the vP, is invisible as a goal for the ϕ-probe on C. Consequently, it can never value this probe. If the *Activity Condition* is abandoned, a constraint along these lines is required independently for other phenomena. In German where the object can scramble above the subject we find the same subject-object-asymmetry with respect to ϕ-Agree: Subject-verb-agreement mediated by a ϕ-probe on T can only target the subject, even if the object is scrambled above it and should thus qualify as closer to T than the subject. The same issue arises for complementizer agreement, triggered by a ϕ-probe on C, which can also target exclusively the subject. This problem is scarcely addressed in the relevant literature, but for these phenomena, we also need to restrict ϕ-Agree to the nominative DP in some way.[6]

3.2.1.3   Variation

The AAE data are opaque if long Ā-movement makes a stop-over in the embedded SpecCP position. However, there is an alternative analysis of the counter-bleeding effect in AAE languages that does not rely on this intermediate movement step: If local movement of a DP to SpecCP causes bleeding, but there is no bleeding under long Ā-extraction, one could propose that wh-movement to the matrix clause does not make a stop-over in SpecCP of the embedded clause but moves directly to the matrix SpecCP position. Under this analysis, it follows that there is full subject-verb-agreement in the embedded clause: When the ϕ-probe on the embedded C searches for a goal, the wh-subject is still in its base position in the embedded clause and thus in the embedded C's c-command domain; it can only move out of the c-command domain of the embedded C when matrix C, bearing [•WH•], is merged.

In what follows, I will show that the assumption that long Ā-movement does not go through the embedded SpecCP causes problems when we look at cross-linguistic variation. There are languages in which *both* local and long-distance Ā-movement of the subject DP bleeds ϕ-Agree. This is the case in some Italian dialects (Fiorentino and Trentino, cf. Brandi and Cordin 1998; Campos 1997) and in Ibibio (cf. Baker 2008*b*). (9-a) vs. (9-b) illustrates that local subject extraction in Fiorentino obligatorily leads to default agree-

---

[6]Note that Activity could solve the first issue (scrambling of the object above the subject): Subject-verb-agreement is triggered by a ϕ-probe on T; the subject may not have been assigned nominative case by T at this point, whereas the object has already been assigned case by v. But Activity still does not solve the complementizer agreement facts: Both the subject and the object are assigned case when C probes, nevertheless, ϕ-Agree is possible. But if Activity isn't the decisive factor, something else must make sure that only the subject DP can ever be a goal for the ϕ-probe on C. The invisibility of non-nominative DPs is a way to achieve the desired result.

ment, just as in Berber. In contrast to Berber, however, default agreement is also obligatory under long subject extraction, compare (9-c) and (9-d).

(9)   *The AAE in Fiorentino (taken from Ouhalla 1993; Campos 1997)*:

    a. Quante   ragazze gli   ha   parlato con te
       how.many girls   CLT.3SG have.3SG spoken to   you
       'How many girls (it) has spoken to you?'

    b. *Quante   ragazze le   hanno   parlato con te
       how.many girls   CLT.3PL have.3PL spoken to   you
       'How many girls have spoken to you?'

    c. Quante   ragazze tu credi che gli   ha   telefonato
       how.many girls   you think that CLT.3SG have.3SG phoned
       'How many girls do you think have phoned?'

    d. *Quante   ragazze tu credi che le   hanno telefonato
       how.many girls   you think that CLT.3PL have.3PL phoned
       'How many girls do you think have phoned?'

Given the assumption that long extraction does not go through the embedded SpecCP position, it is mysterious that there is an AAE effect in the embedded clause in Fiorentino type languages: The subject DP that is to undergo long extraction is still in the c-command domain of the embedded C when this head probes; hence, the ϕ-probe on C should be valued by the subject DP. One solution to this problem could be to say that languages differ in whether long-distance Ā-movement makes a stop-over in the embedded SpecCP: In the Fiorentino type languages it does (and thereby causes bleeding), while in the Berber type languages it does not (movement in one-fell swoop does not cause bleeding). In my view, this analysis is clearly undesirable. First of all, we have seen massive cross-linguistic evidence from reflexes of movement (semantic, syntactic, morphological) in chapter 2 that SpecCP is an intermediate landing site. Furthermore, it is dubious to assume that languages differ so drastically in the size of their locality domains. In the languages in which long Ā-movement does not apply successive-cyclically, Ā-movement would be able to span any distance, creating a non-local dependency between the matrix C and a DP in a potentially very deeply embedded clause; in other languages, Ā-movement would apply in a sequence of very local steps.

      Under the analysis of the opacity effect presented in 3.2.1.2, there is another way to account for variation between AAE languages of the Berber vs. the Fiorentino type: Assume that languages do not have different locality domains for movement; long Ā-movement always makes a stop-over in the embedded SpecCP. Cross-linguistic variation is accounted for by reordering of operation-inducing features on C: In the Fiorentino type languages both intermediate and final movement steps apply *before* Agree, cf. (10). Hence, the wh-subject DP is moved out of the c-command domain of C before the probe on C initiates Agree; consequently, there is always bleeding of ϕ-Agree under Ā-movement. Thus, Fiorentino type languages have the pattern I order (both movement types apply before Agree).

(10)   *Ordering statement in Fiorentino type languages*:
    $[\bullet \text{WH} \bullet]$, $[\bullet \text{EF} \bullet] \succ [*\phi*]$

In Berber type languages, on the other hand, the two types of Merge triggers apply at different points of the derivation relative to Agree, giving rise to opacity, cf. (4).

Languages without the AAE can be described as having the pattern IV order in (11).

(11)   *Ordering statement in languages without the AAE*:
       $[*\phi*] \succ [\bullet\text{WH}\bullet], [\bullet\text{EF}\bullet]$

I am not aware of an AAE-language with the pattern III order $[\bullet\text{EF}\bullet] \succ [*\phi*] \succ [\bullet\text{WH}\bullet]$, but the current account predicts such languages. In such a language, only long-distance Ā-movement would cause bleeding in the clause in which the extracted DP originates; local Ā-movement would not cause bleeding. This is the reverse of the Berber type AAE.[7]

3.2.1.4 Previous analyses

There are numerous analysis of the Anti-agreement effect: Brandi and Cordin (1998); Ouhalla (1993; 2005); Richards (1997); Phillips (2001); Boeckx (2003); Ouali and Pires (2005); Cheng (2006); Schneider-Zioga (2007); Baker (2008a); Ouali (2008); Henderson (2009; 2013); Diercks (2010). I will not go through each of them in detail here. I pick out two of them that are interesting when it comes to the account of cross-linguistic variation because they are similar in spirit to the analyses proposed for the variation between patterns I and II arising under upward Agree (cf. section 2.4).

The analyses I will discuss are Ā-binding approaches (Brandi and Cordin 1998; Ouhalla 1993) and anti-locality approaches (Cheng 2006; Schneider-Zioga 2007). The basic idea of the former type of approach is that subject extraction leaves behind a *pro* in the subject position that must not be bound by its antecedent from an Ā-position (Principle B applied to Ā-binding). It is, however, bound under local subject extraction by the subject in SpecCP. In order to avoid the violation of Principle B, *pro* must not be licensed in the first place. It is licensed by rich agreement; dropping the agreement, resulting in AAE, makes *pro* unavailable and thus avoids the principle B violation.

In the anti-locality approaches, it is argued that local extraction of the subject is too local in the sense of Grohmann (2003). Grohmann argues that movement must cross at least one domain boundary. The relevant domains are the Ω-domain ($\approx$ CP), the Θ-domain ($\approx$ TP), and the Φ-domain ($\approx$ vP). According to the approaches by Cheng (2006); Schneider-Zioga (2007), the landing site of the locally Ā-moved subject DP is in the same domain as the canonical subject position in AAE languages.[8] As a repair at PF, the occurrence of the subject in its lower position is spelled out by a special morpheme, yielding the exceptional AAE morphology.

Apart from a number of general conceptual problems with these approaches that I will not address here, both approaches have problems to account for cross–linguistic variation: So far I have only talked about local extraction; but what about long extraction, where some languages exhibit the AAE but other's don't? As for the Ā-binding approach,

---

[7]In some languages with the AAE, the bleeding effect of local Ā-movement is only optional if the clause contains a sentential negation, i.e. there can be full subject-verb-agreement even if the subject is locally extracted. This holds e.g. in Berber and Celtic, but not in Turkish (see Ouhalla 1993: 499ff.). The negation effect does not fall out in any obvious way from the present analysis; the order of features on C should be independent from the presence of negation. The few existing analyses of the negation effect (Ouhalla 1993; Phillips 2001; Baker 2008a) are not compatible with the present approach because they start from completely different basic assumptions to derive the AAE.

[8]This is the case because the domain sizes adopted by Cheng (2006) differ from those postulated by Grohmann (the subject position SpecTP and the landing site of Ā-movement SpecCP are, by assumption, part of the same domain). Schneider-Zioga (2007) proposes for Kinande (Bantu) that the subject is in a dislocated position in the left periphery before it undergoes Ā-movement, making this movement step too local.

Ouhalla (1993) assumes that long Ā-extraction always passes through SpecCP of the embedded CP, viz. that it applies successive-cyclically. Since languages like Berber do not show the AAE with long extraction, Ouhalla assumes that the trace of the subject in the embedded SpecCP position cannot act as an Ā-binder. To account for the occurrence of the AAE with long extraction in the Fiorentino-type languages, he proposes that the trace can be an Ā-binder in these languages. Hence, languages vary in the ability of traces to be Ā-binders. This is exactly what the quasi-operators approach proposed for the difference between pattern I and pattern II reflexes of upward Agree. I have argued in section 2.4.5.1 that this approach cannot be maintained for conceptual and empirical reasons: First, we need to distinguish two different types of empty elements left by the same movement operation; the ordering approach does not need this additional assumption. Furthermore, the Ā-binding approach cannot easily capture pattern III under upward Agree, and finally, it cannot account for languages where pattern I and pattern II reflexes of upward Agree co-occur.

Let me now turn to the anti-locality approaches. In Berber-type languages where the AAE does not arise under long extraction, the proponents of this approach assume that the subject DP does not make a stop-over in the embedded SpecCP position; rather, movement applies in one-fell swoop. Therefore, movement of the subject is not too local in the embedded clause and the repair, resulting in the AAE, is not necessary. In Fiorentino-type languages in which the AAE also occurs with long extraction, long-distance movement does go through the embedded SpecCP. This analysis of Berber-type languages is an instance of the "where there is no reflex of movement, there is no movement in the first place"- approach, which denies opacity in the data, see section 2.4.1 for general discussion. As I pointed out above, I take it to be undesirable to assume that languages vary so drastically in the locality of Ā-movement.

Without going into the details of his analysis of the AAE, Henderson (2013: 464f., fn.12) also considers the possibility that long Ā-movement in Berber-type languages might apply in one-fell swoop, and therefore, the effect does not occur in the embedded clause. Alternatively, he speculates that the absence of the AAE in embedded clauses might be due to the fact that the agreement relation that results in the reflex of movement (the AAE) does not take place because the embedded C head lacks a probe. The latter analysis is an instance of the "where there is no reflex of movement, there is no agreement with the moved phrase"- approach, see section 2.4.2. We have seen evidence in the aforementioned section that the occurrence of default markers argues against the assumption that there is no probe on the embedded C head. And in many AAE languages, we do find a default marker, suggesting that there has been an attempt to Agree (see the example from Berber in (3-b) and footnote 3).

To summarize, we have seen that the core ideas in the accounts of variation in the AAE are the same as those that have been proposed for reflexes of upward Agree and hence, they encounter the same problems. The reordering approach presented in this thesis is not confronted with these problems because it acknowledges the opacity in the data (for which there is empirical evidence); in addition, it neither requires two different types of empty elements in intermediate landing sites, nor does it have to assume that Ā-movement can be unbounded in some languages. I take this to be a desirable outcome. And let me stress again that the Berber-type AAE is exactly what we expect to exist when the ordering approach with a distinction between final and intermediate movement steps, argued for in chapter 2, is combined with downward Agree.

3.2.2   Movement-related case splits

In this subsection I present another instance of counter-bleeding of the abstract scheme introduced in section 3.1. In contrast to the previous example, the Agree relation will result in case valuation instead of ϕ-agreement. In particular, I look at the encoding of the possessor in Hungarian (Uralic, Hungary). More generally, I will show that ordering of Merge and case-Agree provides an elegant way to derive movement-related case splits.

3.2.2.1   Data: Hungarian possessor encoding

Let me begin with some basic facts about the encoding of possession in Hungarian. The possessor precedes the possessum and the possessum agrees in ϕ-features with the possessor. The possessor exhibits a case split: It can bear either nominative or dative, cf. (12). As argued in Szabolcsi (1994), the two possessors occupy different structural positions: The dative possessor is in the specifier of a functional head in the left periphery of the DP, viz. SpecDP. It is moved to this position from a lower position that is associated with nominative case. Evidence for this analysis comes from three facts: (i) The dative possessor precedes the determiner *a(z)* whereas the nominative possessor follows this element (cf. (12-a) vs. (12-b));[9] (ii) only the dative possessor can be extracted out of the DP, cf. (13); (iii) a wh-possessor must bear dative case and precedes the determiner, cf. (14). Facts (ii) and (iii) show that the position the dative possessor occupies is similar to SpecCP in the clause: It hosts operators and serves as an escape hatch for movement out of the DP. As long as the possessor stays within the DP, the choice between a nominative and a dative possessor is optional, there is no difference in meaning between the constructions (cf. (12-a-b)). Finally, both the dative and the nominative possessor precede a special class of determiner-like elements such as 'each', cf. (12-c); these are lower in the structure than the determiner *a(z)*.

(12)   *Possessors in Hungarian (Szabolcsi 1994: 181, 187, 188)*:

    a.   (a) Mari   kalap-ja
        the Mari.NOM hat-POSS.3SG
        'Mari's hat'

    b.   Mari-nak a   kalap-ja
        Mari.DAT the hat-POSS.3SG
        'Mari's hat'

    c.   a   te   valamenyi tiik-od
        the you.NOM each   secret-POSS.2SG
        'your every secret'

(13)   *Extraction data (Szabolcsi 1994: 181-182)*:

    a.   Mari-nak$_k$ nem ismert-em [$_{DP}$ ___$_k$ növér-é-t   ]
        Mari.DAT not knew-1SG   sister-POSS.3SG-ACC
        'I never knew any sister of Mari.'

    b.   *Mari$_k$ nem ismert-em [$_{DP}$ ___$_k$ növér-é-t   ]
        Mari.NOM not knew-1SG   sister-POSS.3SG-ACC
        'I never knew any sister of Mari.'

---

[9]The variation in the form of the determiner is purely phonological: The final consonant is dropped if the following segment is a consonant.

(14)   *Wh-possessors in Hungarian (Szabolcsi 1994)*:

   a. *ki   kalap-ja
      who.NOM hat-POSS.3SG
      'whose hat?'

   b. ki-nek$_k$ a   ___$_k$ kalap-ja
      who-DAT the   hat-POSS.3SG
      'whose hat?'

Against this empirical background, I make the following assumptions about the structure of the DP in Hungarian, illustrated in (15): The possessor (Poss) is merged as the complement of N (the possessum) for reasons of theta-role assignment, cf. Delsing 1998; de Vries 2006; Georgi and Salzmann 2011. Determiner elements like 'each' take NP as a complement; these elements are of category Det (to distinguish them from the determiner element *a(z)* which I take to be of category D). In order to derive the surface order Poss before Det, the possessor must move to the specifier of a functional projection above Det. I take this to be the specifier of the functional head *n. n* has a trigger for this movement, represented as [•D•] in what follows. This feature is kind of an EPP-requirement of *n* and derives the fact that nominative and dative possessors precede Det. In addition, *n* assigns dative case to the possessor and thus bears a case probe: [*CASE:DAT*]. A case probe is different from the probes we have seen so far in that it already bears a value; to get discharged, it must find a goal with an unvalued case feature. Agree results in case valuation on the goal. Dative is thus the intended possessor case. The nominative is not a case that is assigned under Agree; rather, I assume that it is the default value instantiated on a possessor if it does not receive a case value under case-Agree with *n.* It has been argued independently by Bartos (2001) and Kiss (2002) that the nominative possessor is caseless, viz. it is not assigned a case value.[10] This may fit well with the analysis presented here if caseless means that the case value is not the result of Agree but rather a default value (that has no overt realization). Finally, the head D, realized as *a(z)*, takes *n*P as its complement. D optionally bears a feature that triggers movement of Poss to SpecDP if present: [•D$_2$•] (the index is used to distinguish this feature from the EPP of *n*), cf. the dashed line in (15). Recall that only dative possessors can be moved to SpecDP; nominative possessors stay in the lower Spec*n*P position, preceding Det but following D elements. The Merge triggering feature on D is not specified for the case of the DP it attracts; that it can only attract dative possessors will fall out from the analysis, see the discussion below. As for agreement between Poss and the possessum, I assume that it is mediated by *n*: *n* bears a ϕ-probe that Agrees with the possessor. The exponent is realized on N due to N-to-n movement (which will be ignored in what follows).

---

[10]An argument for this conclusion is that bare demonstratives (occurring without an associated noun) can occur in case-marked positions, e.g. as subjects of the clause (associated with nominative case) or as dative possessors. However, they cannot be used as nominative possessors. The conclusion drawn in the aforementioned literature is that this is because demonstratives need a case-marked associate and the 'nominative' in the possessive construction is not a case assigned to the possessor; the nominative possessor is caseless. See Dékány (2011: ch. 8.7) for an overview of the discussion.

(15)   *Structure of the Hungarian DP*:
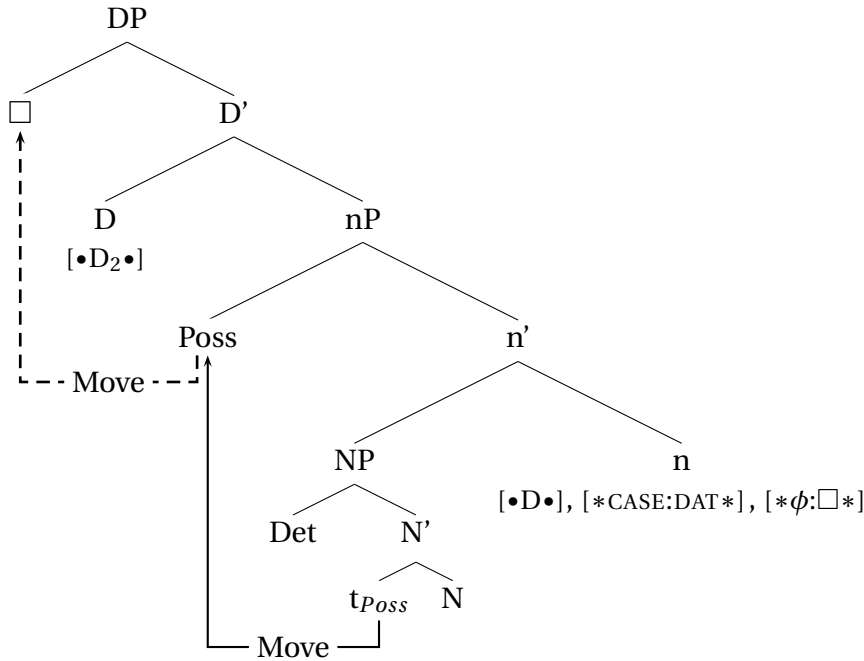
Under these assumptions, the following interactions between case-Agree and Merge arise in the DP: Movement of the possessor to Spec*n*P is obligatory (all possessors precede Det). Since *n* assigns the possessor case under c-command, movement to Spec*n*P may bleed dative case assignment, resulting in default nominative case on the possessor. Opacity arises when looking at the *n*P: All possessors must raise to Spec*n*P; the dative possessor, targeting SpecDP, must make a stop-over in this position if every phrase is a phase. Hence, the possessors occupy the same structural position at the *n*P-stage of the derivation. Nevertheless movement to Spec*n*P has different consequences for case-Agree: If the possessor stays in Spec*n*P, it ends up with nominative case, but if it moves on to SpecDP (optionally), it receives dative case. So, movement to Spec*n*P sometimes bleeds dative case assignment and sometimes it does not (counter-bleeding).[11]

3.2.2.2   Analysis

The bleeding and counter-bleeding case are connected with the nature of the landing site Spec*n*P: If Spec*n*P is the final landing site of the possessor, movement bleeds dative assignment; if, however, the possessor uses Spec*n*P only as an intermediate landing site, moving on to SpecDP, dative case assignment is counter-bled. Thus, we have again evidence for the split of Merge into two types: Intermediate movement steps are triggered by an edge feature on *n*, whereas final movement steps are triggered by [•D•] on *n*. The opacity effect can be derived if the final movement step to Spec*n*P applies *before* Agree, and the intermediate movement step to Spec*n*P applies *after* Agree:

(16)   *A part of the ordering statement in Hungarian*:
$$[•D•] \succ [*\text{CASE:DAT}*] \succ [•\text{EF}•]$$

---

[11] Since the possessor cannot stay in-situ but has to move to Spec*n*P in order to precede Det, we cannot test the direction of Agree. If *n* probed upwards, movement to Spec*n*P would feed Agree; if it probed downwards, movement to Spec*n*P would bleed Agree. I assume downward probing because final movement leads to a less marked case (nominative, zero suffix), which I interpret as a bleeding interaction.

This order of the case-probe relative to the Merge triggers has the following consequences: Final movement of Poss, triggered by [•D•] on *n*, moves Poss to Spec*n*P. Afterwards, *n* triggers dative assignment, but since Poss is no longer in *n*'s c-command domain, it cannot be assigned dative case, cf. (17-b). The result that the case probe on n is deleted by default and Poss is valued by default with nominative, cf. (17).[12]

(17)   *Nominative possessor*:

    a.   Step 1: movement of Poss to Spec*n*P, [•D•] discharged

       [$_{nP}$ Poss$_{[CASE:□]}$ [$_{n'}$ n $_{[•D•]>[*CASE:DAT*]}$ [$_{NP}$ Det [$_{N'}$ ___ N]]]]

       └─────── Move ───────┘

    b.   Step 2: case-Agree fails

       [$_{nP}$ Poss$_{[CASE:□]}$ [$_{n'}$ n $_{[*CASE:DAT*]}$ [$_{NP}$ Det [$_{N'}$ ___ N]]]]
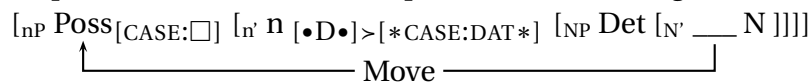
       └── *Agree ──┘

    c.   Result: default case valuation on Poss, default deletion of the probe

       [$_{nP}$ Poss$_{[CASE:NOM]}$ [$_{n'}$ n ~~$_{[*CASE:DAT*]}$~~ [$_{NP}$ Det [$_{N'}$ ___ N]]]]

Possessors that end up with dative case use Spec*n*P as an intermediate landing site; Poss movement to this position is triggered by an edge feature. Due to the order in (16), intermediate steps apply too late to have a bleeding effect: At the point of the derivation where *n* probes (for case), Poss is still in the c-command domain of *n* and is thus assigned dative; it moves to Spec*n*P only afterwards, cf. the derivation in (18).

(18)   *Dative possessor*:

    a.   Step 1: case-Agree

       [$_{nP}$ n $_{[*CASE:DAT*]>[•EF•]}$ [$_{NP}$ Det [$_{N'}$ Poss$_{[CASE:□]}$ N]]]

       └──── Agree ────┘
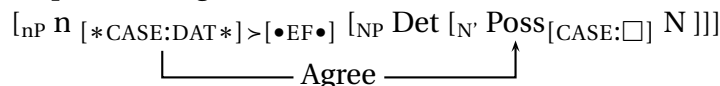
    b.   Result: Poss valued with dative, case-probe discharged

       [$_{nP}$ n ~~$_{[*CASE:DAT*]}$~~$_{>[•EF•]}$ [$_{NP}$ Det [$_{N'}$ Poss$_{[CASE:DAT]}$ N]]]

    c.   Step 2: movement of Poss to Spec*n*P, [•EF•] disacharged

       [$_{nP}$ Poss$_{[CASE:DAT]}$ [$_{n'}$ n $_{[•EF•]}$ [$_{NP}$ Det [$_{N'}$ ___$_k$ N]]]]

       └──── Move ────┘

---
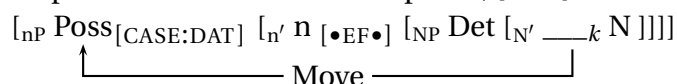
[12]One might wonder why *n* cannot assign dative case to the possessum which is in its c-command domain, too. It does not yet bear a case value at that point of the derivation (its value is determined by the external head that selects the whole DP) and would thus be a potential goal. Note that this is a general problem: In many analysis of DP-internal case assignment, the possessor case is assumed to be assigned by a functional head that has both the possessor and the possessum in its c-command domain. It is unclear why it targets the possessor and not the possessum, a problem that a number of approaches face. See Georgi and Salzmann (2011) for extensive discussion and a number of references to papers that encounter the same question with respect to the possessor doubling construction in German. The following solutions come to mind: (i) Case assignment is tied to XPs and not to heads; (ii) dative assignment is tied to a specific property of the possessor DP, e.g. its thematic role; (iii) case stacking: *n* assigns dative to N, but in addition, N is assigned the case from the external head that selects the DP. Only the case assigned by the external head is realized in the postsyntactic morphological component (cf. Moravcsik 1995; Bejár and Massam 1999; (Merchant 2006; Assmann et al. 2013 on variation in the morphological realization of stacked abstract cases). Under the latter analysis, the dative probe must be able to enter into multiple Agree with the possessor and the possessum (cf. Hiraiwa 2001) because the possessor can in principle be assigned dative, too.

Note that ϕ-Agree between N and the possessor is never bled by movement, there is full ϕ-agreement on N with nominative and dative possessors.[13] The fact that case-Agree can be bled by possessor movement but not by ϕ-Agree suggests that the two are not the reflex of a single Agree operation. Rather, the case- and the ϕ-probe are separated and can thus be ordered differently relative to Merge (see e.g. Marantz 1991; Bobaljik 2008; Baker and Vinokurova 2010; Keine 2010; Preminger 2011*b* for the conclusion that case and ϕ-Agree are triggered by separate probes). The facts follow if ϕ-Agree by *n* applies before *n* triggers any kind of movement of the possessor, cf. (19). Consequently, ϕ-Agree cannot be bled by movement: Poss is always in the c-command domain of the ϕ-probe.

(19)   *A part of the ordering statement in Hungarian (extended version)*:
$$[*\phi{:}\square*] \succ [\bullet D\bullet] \succ [*\text{CASE:DAT}*] \succ [\bullet\text{EF}\bullet]$$

Up to this point, I have left out the final movement step of the possessor to SpecDP. This movement is obligatory if Poss bears dative case, but prohibited if Poss bears nominative. How does this follow from the analysis? Recall that I assume that the $[\bullet D_2 \bullet]$-feature on D is optionally present.[14] There are two scenarios. In the first scenario, D bears the structure-building feature. As a consequence, an edge feature is inserted on *n* when it is selected in the numeration. *Phase Balance* requires edge feature insertion on *n* because D, bearing $[\bullet D_2 \bullet]$, is still in the numeration at this point (cf. the definition of *Phase Balance* in (148), chapter 2). Furthermore, $D \neq n$, and there is no matching [D]-feature in the numeration: The possessor (and further optional DP arguments of N) is (are) already part of the structure. In the second scenario, D does not bear a structure-building feature. Consequently, no edge feature will be inserted on *n* in the numeration. This leaves only the EPP-feature of *n* to attract Poss. Due to the order of operations in (16), this leads to bleeding of dative assignment and default nominative valuation. In the first scenario where an edge feature is inserted on *n*, the order in (16) predicts that this feature is discharged after case-Agree, resulting in dative on Poss. The dative Poss then moves to SpecDP to check D's structure-building feature. Since Poss can only bear dative if it checks an edge feature, and since the edge feature can only be present when D bears $[\bullet D_2 \bullet]$, it follows that only dative possessors can move to SpecDP. The nominative possessor cannot do so because D does not bear the $[\bullet D_2 \bullet]$-feature in the first place; otherwise, the possessor could not bear nominative. Hence, we do not have to state an explicit restriction on which possessors D can attract; that only dative possessors can be moved to SpecDP follows from *Phase Balance* and the order of operations.

An obvious question is why the possessor must check the edge feature once it is present. The EPP-feature of *n* is inherent, it co-occurs with the edge feature, so it could in principle also trigger movement of the possessor. This must be excluded because then

---

[13]The agreement facts between the possessor and the possessum are actually more complicated. There is a number effect: 3rd person pronominal possessors in the nominative cannot be marked for plural; similarly, pronominal possessors in the nominative cannot trigger plural agreement on the possessum. As for datives, there is dialectal variation for this effect depending on a number of factors. I do not attempt to derive these effects here as it is case Agree that is of central interest for the present discussion of opacity. See Kiss (2002) as well as den Dikken (1999) for an overview of the variation and an analysis of these phenomena.

[14]In contrast to the optionality discussed in section 2.5.3.2, it is Merge that is optional here and not Agree. The optional movement-inducing feature on D is not involved in the interaction of Merge and Agree that leads to opacity; opacity arises through the interaction of operations on *n*. Therefore, the optionality here cannot be derived by partial ordering; rather, the feature itself must be optional. Recall that simple optionality of a structure-building feature was untenable for empirical reasons for the languages with an optional reflex of movement discussed in the last chapter, not for general conceptual reasons.

the possessor would bear nominative and could nevertheless be moved to SpecDP. I do not have an explanation why this is so. I have to assume a constraint that demands that in case there are two features on a head H that could trigger movement to SpecHP and one of them is an edge feature, the edge feature must be used as a trigger. Recall that the same situation arises in Dinka on the vP, discussed in section 2.6.1.3: v has an inherent EPP-feature, but if a VP-internal phrase uses SpecvP as an intermediate landing site, it discharges an edge feature and not the EPP. This is evident because edge feature-driven movement to SpecvP has a morphological reflex in Dinka that pure EPP-driven movement has not; furthermore, the residual EPP can be discharged by a different XP (under certain circumstances) in this context. We will see another example of exactly the same abstract type in section 3.3 in the B-dialect of Icelandic. In all cases, it is the edge feature of a head H that must be discharged, and not its inherent EPP-feature, so the constraint seems to hold more generally and not only for Hungarian. In contrast to Dinka (and also to Icelandic B, as we will see), the residual EPP-feature on the edge feature-bearing head cannot be checked off by movement of another phrase to its specifier in Hungarian. The EPP-feature of n seems to be deleted as a last resort in Hungarian.

Finally, note that the version of *Phase Balance* in (148) in chapter 2 predicts edge feature insertion on n if D bears a structure-building feature, as required for the present analysis of opacity. As discussed in the analysis of Dinka where the same abstract configuration arises, edge feature insertion seems to be superfluous at first sight: The [•D•]-feature, inherently present on n, would enforce movement of Poss to the edge of nP anyway; this movement step is necessary for locality reasons (otherwise, Poss would not be accessible for the [•D•]-feature on D). Recall that the purpose of edge feature insertion is to get a phrase into the edge of a phase in order to keep the phrase accessible for a higher head. But since the EPP-feature of the phase head n can do the job, edge feature insertion should be superfluous and thus avoided. But it applies nevertheless. This follows since *Phase Balance* applies so early, viz. in the numeration. When n is selected, the system does not "know" that the phrase with the matching feature [D] for D's [•D•]-feature (the possessor) will be the one that the inherent EPP-feature of n would attract to Spec$n$P (which in turn would make edge feature insertion redundant). Therefore, *Phase Balance* predicts edge feature insertion, as desired (otherwise, Poss would always check n's EPP and we could not derive the opacity).

Let me summarize the analysis of the Hungarian case split. The abstract pattern of interaction is identical to the one in AAE languages: Intermediate movement steps counter-bleed Agree because they apply after Agree; final movement steps bleed Agree because they apply before Agree. The only difference between the AAE and possessor case in Hungarian is the functional head that triggers internal Merge and Agree, as well as the morphological reflex of the Agree relation. It is head-marking (agreement) in AAE languages and dependent-marking (case) in Hungarian.

3.2.2.3 Variation

Another instance of opacity with possessor extraction can be found in other Uralic languages for ϕ-Agree instead of case Agree (cf. Nikolaeva 2002): In Nenets (Samoyedic branch of the Uralic languages), for example, it is not the case value of the possessor that varies with final or intermediate movement of the possessor but ϕ-Agree between the possessor and the possessum: There is overt agreement on the possessum with so-called external possessors (cf. (20-a)) but not with internal possessors (cf. (20-b)):

(20)   *Possessive constructions in Nenets (Nikolaeva 2002: 277f.)*:

    a.   Wata-GEN tyuku˚ te-da
         Wata-GEN this   reindeer-3SG
         'this reindeer of Wata'   *external Poss + Agr*

    b.   tyuku˚ Wata-h   ti
         this   Wata-GEN reindeer
         'this reindeer of Wata'   *internal Poss without Agr*

    c.   *tyuku˚ Wata-h   te-da
         this   Wata-GEN reindeer-3SG
         'this reindeer of Wata'   *internal Poss + Agr*

    d.   tyuku˚ nye   ngøcyeki-h yetryi bantø-da   / *bant˚ -Ø ngarka
         this   woman child-GEN   always ribbon-3SG / ribbon-Ø big
         'The ribbon of this girl is always big.'   *external Poss, extracted from DP*

External possessors that trigger Agree on the possessum pattern with dative possessors in Hungarian: They precede the D element (compare (20-a) and (20-c)), and they can be extracted out of the DP (cf. (20-d)). Internal possessors that do not Agree with the possessum pattern with nominative possessors in Hungarian in that they follow the D element and cannot be extracted out of the DP (cf. (20-c) which is ungrammatical without agreement on the possessum). The case of the possessor is constant; it always bears genitive, regardless of how far the possessor moves. Nenets is thus very similar to Hungarian. The only difference is that in Nenets it is head-marking (agreement on the possessor) that reflects the difference between the two movement types, whereas in Hungarian is is dependent-marking (case on the possessor). The Nenets pattern is derived if the ϕ-probe and the case probe on n in Hungarian are interchanged:

(21)   *Ordering statement in Nenets*:
    $[*\text{CASE:GEN}*] \succ [\bullet\text{D}\bullet] \succ [*\phi{:}\square*] \succ [\bullet\text{EF}\bullet]$

As a consequence, the possessor is always assigned genitive because it is still in the base position when the case probe on n searches for a goal. A final movement step to Spec$n$P applies before ϕ-Agree and thus *bleeds* Agree. Hence, internal possessors do not show agreement on the possessum. External possessors that use Spec$n$P only as an intermediate landing site move after n initiated ϕ-Agree; Agree is thus *counter-bled*.

    Selkup (Samoyedic branch of the Uralic languages) exhibits a combination of the Hungarian and the Nenets system because both head-marking and dependent-marking are sensitive to the type of movement the possessor undergoes: Internal possessors neither have an overt case marker nor do they trigger agreement; external possessors bear an overt case marker (locative) and trigger agreement (Nikolaeva 2002: 277, fn.8). This is derived under the order in (22) where both Agree operations apply in between final and intermediate movement steps:

(22)   *Ordering statement in Selkup*:
    $[\bullet\text{D}\bullet] \succ [*\phi{:}\square*], [*\text{CASE:LOC}*] \succ [\bullet\text{EF}\bullet]$

As a consequence, final movement to Spec$n$P bleeds both ϕ- and case-Agree, whereas intermediate movement to Spec$n$P counter-bleeds both Agree relations. Thus, there is opacity of the pattern II type in several Uralic languages: Final movement steps apply before Agree, whereas intermediate movement steps apply after Agree ($[\bullet\text{F}\bullet] \succ [*\text{F}{:}\square*]$ / $[*\text{F:V}*] \succ [\bullet\text{EF}\bullet]$)

What about the other patterns? Languages without case splits could be analysed as having a pattern IV ordering: No movement type bleeds case-Agree because both final and intermediate movement steps apply after case-Agree. Note that that Hungarian ϕ-Agree behaves like that: It is never bled by movement because both types of internal Merge follow Agree. Other movement-related case splits where movement of the possessor always results in bleeding (occurrence of a less specific case marker under movement) are instances of pattern I if analysed as an ordering effect, see section 3.2.2.5: Both final and intermediate movement steps apply before Agree. I have not found data that illustrate pattern III (intermediate steps apply before Agree which in turn applies before the final step). It would be a language with a case split of the following type (under downward Agree): The DP bears default case only if it undergoes movement through the specifier of the head that assigns case to it, but it bears the regular case valued under Agree if it surfaces in the specifier of this head or if it is not moved at all.

3.2.2.4 Previous analyses

There are only few (more or less) explicit formal analyses of the case split in Hungarian. Let me briefly summarize what can be found: Szabolcsi (1994) also assumes that the possessor is merged as a complement of N and moves to the specifier of a functional projection Agr to receive nominative case. This is the low possessor position (Spec*n*P in the present analysis). Above Agr, there is another functional projection SpecDP to whose specifier the possessor can move (the higher position). This position is associated with dative case, assigned by D. Szabolcsi (1994: 203) suggests that the dative marker "is presumably not a real case marker here, however. On the one hand, the possessor is moved into SPEC of DP from an already case marked position [SpecAgrP where nominative is assigned, D.G.]." But she does not offer an alternative analysis of dative assignment. She speculates that the source of this morpheme is DP-internal, probably coming from D. The theoretical problem with dative assignment in SpecDP she alludes to in that quote is the following: The possessor always gets nominative case in the low position and movement to SpecDP feeds dative assignment. But due to the *Earliness Principle*, the DP is already case marked when it moves to SpecDP and should not be able to get dative case in addition. Case stacking might be a way to make it work, cf. footnote 12. Alternatively, nominative assignment in the low position must be suppressed in case the possessor moves up to SpecDP. This can also naturally be analysed as bleeding: Intermediate movement steps of the possessor bleed nominative case assignment. The possessor can then receive dative in SpecDP, probably from D. Whichever is ultimately the correct solution, if we do not want to assume case stacking or give up *Earliness* completely, bleeding seems to be involved in the derivation, and this is naturally accounted for by ordering of operation-inducing features.

A completely different approach is to deny that there is a case split in the first place. One such proposal is entertained by Gavruseva (2000): She assumes that the possessor only gets nominative case (in the specifier of a low functional projection, corresponding roughly to *n*P in my analysis). The dative marker is not a case suffix; indeed, it is the morphological reflex of checking a Q-feature on D, which is the trigger for possessor movement to SpecDP. So, there is actually no case split under this analysis. However, the argumentation is mainly theory-internal: DP is assumed to be the nominal counterpart of CP; dative assignment in SpecDP is assumed to be undesirable because "no case-checking usually takes place in the CP-like projection" (p. 759). Furthermore, in some languages the form of the determiner changes under possessor extraction (an indicator of Q-checking).

But since it does not change its form in Hungarian, the reflex must be found somewhere else; Gavruseva concludes that the locus is the possessor itself: the reflex is the "dative" suffix. Empirical arguments from Hungarian are lacking, however. So I do not follow this mainly theory-internal line of reasoning.

Another approach that denies the split is proposed by den Dikken (1999): He claims that the possessor is always the complement of a preposition. The "dative" suffix is analyzed as the overt realization of a preposition. However, the preposition can also be null ("nominative" possessor). In this case, it requires a special licensing movement to the specifier of a functional projection FP, an instance of predicate inversion. This movement is not necessary if the P head is overt. PPs with an overt preposition ("dative" possessor) can undergo movement to another position SpecDP, which is higher than FP, accounting for the fact that dative possessors are in a higher position than nominative possessors. Since there is no real case split, we do not get an *Earliness* problem with case assignment.

I will not evaluate these alternative proposals to the nominative-dative alternation here in any detail. It is fair to say that Gavruseva (2000) does not primarily aim at accounting for Hungarian possession, the scope of her paper is broader; and den Dikken's (1999) approach tries to account for a number of puzzling agreement facts across different dialects of Hungarian that I do not attempt to analyze. So, the papers have their merits regardless of how well they can handle the case split. The point is that if the case split on Hungarian possessors is taken to be real, a bleeding analysis is a straightforward option to circumvent the *Earliness* violation and case stacking which we get if nominative and dative are assigned by different heads in the structure. I will briefly elaborate on this general idea in the next section.

3.2.2.5   Extensions to other case splits

I have argued that the case split on Hungarian and Selkup possessors can be analysed as a consequence of the interaction of internal Merge and Agree: Movement bleeds case assignment. As long as movement is involved, the ordering approach can be transferred to other case splits. Often, however, case splits are not due to movement but are rather scale-driven: The case of an argument DP depends on the properties of the argument (local case split) or on those of its coargument (global case split). The choice of the case marker is driven by the location of these properties on a Silverstein scale (Hale 1972; Silverstein 1976), e.g. the person, animacy, or definiteness scale. In Hungarian, however, the case split is purely movement-related; the choice of the case marker does not depend on the animacy, definiteness, etc. of the possessor. Purely movement-related splits seem to be rare, but sometimes, a scale-driven split goes hand in hand with movement: It has been claimed that in some languages, there is a correlation between the definiteness of an XP and its structural position (inside or outside the verb phrase, cf. Diesing 1992). If there is a definiteness-driven split in these languages in addition, the split is correlated with movement, and consequently, it can be accounted for by the present approach. In Turkish, for instance, the case marking of the object depends on its definiteness / specificity. Definite objects take the overt accusative suffix -*(y)i*, whereas indefinites take a zero suffix (nominative). The definiteness of the object correlates (to a certain) extent with its structural position. Heusinger and Kornfilt (2005) show that objects that are moved out of the VP must bear accusative, whereas objects that are adjacent to the verb bear nominative; since a more marked case is assigned under movement (in contrast to Hungarian), the Turkish case split probably involves feeding of accusative assignment by movement, which probably involves upward Agree. There are certainly more factors that influence the case split

in Turkish, but it illustrates the type of interaction of movement and definiteness that could be analysed along the lines of the present approach. By the same logic, case alternations that arise as a consequence of grammatical function-changing processes could be reinterpreted as the result of rule ordering. The alternation between nominative and accusative on the internal argument of a transitive verb in the active / passive alternation could arise as follows: The head v assigns accusative to an argument in its c-command domain, in both the active and the passive. But in the passive, the argument moves out of the vP before v triggers case assignment. Accusative assignment is thus bled. Under the reverse order of operations, we would get accusative on the sole argument in the passive, which is indeed attested in a number of languages (e.g. Ukrainian). It would be interesting to see whether the choice between accusative and nominative correlates with movement of the sole argument of a passivized verb in some languages.

The ordering approach to movement-related case splits seems worth pursuing because it has advantages over alternative derivational analyses of the phenomenon: For example, one could also handle case splits by assuming that the case assigning head H is born with a case value X that is changed to or replaced by a case value Y in a certain syntactic context; however, this would violate *Inclusiveness* (cf. Chomsky 1995). Alternatively, the case split might arise because in certain contexts a structurally higher head W assigns a case value (cf. some of the previous analyses of Hungarian where it has been claimed that nominative is assigned by the equivalent of *n*, whereas dative is assigned by D). However, this is in conflict with *Earliness*: By cyclicity, the DP should be assigned the case by H because H is merged earlier than W. If one allows for case stacking, this might not be problematic; DP can be assigned the case by W in addition. But if we do not want to allow for case stacking, W should not be able to assign case after H. The ordering approach that treats case splits as the result of bleeding (or maybe also feeding) does not have these problems and is thus a viable alternative to the aforementioned approaches. However, a detailed investigation of other movement-related case splits is beyond the scope of this thesis; I will leave it for future research.[15]

---

[15]Let me briefly address two questions that should be pursued in this line of research: One questions would be if there is also evidence for different directions of Agree (upward and downward) in case splits? If Agree applies upwards, we get feeding instead of bleeding: Movement of an XP to SpecHP enables upward Agree of H with the XP in SpecHP. Upward Agree seems to be required for languages where movement of XP results in the occurrence of a more marked case on XP. Turkish accusative would be a candidate here. Another interesting question is whether there are movement-related splits where two overt case markers alternate, and not a zero marker and an overt marker (scale-driven splits of this type exist, e.g. Kolyma Yukaghir, cf. Maslova 2003). The ordering approach could handle them if case features are decomposed into bundles of binary features, e.g. into $[\pm\alpha]$ and $[\pm\beta]$. If each of them is assigned independently of the other by a head H, a movement operation triggered by H can be interleaved with them. Consequently, movement can feed / bleed assignment of only one of the two case features but not of the other, resulting in opacity (some case assignment operations are fed / bled by movement, but others are not):

(i)   *Order of operations on head H*:
       $[+\alpha]$-assignment ≻ movement ≻ $[+\beta]$-assignment

If we find interactions of this type they would provide an argument for different subtypes of (case) Agree, as well as the presence of decomposed case features in syntax.

3.3   (Counter-)feeding: defective intervention

In this section, we will see an opaque interaction of Merge and downward Agree that leads to (counter-)feeding instead of (counter-)bleeding. The crucial difference between the feeding and bleeding cases is the following: In the latter, the XP moved to SpecHP is also the XP that would be the goal for the probe on H; therefore, early movement of XP bleeds downward Agree with XP, as abstractly illustrated in 3.1. However, in the feeding case we will look at now, the XP that is moved to SpecHP is *not* the XP that H wants to Agree with; it wouldn't be a goal for H even if XP was not moved and remained in H's c-command domain. Rather, XP blocks Agree between H and a lower YP if XP is not moved at all. Consequently, movement of XP to SpecHP enables Agree between H and YP, a case of feeding.

3.3.1   Data: raising in Icelandic

In many languages, a dative XP blocks Agree relations between a probe on a head H and a lower DP if H c-commands both XP and DP, and if XP asymmetrically c-commands DP:

(23)   *Intervention configuration*:
$$[_{HP} \text{ H } [_{YP} \textbf{XP}_{dat} [_{Y'} \text{ Y } [_{WP} \text{ W DP }]]]]$$
└────── *Agree ──────┘

This looks like a classic intervention / minimality configuration: The dative XP is closer to H than DP. However, it is remarkable because the dative XP itself cannot value the probe on H and should does not count for the computation of minimality; nevertheless, it seems to intervene. This effect is known as *defective intervention* (Chomsky 2001: 123). There are various proposals why datives block Agree across them; I will not go through them here, see Preminger (2011*b*: ch.4) for an overview. For present purposes, it suffices to take it as a fact that datives block Agree in some languages. Anand and Nevins (2005: 6) formulate a parameter on the possibility of agreement with an inherently case-marked DP:

(24)   *The Visibility of Inherent-Case to Verbal Agreement Parameter (VIVA)*:
A language will differ as to whether the verb can agree with an inherently case-marked DP.

In Icelandic, the verb cannot Agree with dative DPs. Thus, dative experiencers (Exp), base-generated in SpecvP, are defective interveners for Agree between a ϕ-probe on the T head and a nominative DP in its c-command domain. This is illustrated in (25) with raising constructions. In (25-a), the raising verb c-commands both a nominative DP and a dative DP (set in bold). The dative DP is a closer goal but since it is a defective intervener, it blocks agreement between the verb (more precisely, the T head which bears the ϕ-probe feature) and the 3rd person plural nominative DP; in addition, the dative DP can also not be a target for the Agree relation. Hence, the ϕ-probe does not find a goal and consequently, the verb bears default 3rd singular agreement. In some dialects of Icelandic, the intervention effect is canceled with respect to number agreement if the dative experiencer moves out of the c-command domain of the T head, cf. (25-b) (I will come to person agreement below).

(25) *Raising constructions in Icelandic B (Holmberg and Hróarsdóttir 2003)*:

    a. það virðist/*virðast **einhverjum manni** [TP hestarnir   vera
       there seem.3SG/seem.3PL some   man.DAT   horses.DEF.NOM be
       seinir]
       slow
       'It seems to some man that the horses are slow.'

    b. **Mér**$_k$ virðast   ___$_k$ [TP hestarnir   vera seinir]
       me.DAT seem.3PL   horses.DEF.NOM be   slow
       'It seems to me that the horses are slow.'

    c. [**Hvaða manni**   ]$_k$ veist þú að virðist/*viðast   ___$_k$ [TP
       which man.DAT   know you that seem.3SG/seem.3PL
       hestarnir   vera seinir]
       horses.DEF.NOM be   slow
       'To which man do you know that the horses seem to be slow?'

Sigurðsson and Holmberg (2008) describe three Icelandic dialects that pattern differently with respect to the presence vs. absence of the intervention effect for number agreement in the context of experiencer movement. An opaque interaction of number Agree and internal Merge is found in the dialect Icelandic B, illustrated in (25-b) vs. (25-c) (see also Holmberg and Hróarsdóttir 2003 for a description of that dialect): Whereas movement of the non-wh-experiencer cancels the intervention effect (cf. (25-b)), wh-movement (and relativization) of the experiencer does not, cf. (25-c). This is opaque: In both cases the experiencer is no longer in the c-command domain of the matrix T-head on the surface and should thus no longer block Agree between T and the nominative DP in the infinitive. However, the wh-experiencer behaves as if it still intervened. Put differently, experiencer movement should feed Agree, but it does not if the experiencer undergoes wh-movement; wh-movement thus counter-feeds Agree.

    To understand the nature of the difference between (25-b) and (25-c), a few remarks on Icelandic syntax are in order. Icelandic has the V2-property: In main clauses, the finite verb V_{fin} occupies the second position and is preceded by exactly one constituent. This constituent can be any XP, an argument or an adjunct. It is often assumed that in V2-clauses, the finite verb moves to C and the constituent that precedes it is moved to SpecCP (though I will reject this assumption for subject-initial main clauses below); the latter movement is called topicalization. The V2 order XP-V_{fin} is also possible in most embedded clauses, except for embedded questions (Vikner 1995). It is not possible to topicalize a non-subject XP in an embedded question, cf. (26) (this will become relevant below when I discuss which movement step actually feeds Agree in (25-b)).

(26) *Topicalization in embedded questions (Vikner 1995: 74)*:

    *[CP Ég veit ekki [CP hvar í gær   hefur kýrin   staðið]]
      I know not   where yesterday has   cow.DEF stood
    'I don't know where the cow has stood yesterday'

In addition, there is evidence that Icelandic T has the EPP-property, i.e. that the subject of the clause must undergo movement to SpecTP, and that the finite verb moves to T (before it undergoes further movement to C in V2-clauses). We can see this in embedded questions as well: The finite verb in embedded questions must precede all VP-adverbs, indicating that the verb moves (at least) to T, cf. (27-a) vs. (27-b). However, the finite verb must not precede the subject, cf. (27-c). Since there is no topicalization in embedded

questions (cf. (26)), the latter indicates that the subject is in SpecTP and consequently, that the finite verb stays in T.

(27)   *Movement to SpecTP in Icelandic (Vikner 1995: 139)*:
   a.   [$_{CP}$ af hverju [$_{TP}$ Helgi hefði [$_{VP}$ oft   lesið þessa bók   ]]]
          why   Helgi had   often read this   book
   b.   *[$_{CP}$ af hverju [$_{TP}$ Helgi [$_{VP}$ oft   hefði lesið þessa bók   ]]]
          why   Helgi   often had   read this   book
   c.   *[$_{CP}$ af hverju hefi [$_{TP}$ Helgi [$_{VP}$ oft   lesið þessa bók   ]]]
          why   had   Helgi   often read this   book
       '(I asked) why Helgi had often read this book.'

With this in mind, let me come back to the contrast between (25-b) and (25-c). Assume for the time being that in V2-main clauses the verb is in C and the XP preceding it is moved to SpecCP. So there is both movement of the subject to SpecTP and of an XP to SpecCP in main clauses. In the two critical examples, it is the dative that undergoes both of these movements (under the current set of assumptions, but see the discussion below). Since the dative has subject properties (see e.g. Thráinsson 1979; Maling 1980; Zaenen et al. 1985; Sigurðsson 1989; Jónsson 1996), it can check the EPP on T; in addition, it undergoes topicalization because it precedes the finite verb. So the question is, which movement step of the dative experiencer actually feeds Agree – movement to SpecCP (topicalization) or movement to SpecTP (EPP-movement)? We can use embedded questions, which do not allow for topicalization, to test this. (28) shows that movement of a non-wh-experiencer to a position preceding the finite verb and following the wh-word enables Agree between T and the lower nominative. Since this can only be EPP-driven movement, we know that it is movement to SpecTP that has the feeding effect; subsequent movement to SpecCP (topicalization) is irrelevant for the Agree relation between T and a lower DP.

(28)   *Feeding in an embedded interrogative clause (Halldór A. Sigurðsson, p.c.)*:[16]

   Ég spurði [$_{CP}$ hvers vegna henni   virtust   myndirnar   vera ljótar]
   I   asked   why   her.DAT seem.PL.PST paintings.DEF.NOM to.be ugly
   I asked why the paintings seemed to be ugly to her

To summarize the discussion so far, movement of the experiencer to SpecTP can in principle feed Agree. Wh-movement clearly targets SpecCP; if every phrase is a phase, it must go through SpecTP as well (for independent evidence that SpecTP is an intermediate landing site for movement see the evidence from binding presented in section 2.1.2.2, and section 3.3.4 for evidence from variation). But movement of the wh-experiencer through SpecTP does not feed Agree in Icelandic B. Thus, opacity arises at the TP-level of the matrix clause, as depicted in (29) ('#' is the symbol for the number feature). In both (25-b) and (25-c), the experiencer undergoes movement to SpecTP, but it has has a feeding effect in only one of the two cases.

---

[16]Halldór A. Sigurðsson informs me that he is a speaker of the A-dialect, which I will turn to in section 3.3.4. But as for EPP-movement of a non-wh-experiencer in a main clause the A- and the B-dialect pattern alike; it's only for wh-movement of the experiencer that they differ. So it seems reasonable to assume that (28) also holds in the B dialect. Unfortunately, B speakers are a minority, and I did not get the required information from the B-speakers I contacted.

(29)   *Surface representation of the TP*:

a. Feeding (EPP-movement):

```
              TP
           /      \
        Exp_k      T'
                 /    \
            T [#:PL]   vP
                     /    \
                   t_k     XP
                          / \
                     … DP_[#:PL] …
```
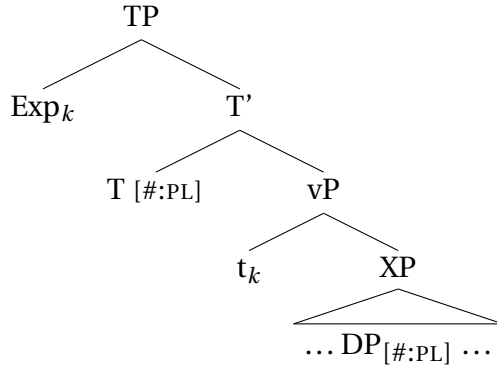
b. Counter-feeding (wh-movement):

```
              TP
           /      \
      wh-Exp_k      T'
                 /      \
          T [#:SG_def]   vP
                       /    \
                     t_k     XP
                            / \
                       … DP_[#:PL] …
```
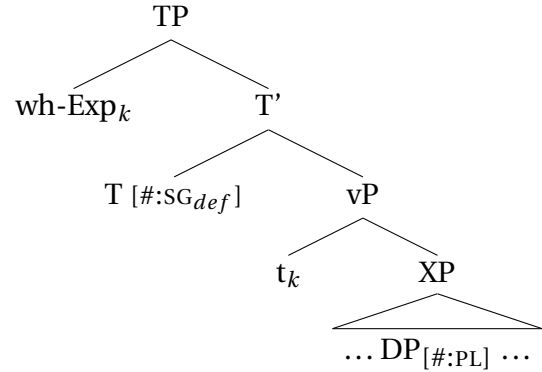
Before I close the descriptive part, let me briefly mention person agreement. So far, we have seen the feeding effect of movement to SpecTP only for number agreement. Sigurðsson and Holmberg (2008) show that person agreement patterns differently: Person agreement between T and the nominative DP is blocked in all three dialects of Icelandic, regardless of the position of the dative experiencer. There can only be default 3rd person agreement, even if the lower DP is 1st or 2nd person and the experiencer is moved out of the c-command domain of T (EPP-moved or wh-moved).

3.3.2   Analysis

What I want to propose is the following: The crucial difference between the feeding and the counter-feeding example with respect to number Agree is whether the experiencer checks the EPP-feature of T or not. If its does, this movement step to SpecTP is a final step; the experiencer will not move on to SpecCP. Such a final movement step feeds Agree (recall that movement of the experiencer to SpecTP in (28) is obviously final and triggers feeding).[17] If, however, the experiencer moves to SpecCP, it uses SpecTP only as an intermediate landing site and checks an edge feature on T instead of the EPP-feature. Such an intermediate movement step counter-feeds Agree between T and the lower DP. Thus, we have the same split of movement types as before: final vs. intermediate movement steps. But in Icelandic, the split has consequences at the TP-level (and not at the CP-, vP-, or nP-level). The number agreement facts follow from the order of operations in (30), where Agree splits up the two movement types (the EPP is represented as [•F•] because it can be checked by XPs of different categories, see the discussion of *Stylistic Fronting* below):

(30)   *A part of the ordering statement in Icelandic B*:
        $[\bullet \text{F} \bullet] \succ [* \#:\square *] \succ [\bullet \text{EF} \bullet]$

As a consequence, final movement steps apply early enough to take the experiencer out of the c-command domain of T. Consequently, T can Agree with the nominative DP, the experiencer does not intervene anymore, cf. the derivation in (31) (following the literature on Icelandic intervention effects, I take the experiencer to be base-generated in SpecvP).

---

[17]The EPP qualifies as a final feature because it is not present for pure locality reasons, inserted to satisfy *Phase Balance*; rather, it is an inherent structure-building feature of T in Icelandic.

(31)   *EPP-movement of the experiencer, full agreement*:

    a.   Step 1: Movement of the experiencer to SpecTP, EPP discharged

       $[_{\text{TP}}\ \square\ [_{\text{T}'}\ \text{T}_{\{\cancel{[\bullet\text{EF}\bullet]}\ \succ\ [*\#:\square*]\}}\ [_{\text{vP}}\ \boxed{\text{Exp}}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ t_{Exp}\ [_{\text{V}'}\ \text{V}\ [_{\text{TP}}\ \text{DP}_{[\#:\text{PL}]}\ [_{\text{T}'}\ ...\ ]]]]]]]]$

                 —— Move ——

    b.   Step 2: Agree between T and the lower DP

       $[_{\text{TP}}\ \text{Exp}\ [_{\text{T}'}\ \text{T}_{[*\#:\square*]}\ [_{\text{vP}}\ t_{Exp}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ t_{Exp}\ [_{\text{V}'}\ \text{V}\ [_{\text{TP}}\ \text{DP}_{[\#:\text{PL}]}\ [_{\text{T}'}\ ...\ ]]]]]]]]]$

               —— $\checkmark$ Agree ——

    c.   Result: ϕ-probe valued

       $[_{\text{TP}}\ \text{Exp}\ [_{\text{T}'}\ \text{T}_{[\#:\text{PL}]}\ [_{\text{vP}}\ t_{Exp}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ t_{Exp}\ [_{\text{V}'}\ \text{V}\ [_{\text{TP}}\ \text{DP}_{[\#:\text{PL}]}\ [_{\text{T}'}\ ...\ ]]]]]]]]]$

However, intermediate movement steps apply too late to feed number Agree: The experiencer is still in the c-command domain of T when it probes, and thus still intervenes. It is moved out of the c-command domain only *afterwards*, giving rise to opacity on the surface, cf. the derivation in (32).

(32)   *wh-movement of the experiencer, default agreement*:

    a.   Step 1: Agree initiated by T fails, Exp_{wh} still intervenes

       $[_{\text{T}'}\ \text{T}_{\{\,[*\#:\square*]\ \succ\ [\bullet\text{EF}\bullet]\}}\ [_{\text{vP}}\ \boxed{\text{Exp}_{wh}}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ t_{Exp}\ [_{\text{V}'}\ \text{V}\ [_{\text{TP}}\ \text{DP}_{[\#:\text{PL}]}\ [_{\text{T}'}\ ...\ ]]]]]]]]$

               —— *Agree ——

    b.   Result: default valuation, probe disacharged

       $[_{\text{T}'}\ \text{T}_{\{\,[\#:3\text{SG}]\ \succ\ [\bullet\text{EF}\bullet]\}}\ [_{\text{vP}}\ \boxed{\text{Exp}_{wh}}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ t_{Exp}\ [_{\text{V}'}\ \text{V}\ [_{\text{TP}}\ \text{DP}_{[\#:\text{PL}]}\ [_{\text{T}'}\ ...\ ]]]]]]]]$

    c.   Step 2: Intermediate movement step to SpecTP, edge feature discharged

       $[_{\text{TP}}\ \text{Exp}_{wh}\ [_{\text{T}'}\ \text{T}_{[\bullet\text{EF}\bullet]}\ [_{\text{vP}}\ t_{Exp_{wh}}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ t_{Exp}\ [_{\text{V}'}\ \text{V}\ [_{\text{TP}}\ \text{DP}_{[\#:\text{PL}]}\ [_{\text{T}'}\ ...\ ]]]]]]]]]$

               —— Move ——

The same result obtains when the experiencer is not moved to SpecTP at all: It intervenes between T and the lower nominative DP, cf. (33).

(33)   *No movement of the experiencer: default agreement*:

       $[_{\text{TP}}\ \text{T}_{[*\#:\square*]}\ [_{\text{vP}}\ \text{v}\ [_{\text{VP}}\ \boxed{\text{Exp}}\ [_{\text{V}'}\ \text{V}\ [_{\text{TP}}\ \text{DP}_{[\#:\text{PL}]}\ [_{\text{T}'}\ ...\ ]]]]]]$

               —— *Agree ——

To complete the picture, remember that person Agree is always blocked. This can be implemented if the person and the number probe are separate, each probing on its own, as suggested by Sigurðsson and Holmberg (2008) on the basis of the different behaviour of person and number with respect to intervention (see also Béjar 2003; Béjar and Řezáč 2009 for a separation of person and number Agree, though on a different empirical basis). I adopt their conclusion here. The distribution follows if person Agree applies before all movement operations to SpecTP, cf. (34) (the person probe is represented as π).

(34)   *A part of the ordering statement in Icelandic B (extended version)*:

       $[*\pi:\square*]\ \succ\ [\bullet\text{F}\bullet]\ \succ\ [*\#:\square*]\ \succ\ [\bullet\text{EF}\bullet]$

As a consequence, the dative experiencer always intervenes; it is still in its base position when the person probe initiates Agree. Any kind of movement comes too late to feed person Agree.[18]

Several issues arise for the present account. First of all, is is important that a wh-experiencer does not check the EPP on T. If it did, the movement would apply early and thus feed Agree, contrary to fact. Indeed, there is independent evidence that the wh-experiencer does not check the EPP-feature. Icelandic exhibits *Stylistic Fronting* (SF), an operation that moves the closest overt VP-internal category to SpecTP. *Stylistic Fronting* is optional. According to Maling (1980); Platzack (1987); Rögnvaldsson and Thráinsson (1990); Holmberg (2000), SF targets the subject position SpecTP (or its equivalent in previous analyses); Clements et al. (1983); Holmberg (2000); Ott (2009) argue that SF targets SpecTP because it applies to check the EPP-feature on T (see Ott 2009 for a detailed overview of the literature on SF, including tests that show that SF patterns differently from Ā-dependencies like topicalization). Evidence for this analysis comes from the observation that SF is possible whenever the canonical subject position SpecTP is not occupied (Maling 1980). This is the case if either the construction does not have a subject in the first place (impersonal constructions which would have a dummy subject without SF, see Clements 1984*b*: 4-5 for examples), if the subject is demoted (in the passive), or if the subject is indefinite and stays within the vP. In addition, SF can apply in embedded sentences if the subject position is emptied by Ā-movement of the subject (wh-movement, relativization, see Maling 1980 for examples). In (35), we see that SF cannot apply in the embedded clause if the subject is present:

(35)   a.   Verðbólgan varð verri en   ríkistjórnin   hafði búist   við
            inflation.DEF was worse than government.DEF have expectation had
            'Inflation was worse than the government had expected.'
       b.   *Verðbólgan varð verri en   *búist*   *hafði* ríkistjórnin   við
            inflation.DEF was worse than expectation have government.DEF had
            'Inflation was worse than the government had expected.'

However, SF with transitive verbs is possible if the subject is questioned (cf. (36-a) vs. (36-b)); but when the object is questioned, with the subject in SpecTP (following the questioned object but preceding the auxiliary in T), SF still cannot apply (cf. (36-c) vs. (36-d)):

(36)   *Stylistic fronting under wh-movement (Clements et al. 1983: 6)*:
       a.   Lögreglan veit ekki hver   hafi framið   glæpinn
            police.DEF know NEG who.NOM has committed crime.DEF
            'The police does not know who committed the crime.'   *SU question, no SF*
       b.   Lögreglan veit ekki hver   *framið*   hafi glæpinn
            police.DEF know NEG who.NOM committed has crime.DEF
            'The police does not know who committed the crime.'   *SU question, SF*
       c.   Lögreglan veit ekki hvað   glæpamaðurinn hafi framið
            police.DEF know NEG what.ACC criminal.DEF   has committed
            'The police does not know what the criminal has committed.'
                                                                    DO question, no SF

---

[18]In the present analysis, the effect follows from different ordering of the person and number probe on a single head. In Sigurðsson and Holmberg (2008), it follows from the location of the probes on different heads: The person probe bearing head is structurally higher than the number probe bearing head. Dative movement can target a position above the number probe but never above the person probe. It then follows that dative movement above the number head can feed number agreement, but not person agreement.

d. *Lögreglan veit ekki hvað *framið* glæpamaðurinn hafi
   police.DEF know NEG what.ACC committed criminal.DEF   has
   'The police does not know what the criminal has committed.'

DO question, SF

Hence, SF applies to check the EPP on T if the subject did not do so. Crucially, SF is possible in the main clause if the dative experiencer (that behaves like a subject for all relevant tests) is wh-moved. Holmberg and Hróarsdóttir (2003: 1009) provide the following example, where *Òlafur* is raised to the matrix clause to check the EPP-feature, as evident from the use of a compound tense (the raised DP is located between the finite verb in C and above the matrix raising verb):

(37)   Hverjum hefur Òlafur   virst   vera gàfaður
        who.DAT has   Olaf.NOM seemed be   intelligent
        'To whom has Olaf seemed to be intelligent?'

This supports the hypothesis of the present analysis that wh-experiencers do not check the EPP-feature on T. I proposed that it rather checks an edge feature; the residual EPP can be checked by movement of another phrase (SF).

Another question that arises in this context is the following: If there is a wh-experiencer, why is an edge-feature inserted on T in the first place? T inherently bears a Merge-trigger, the EPP. This feature would enforce movement of the wh-experiencer to the edge of TP anyway, as required by the PIC: Recall that the purpose of edge feature insertion is to keep an XP, here the wh-experiencer, accessible for an operation-inducing feature on a structurally higher head. So edge feature insertion should be superfluous to get the experiencer to SpecTP. Put differently, edge feature insertion should be bled by the presence of the EPP, but it isn't. Note that this is exactly the same situation we found in Dinka at the vP-level and in Hungarian at the nP-level: Although v and n bear an EPP-like feature (an inherent trigger for movement to their specifiers), movement through Spec$v$P and Spec$n$P checks an edge feature and not the EPP. As pointed out in the analyses of Dinka and Hungarian, the version of *Phase Balance* I propose (see (148) in chapter 2) correctly predicts the insertion of an edge feature in such a context: When T is selected for Merge in the numeration, there is still another head in the numeration, viz. C, that bears a structure-building feature [•WH•]; the experiencer with the matching feature is already part of the structure (merged in Spec$v$P), and therefore not accessible. Consequently, an edge feature will be inserted on T. The crucial assumption to enforce edge feature insertion here is that *Phase Balance* is checked early and does *not* only apply at the phrase level, after the head in question has discharged all of its inherent features (application at the phrase level is proposed in the version of *Phase Balance* by Heck and Müller 2003). If it did, the wh-experiencer would be in SpecTP due to the EPP, and edge feature insertion would be blocked. That it happens nevertheless provides evidence for a strictly derivational system without output-oriented constraints (early application of *Phase Balance*). A different question is why the wh-experiencer must not check the EPP, but has to check the edge feature on T when both are present. I don't know a good reason why this should hold, but it obviously does in Dinka, Hungarian and Icelandic. At the moment, I have to stipulate the preference for edge feature checking.

Having said this, a problem arises under the assumption that the there is always topicalization (movement to SpecCP) in main clauses. If this was the case, the C head of this clause would have a structure-building feature, call it [•TOP•]. It attracts a constituent with a matching feature [TOP]; assume that the non-wh-experiencer in (25-b) bears this

feature. As a consequence of [•TOP•] on C, *Phase Balance* predicts the insertion of an edge feature on T because the experiencer DP with [TOP] is not accessible (it is already part of the current phrase marker vP when T is selected in the numeration). But then T bears both an edge feature and an EPP-feature, the XP that ultimately moves to SpecCP checks the edge feature. But then wh-movement of an experiencer and topicalization should pattern alike: Both should counter-feed Agree because both undergo edge feature-driven movement to SpecTP; *Phase Balance* does not care whether the feature on C that causes edge feature insertion on T is a [•WH•]- or a [•TOP•]-feature. Note that this problem does not arise if a non-subject is topicalized, e.g. the direct object. In this case, edge feature insertion must apply, as predicted. The object checks the edge feature in SpecTP and the experiencer checks the residual EPP-feature on T. But if is is the experiencer that is topicalized, edge feature insertion must be prevented in order to get feeding. To solve this problem, I assume that subject-initial main clauses are TPs and not CPs, i.e. there is no topicalization in these clauses; if there is no C head with a [•TOP•], there will be no edge feature insertion on matrix T and the subject can only check the EPP, undergoing a final movement step. Indeed, it has been argued independently that subject-initial main clauses in Icelandic are smaller than CPs by Rögnvaldsson and Thráinsson (1990) (see also the discussion in Thráinsson 2007). Comparing the predictions of the two competing analyses of such clauses (CP-structure with subject in topic position SpecCP vs. TP-structure with subject in SpecTP), the authors conclude that there is at best no difference in empirical coverage. If anything, there is a slight advantage for the TP-approach, at least there is no clear evidence for the presence of a CP-shell. Note that the TP-analysis of subject initial main clauses has also been proposed for other Germanic languages with the V2-property, see Zwart (1993) on Dutch, and Travis (1984) on German and Yiddish (contra the CP-analysis first proposed by den Besten 1977).

Moreover, the agreement facts in Icelandic are instructive for the nature of Agree. First, since T can Agree with a nominative DP in an embedded infinitive if the experiencer is EPP-moved, it is clear that Agree is not subject to the same locality restrictions as (internal) Merge. Agree in Icelandic can apply over a number of intervening XPs, which movement cannot: Even though Agree does not cross a CP-boundary in the raising examples, it applies over at least one vP-boundary, and we have seen a number of reflexes of movement that suggest that (Ā-)movement must makes a stop-over in SpecvP. And if TP is a also a phase, as I assume, Agree crosses even more nodes that are locality boundaries for movement. For this reason, I assumed that Agree is not subject to the PIC (cf. section 2.5.2.1.2). Second, feeding in Icelandic shows us that the relevant cyclic domain (at least for Agree) is the XP, see the definition of the SCC in (143). To get feeding, the wh-experiencer must move to SpecTP *before* T can Agree with the lower nominative DP (and this is the standard assumption in the literature on feeding effects of this sort, see e.g. Holmberg and Hróarsdóttir 2003; Anand and Nevins 2005; Sigurðsson and Holmberg 2008; Asarina 2011). If every projection was a cyclic node, Agree between T and the nominative DP would be counter-cyclic because it only affects T′, but TP is already projected due to prior experiencer movement to SpecTP.

Note finally that Preminger (2011*a*) presents an argument against a timing approach to Icelandic B intervention effects: One could assume that EPP-movement of the dative experiencer to SpecTP applies before T triggers Agree, as I did in the present account. Therefore, EPP-movement feeds Agree of T with the lower nominative DP. However, Preminger (2011*a*: 7ff.) refutes the ordering approach based on the following observation: In clauses with nominative subjects and accusative objects, movement of the subject to SpecTP should then also feed Agree between T and the accusative object (lead-

ing to an ergative agreement pattern), which is, however, not the case. T enters into Agree with the nominative subject in this scenario. I would like to contend that this does not undermine the ordering approach because the unattested ergative agreement pattern is excluded for independent reasons: Crucially, it is the case that verbal agreement in Icelandic can always only target nominative XPs. Icelandic allows for subjects with quirky (i.e. non-nominative) case and lexical case on internal arguments, but agreement can never be with those non-nominative arguments. If none of the two arguments of a transitive verb is nominative (quirky dative subject and lexical case on the object), default agreement arises on the verb (see e.g. Thráinsson 2007: 166ff.). Whatever is the reason for this restriction in Icelandic, it explains why EPP-movement of a nominative subject cannot feed Agree between T and the accusative object, leading to an ergative pattern.[19]

3.3.3   Previous analyses

Recall that in Icelandic B wh-moved and EPP-moved experiencers behave differently with respect to their intervention potential. Holmberg and Hróarsdóttir (2003) propose the following: When looking at the surface structures in (25-b) and (25-c) one could state the descriptive generalization that NP-traces (left by EPP-movement of the experiencer) do not intervene, but wh-traces (left by wh-movement of the experiencer) do intervene. This is similar in spirit to the "quasi-operators" approach to pattern II, discussed in section 2.4.5.1: There are two types of traces that behave differently when it comes to Agree. As discussed in the aforementioned section, it is conceptually undesirable to increase the number of empty elements in Minimalism.

Indeed, Holmberg and Hróarsdóttir (2003) refine this analysis. They do not actually state constraints over different trace types. Rather, they claim that wh-experiencers do intervene because the wh-experiencer is still in the c-command domain of T when T probes. This idea is similar to my proposal, however, it is derived in a different way: In Holmberg and Hróarsdóttir's approach, wh-movement does not make a stop-over in SpecTP and hence, does not check the EPP; rather, wh-experiencers move directly to SpecCP. If it moved through SpecTP "[...] the relevant trace would not be a wh-trace, by usual assumptions" (Holmberg and Hróarsdóttir 2003: 998). According to the authors, wh-movement would leave a NP-trace in SpecTP and should thus also lead to feeding, which it doesn't; hence, wh-movement does not target SpecTP. Since C is merged later than T in a derivational system, T probes *before* a wh-experiencer can be moved to SpecCP. Consequently, the wh-experiencer is still in SpecvP when T probes and thus intervenes. EPP-movement happens before T probes and thus feeds Agree. As I will show in the next subsection, assuming that the wh-experiencer does not go through SpecTP causes problems when it comes to cross-linguistic variation of the interaction of movement and Agree in languages with defective intervention effects.[20]

---

[19] One can implement this restriction in the same way that I rules out agreement with direct objects in AAE languages under (local) subject extraction in section 3.2.1.2: Non-nominative XPs are opaque for agreement (cf. Keine's (2010) *agreement opaqueness*. Indeed, Preminger (2011*a*: 10) proposes essentially the same solution; he calls it *goal filtering*: The ϕ-probe on T can only target goals with a certain property (Relativized Minimality), here nominativ case). Alternatively, we could follow Preminger (2011*c*: 27) and say that that the restriction arises due to case filtration, i.e. the probe is picky, it can only Agree with an XP that bears a certain case value. Whether we use agreement opaqueness of goals or selectivity of probes, the outcome is the same: Only nominative XPs can be a goal for the ϕ-probe on T in Icelandic.

[20] For Holmberg and Hróarsdóttir's (2003)'s and my approach, the question is why wh-experiencers cannot discharge the EPP. If they could, this movement should cause feeding. Holmberg and Hróarsdóttir (2003) claim that this is the case because the EPP attracts only visible material, i.e. material that is already

In his discussion of defective intervention effects Řezáč (2004) also assumes that wh-movement does not go through SpecTP. With respect to the opaque Icelandic B data he states (Řezáč 2004: 65): "One thing that blocking of ϕ-Agree by an Ā-moved experiencer in (68) [= my (25), D.G.] shows is that Ā-extraction cannot be taking place from [Spec, TP]; an experiencer in [Spec, TP] permits ϕ-Agree by cyclic displacement [...]. So something must require Ā-extraction from the in situ position of the experiencer, or its successive-cyclic movement position at the edge of vP." Both Holmberg and Hróarsdóttir (2003) and Řezáč (2004) claim that movement through SpecTP cannot take place because then the wh-experiencer could not be distinguished from a non-wh-experiencer which checks the EPP on T; both would be in the same position SpecTP, but they do not pattern alike with respect to the possibility of creating new Agree relations. I propose that once we shift the attention from the moved elements to the head that triggers the movements, this problem vanishes. We cannot distinguish the wh- and non-wh-experiencer by the *position* they occupy, as these authors correctly point out (this is the output-oriented view: When looking at TP, the two constructions look the same, cf. (29)). However, we can distinguish wh- and non-wh-experiencers by their triggers on T, as the present analysis does (the strictly derivational view). If final and intermediate movement steps have distinct triggers, they can apply at different points of the derivation, with exactly the same result that wh-experiencers still intervene when T probes – although they do go through SpecTP. I conclude that the Icelandic B data do not necessarily show that wh-experiencers do not target SpecTP. This is a desirable result, as I will show in section 3.3.4.

Another analysis is proposed by Chomsky (2008: 152f.). The asymmetry between experiencer movement to SpecTP (A-movement) and movement to SpecCP (Ā-movement) can be derived under parallel Merge. Both A and Ā-movement are triggered by operation-inducing features on C. C triggers internal Merge of operators to SpecCP; in addition, when C is merged with TP, T inherits another operation-inducing feature from C that is responsible for attraction of a DP to SpecTP (the EPP-property). In case we have a wh-experiencer, it undergoes A-movement to SpecTP and Ā-movement to SpecCP *in parallel*, i.e. the wh-experiencer in its base position SpecvP is the foot of an A-chain whose head is in SpecTP, and it is the foot of an Ā-chain whose head is in SpecCP. Unlike Holmberg and Hróarsdóttir 2003, Chomsky does not want Ā-movement to skip SpecTP (for independent reasons); parallel movement guarantees that there is a representation of the wh-experiencer in SpecTP under Ā-movement to SpecCP, but without movement from SpecTP to SpecCP. By assumption, the copy of the experiencer becomes invisible as a goal for Agree if all members of the chain it is contained in are in an A-position. This is the case if the experiencer only moves to SpecTP (non-wh-experiencer), giving rise to feeding. If, however, parallel movement applies to a wh-experiencer, the copy in the base position is part of an A-chain, but it is also part of a chain whose head is in an Ā-position (due to

---

spelled out. A category is spelled out as soon as all of its uninterpretable features are checked. For NPs, this is fulfilled if they are assigned case. Hence, an NP-experiencer is visible to the EPP on T because its case feature is interpretable (experiencers bear inherent case). A wh-experiencer, however, is not visible to T because it bears an uninterpretable wh-feature (called 'c' in the paper) that can only be checked by C. I do not adopt this proposal because it has conceptual problems: It is not clear to me how something that is spelled out can be visible for syntactic operations, while something that is not spelled out is invisible. If spell-out means sending a constituent to the postsyntactic PF component, we would expect it to be exactly the other way around: Spelled out material should no longer be accessible to syntax. Likewise, if one assumes the *Activity Condition*, it is elements with *un*valued features that are visible for syntactic operations and not those with valued features; a case-marked phrase like the experiencer should thus be invisible, the opposite of what is assumed in Holmberg and Hróarsdóttir's approach. Though I do not have a better explanation for why the wh-phrase cannot check T's EPP-feature, I do not adopt their proposal.

parallel movement). Therefore, the copy of the wh-experiencer is visible for Agree and thus counts an an intervener. Apart from the fact that feature inheritance from C to T is counter-cyclic (though there might be reasons to adopt it, see Richards 2007), this analysis requires that whole movement chains are taken into account in order to be able to determine whether the lower copy of the experiencer intervenes: The copy of the experiencer in its base position is exactly the same under both pure A-movement and Ā-movement (involving parallel movement). What one has to look at is the nature of the landing site of the moved phrase (A- or Ā-position) and whether all chains the copy is part of are uniformly A-chains. It must thus be possible to scan a certain amount of structure after the movement operations in question have taken place. One must wait until the CP is built in order to determine whether a copy of the experiencer in SpecvP intervenes for Agree between T and a lower DP. The analysis thus has a representational character. Furthermore, if every XP is a cyclic node, Agree must apply in a counter-cyclic manner because it affects only the TP-cycle that is already dominated by CP. In the present approach, Agree is cyclic and we do not have to wait until the CP is projected to determine whether the experiencer intervenes or not.[21]

Finally, Sigurðsson and Holmberg (2008) suggest a timing account. They assume the clause structure in (38) with two functional heads above TP: Pn bears the person probe and Nr bears the number probe.

(38) $[_{Pn} [_{Nr} [_T [_{vP} \text{Dat ... Nom}]]]]$

The experiencer can target the specifier of the number head. In order to derive feeding of number Agree, they assume that movement of the experiencer to SpecNrP can apply *before* Nr probes. Hence, early movement gives rise to feeding, as in the present approach. With respect to the intervention of wh-experiencers, Sigurðsson and Holmberg (2008: 268) speculate that movement of the experiencer to SpecNrP applies *after* Nr probes and therefore, the experiencer still intervenes. This is similar in spirit to my proposal: Wh-experiencers and non-wh-experiencers move at different points relative to Agree. However, the analysis is not worked out. It is not shown how the status of the experiencer (wh/non-wh) influences the order of operations in the NrP. If the experiencer is a wh-word, Nr must probe first, if it is a non-wh-word, Nr must trigger Merge first. But this is not formally implemented, no complete derivations are shown for either wh- or non-wh-experiencer movement in Icelandic B. As it stands, the analysis is contradictory because it is assumed that both Merge before Agree and Agree before Merge hold for the same head; the conditions under which the order is chosen are not identified. In the present approach, the contradiction is canceled by postulating two distinct Merge triggers for final and intermediate movement steps, and these can be ordered are ordered differently relative to Agree.

3.3.4   Variation

On the surface, the interaction of number Agree and internal Merge in Icelandic B is opaque because, when looking at the output structure, neither the EPP-moved nor the wh-moved experiencer intervenes between the matrix T and the embedded subject; nevertheless, the wh-experiencer behaves as if it intervened. Derivationally, this could be de-

---

[21]Interestingly, Chomsky (2008: 153) suggests at the end of the discussion of the parallel movement account to Icelandic B that "[o]ne might explore whether the variation in judgements can be attributed to the timing of the edge and Agree probes.", which is exactly what the present account explicitly explores.

rived by assuming that the wh-experiencer moves directly to SpecCP without a stop-over in SpecTP (cf. the analyses by Holmberg and Hróarsdóttir 2003; Řezáč 2004 introduced above). Under this analysis, it is clear why the wh-experiencer causes an intervention effect: When T probes, the wh-experiencer is still in the c-command domain of T and thus blocks Agree. It is moved to SpecCP only after C is merged. This analysis might seem simpler than the approach I have proposed. However, I will argue in this subsection that it is problematic to make this assumption with respect to cross-linguistic variation in defective intervention, both within Icelandic and outside of Icelandic: In another Icelandic dialect, Icelandic A, *both* wh-movement and EPP-driven movement of the experiencer feed number Agree, i.e., (25-c) is grammatical with 3rd person plural agreement on the verb (see Sigurðsson and Holmberg 2008). Under the assumption that wh-movement does not go through SpecTP, this feeding effect is mysterious: How can the wh-phrase feed Agree if it is still in the c-command domain of T when T probes? Wh-movement is triggered by C at a later stage of the derivation.

A similar configuration, though not involving a (visible) ϕ-Agree relation, holds in Romance languages and Greek (see Rizzi 1986*b*; Torrego 1996; McGinnis 1998; Boeckx 1999; Anagnostopoulou 2003; Hartman 2012). In these languages, an experiencer blocks movement of (instead of ϕ-Agree with) a lower nominative DP to SpecTP. This movement is necessary to check the EPP on T. I illustrate this for French in (39-a) (the intervener is set in bold, *vous* is the nominative subject). The experiencer itself cannot check the EPP, cf. (39-b). But if the experiencer is moved to T (cliticized to T in the case of French, cf. (39-c)), or if it is Ā-moved (cf. (39-d)), the lower DP can be moved to SpecTP.

(39)   *Defective intervention in French (Řezáč 2004: 60)*:

    a.   Vous   semblez (* **à Obelix**) être   drôles
        you.PL seem   to Obelix   to.be funny.
        'You seem to Obelix to be funny.PL'

    b.   ***A Obelix** semblez vous   être   drôles
        to Obelix seem   you.PL to.be funny.
        'You seem to Obelix to be funny.PL'

    c.   Vous   **lui** semblez être   drôles
        you.PL him seem   to.be funny.PL
        'You seem to him to be funny.'

    d.   le   Gaulois **à qui**   vous   semblez être   drôles
        DEF Gaul   to whom you.PL seem   to.be funny.PL
        'The Gaul to whom you seem to be funny.'

For languages with this pattern, the same problem arises as in Icelandic A: If Ā-movement directly targets SpecCP and does not go through SpecTP, how can Ā-movement to SpecCP feed Agree initiated by T? The experiencer does intervene when T probes. Several solutions to this problem have been proposed in the literature for Greek and the Romance languages:

  (i)   phase-internal counter-cyclicity (Anagnostopoulou 2003)
  (ii)   intervention is evaluated at the phase-level (McGinnis 2001)
 (iii)   covert movement of the wh-experiencer to a low Ā-position (Legate 2002)

Solution (i) gives up the Strict Cycle Condition within a phase: Phases are assumed to be CP and vP. Within a phase, any order of operations is allowed, the Cycle only holds for operations across phases. Thus, Agree between T and the lower DP can take place *after*

the wh-experiencer has been moved to SpecCP. Agree only affects a subtree of the current CP, namely TP; but since CP and TP are in the same phase, this is licit. According to solution (ii), intervention is evaluated at the phase level, after all operations in a phase have applied. This means that minimality is read off of the surface structure of an expression. When looking at the structure *after* Ā-movement of the wh-experiencer within the CP-phase, the experiencer does not intervene anymore between T and the lower DP and the prior Agree relation is thus legitimate. Finally, solution (iii) resorts to a covert movement step: Operator phrases are moved to a low Ā-position between VP and TP. When T probes for A-features (ϕ-Agree, EPP-checking) it ignores elements in Ā-positions. All of these solutions have shortcomings: Solution (i) gives up a core constraint on derivations, viz. the Strict Cycle Condition. Solution (ii) delays evaluation of constraints and thus needs a highly representational version of the Minimal Link Condition. Both of them are thus not compatible with a strictly derivational model of grammar. The movement step proposed in solution (iii) is not motivated independently.

In order to avoid all these problematic assumptions one could propose the following solution: The different feeding patterns in Icelandic B vs. Icelandic A / Romance are a consequence of the different locality domains in these languages: In Icelandic B-type languages, wh-movement does not make a stop-over in SpecTP, which accounts for the absence of feeding with wh-experiencers. In languages of the Icelandic A- / Romance / Greek type, however, wh-movement must proceed in smaller steps and does go through SpecTP, explaining the feeding effect with any kind of movement. The question is whether we want to assume that languages differ in such a fundamental property like the locality domains for movement.

Under the present ordering approach to opacity, another solution is possible that (a) neither weakens the Strict Cycle Condition, nor (b) invokes strongly representational constraints (evaluated at the end of the derivation), nor (c) postulates different locality domains in different languages. The core idea is again that variation simply arises by reordering of operation-inducing features. In contrast to Icelandic B, the EPP and the edge feature are ordered *before* the number probe feature in Icelandic A and Romance/Greek, cf. (40):

(40)   *A part of the ordering statement in Icelandic A*:[22]
        $[*\pi{:}\square*] \succ [\bullet F\bullet], [\bullet EF\bullet] \succ [*\#{:}\square*]$ *pattern I order*

The experiencer, whether it is a wh-phrase (using SpecTP as an intermediate landing site) or not (using it as a final landing site) moves early to SpecTP and does not intervene anymore at the point when the number probe on T seeks for a goal. Thus, both types of movement feed number Agree between T and the lower DP.[23] In Icelandic B, the two movement triggering features align differently with respect to the probe feature, see the order repeated in (41).

---

[22]Note that person agreement is blocked in all Icelandic dialects, regardless of whether the experiencer undergoes wh-movement or EPP-movement. This follows if person Agree applies before any type of movement, as indicated in (40).

[23]In section 4.2 I will show that Italian differs from the other Romance languages mentioned above in that it also exhibits opacity of the Icelandic B type in a special context: Cliticization of the experiencer always feeds EPP-movement to SpecTP, as in other Romance languages. Certain types of Ā-movement do so as well, again as in other Romance languages. However, relativization does not lead to feeding, just like Icelandic Wh-movement does not, hence this is a case of counter-feeding.

(41)   *A part of the ordering statement in Icelandic B*:
         [∗π:□∗] ≻ [•F•] ≻ [∗#:□∗] ≻ [•EF•]   *pattern II order*

This gives rise to opacity: Only final EPP-movement feeds Agree, intermediate movement steps come too late to feed it.

What about the other permutations of operation-inducing features? There is a third dialect of Icelandic described in Sigurðsson and Holmberg (2008), Icelandic C. In this variety, EPP movement does not feed number Agree. Hence, final movement must apply after number Agree, in contrast to Icelandic A and B. It is not clear however, whether wh-movement feeds number Agree (Halldór A. Sigurðsson, p.c.). If it feeds Agree, we need the order in (42-a) (intermediate steps apply before number Agree); if it does not feed Agree, we need the order in (42-b) (intermediate steps apply after number Agree).

(42)   *Part of the ordering statement in Icelandic C*:
         a.   [∗π:□∗] ≻ [•EF•] ≻ [∗#:□∗] ≻ [•F•], or   *pattern III order*
         b.   [∗φ:□∗], [∗#:□∗] ≻ [•F•], [•EF•]   *pattern IV order*

In Romance languages, Greek, and Icelandic A, movement interacts transparently with number Agree: Movement (both final and intermediate) always feeds Agree. In Icelandic B, we get an opaque interaction: Movement sometimes feeds Agree and sometimes it does not. Note that the languages in which wh-movement of the experiencer feeds Agree require that wh-movement goes through SpecTP if we do not want to give up strict cyclicity for Agree. If this holds universally, then the Icelandic B data are indeed opaque because both wh- and non-wh-experiencers land in SpecTP, but with different consequences for Agree. Reordering of operation-inducing features provides a simple way to account for the variation in the cancellation of intervention effects between Icelandic A / Romance / Greek and Icelandic B without assuming different locality domains for wh-movement in different languages. Furthermore, the split in movement types the analysis requires is exactly the same I argued for in chapter 2 for upward Agree. We thus expect it to be be active with downward Agree as well. The approach has the additional advantages that standard locality constraints like the Strict Cycle Condition are obeyed and that the strongly representational version of the MLC, evaluated at the end of the derivation can be avoided.

3.4   Interim conclusions

In this chapter I have shown that internal Merge can interact opaquely with downward Agree when both operations are triggered by the same head H: Some movement steps to SpecHP bleed downward Agree whereas others counter-bleed it. Empirical evidence suggests that the split of movement types is the same that we found for upward Agree in chapter 2: final vs. intermediate movement steps. Under the ordering approach, movement steps that apply before Agree bleed Agree, whereas movement steps that apply after Agree counter-bleed it. However, movement cannot only have a (counter-)bleeding effect but also a (counter-)feeding effect on downward Agree. The former arises if the XP moved to SpecHP is also the potential goal of a probe on H. (Counter-)feeding arises if an intervening XP moved to SpecHP usually blocks Agree between H and a YP c-commanded by XP. Reordering of operation-inducing features derives the range of variation found cross-linguistically (cf. the summary in table 3.1).

	pattern I order	pattern II order	pattern III order	pattern IV order
AAE	Fiorentino Trentino Ibibio	Berber Celtic	–	languages without AAE (e.g. French)
case split	Turkish	Hungarian Selkup	–	Nenets
defective intervention	Icelandic A Romance, Greek	Icelandic B	Icelandic C (?)	

Table 3.1: Summary: Interactions of internal Merge and downward Agree

Moreover, the ordering approach provides a uniform analysis for a number of seemingly unrelated phenomena in addition to the reflexes of movement treated in chapter 2: Anti-agreement in Berber type-languages, movement-related case splits in Hungarian, and cancellation of intervention effects in Icelandic B. What they all have in common is that final and intermediate movement steps apply at different points relative to Agree, giving rise to opacity.[24]

---

[24]In contrast to the interaction of Merge and upward Agree, I did not find an example of pattern III for the interaction of Merge and downward Agree, cf. table 3.1. In Georgi (2013), where I was solely concerned with interactions of the latter type, I hypothesized that the absence of pattern III is due to Specificity (for application of this concept to syntax see Sanders 1974; Pullum 1979; Lahne 2012; van Koppen 2005), according to which more specific rules apply before less specific rules (see section 5.2 for extensive discussion of principles that determine rule ordering). With respect to the order of operation-inducing features this means that the more specific internal Merge triggering feature must be discharged before the less specific one. In Georgi (2013) I suggest that features that trigger final movement steps are more specific than edge-features because final features attract elements of a certain category or with a certain effect on interpretation (e.g. [•D•], [•WH•], [•TOP•]), whereas edge features are categorially underspecified structure-building features that do not ask for certain properties of the XP they attract. Hence, an order in which an edge feature is discharged before a non-edge feature, required to derive pattern III, would be excluded by Specificity. Since there is pattern III for upward Agree, this cannot hold universally. Ideally, we will find instances for the predicted pattern III order for the phenomena discussed in the present chapter, which would mean that Specificity plays no role for ordering of operation-inducing features; or, if PIII is indeed unattested for downward Agree, I would be forced to say that Specificity regulates the order of operation-inducing features only in some languages or in some constructions but not in others, clearly a less elegant solution. For the time being, I take the absence of examples for PIII with downward Agree to be an accidental gap.

Chapter 4

Extension to other types of Merge

In this chapter I investigate whether we need to distinguish even more subtypes of Merge. So far, I have solely considered cases where Agree interacts with internal Merge. I will argue that the same configuration where Agree is interleaved with two structure-building operations (producing opacity) is also found with other types of Merge. In section 4.1 I present examples that suggest a split between internal and external Merge, and in section 4.2 I argue for a split between different types of final and intermediate internal Merge. Section 4.3 pursues the question whether a fifth pattern of reflexes of upward Agree, not considered so far, provides evidence for yet another type of internal Merge.

4.1   External vs. internal Merge

In this section I show that external and internal Merge can apply at different points in the derivation relative to Agree, in either order. The conclusion is the same as before: If all Merge operations were triggered by the same type of structure-building feature, they could not be interleaved with Agree. Therefore, external and internal Merge triggers need to be distinguished.

4.1.1   Internal > external Merge

I have argued in section 2.6 for a split of Merge into external vs. internal Merge on the basis of the abstract pattern in (1):

(1)   *Opaque interaction of external and internal Merge with upward Agree*:
        [•EF•] ≻ [∗F:□∗] ≻ [•D•]

The consequences of this order are as follows: If Agree applies upwards, internal Merge feeds Agree but external Merge counter-feeds Agree. We have seen empirical evidence for this split in section 2.6.1 from Duala and Wolof. These languages exhibit a subject/non-subject asymmetry. In Duala the reflex of movement only occurs if a non-subject is extracted. In Wolof the default class marker, which surfaces if class Agree with C is counter-fed, cannot appear if the extracted element is the subject of the clause. I analysed this split as follows: The crucial difference between subject and non-subject extraction is that non-subjects are moved through SpecvP, whereas subject are externally merged in SpecvP. Under the order of operations in (1), internal Merge to SpecvP can feed Agree with v; external Merge comes too late, it counter-feeds Agree. I proposed that Agree involves valuation of the category of the v-head on the XP moved to SpecvP. This information is passed on

from the moving XP to the C-head under category Agree at the CP-level. The VIs *no* in Duala and *l-* in Wolof are sensitive to the presence of the category feature [CAT:V] on the C head. Therefore, only non-subjects can trigger *no-* and /*l*-insertion, respectively, when they undergo Ā-movement.

4.1.2   External > internal Merge

In this subsection I argue that there is also evidence for an opaque interaction of external / internal Merge and Agree with the reverse order of Merge operations, cf. (2):

(2)   *Opaque interaction of external and internal Merge with upward Agree*:
        [•D•] ≻ [∗F:□∗] ≻ [•EF•]

The argument is based on the analysis of the *ban on ergative movement* in Assmann et al. (2012). Let me first introduce the phenomenon. In many morphologically ergative languages the external argument of a transitive verb (the ergative DP) cannot be Ā-moved, whereas the internal argument of a transitive verb and the sole argument of an intransitive verb can, cf. the data from wh-movement in Kaqchikel (Mayan) in (3):

(3)   *Wh-movement of DP_{erg} vs. DP_{abs} in Kaqchikel*:
        a.   N-Ø-u-löq'   jun sik'iwuj ri   a   Karlos
             INCOMPL-3SG.ABS-3SG.ERG-buy INDEF book   DET CLT Carlos
             'Carlos buys a book.'
        b.   Atux n-Ø-u-löq'   a   Karlos
             Q   INCOMPL-3SG.ABS-3SG.ERG-buy CLT Carlos
             'What does Carlos buy?'
        c.   *Achike n-Ø-u-löq'   jun   sik'iwuj
             Q   INCOMPL-3SG.ABS-3SG.ERG-buy INDEF book
             'Who buys a book?'

(4)   *Wh-movement of DP_{abs} in Kaqchikel*:
        a.   N-Ø-tze'en   a   Karlos
             INCOMPL-3SG.ABS-laugh CLT Carlos
             'Carlos laughs.'
        b.   Achike (ri)   n-Ø-tze'en
             Q   DET INCOMPL-3SG.ABS-laugh
             'Who laughs?'

There is no corresponding extraction asymmetry in morphologically accusative languages. Assmann et al. (2012) propose an analysis of the ban on ergative movement and the asymmetry between ergative and accusative languages that makes crucial use of the timing of Merge and Agree (but they do not distinguish between different subtypes of Merge). The core idea is that the ban on ergative movement arises because when the ergative DP undergoes movement, it moves *too early* and thereby absorbs the case value (absolutive) that was originally provided for its co-argument DP_{int}; hence, movement of the ergative DP *bleeds* case assignment to DP_{int}. The assumptions of the analysis are the following:

(i) T assigns the morphologically unmarked external case [∗CASE:EXT∗] (nominative / absolutive) and v assigns the morphologically marked internal case [∗CASE:INT∗] (ergative / accusative).

(ii) In an intransitive context, only T is active and hence, the sole argument will get the unmarked case.

(iii) Case filter: Every DP needs to be assigned a case value (no default case assignment).

(iv) Case is assigned under Agree.

(v) Spec-Head bias: A head can in principle Agree with a DP in its specifier (upward Agree) and with a DP in its c-command domain (downward Agree), but in case there is a potential goal for the head H in its specifier and in its c-command domain, there is a preference for Agree with the specifier (viz. upward Agree is preferred over downward Agree).

(v) There is no minimality condition o Agree: If there are several goals in the probing domain of a head, the head can freely choose a goal.

(vii) DPs can be assigned more than one case value (case-stacking, only the case assigned first is morphologically realized).

(viii) Every phrase is a phase and movement to the edge of a phase is triggered by edge features.

(ix) If a head triggers both Merge and Agree, the operations are ordered (language-specifically).

Morphological ergativity is derived as follows: v triggers both external Merge of DP_{ext} and case assignment (Agree): v { [•D•], [∗CASE:INT∗] }. Hence, the operations need to be ordered. Ergativity arises if Merge applies before Agree (cf. Müller 2009).[1] When v is merged with VP (which contains DP_{int}), it first triggers external Merge of DP_{ext}. Afterwards, it triggers case assignment (Agree). Due to the Spec-head bias, the marked internal case is assigned to DP_{ext} in SpecvP and not to DP_{int} in v's c-command domain, cf. (5-a) (solid lines indicate case assignment, dashed lines indicate movement). Hence, only the unmarked external case from T remains for DP_{int}. Since in intransitive contexts only T assigns case, DP_{int} ends up with the same case as the sole argument of an intransitive verb; DP_{ext} bears a different case value, viz. we get morphological ergativity. Let me now turn to the extraction asymmetry in ergative languages. If an argument is Ā-moved to SpecCP, it must make a stop-over in SpecTP because every phrase is a phase. Thus, T must bear an edge feature. In addition, T is a case assigner. Consequently, the same conflict arises on T that we encountered on v: T triggers both Merge (internal Merge) of a DP and Agree (case assignment): T { [•EF•], [∗CASE:EXT∗] }. Assmann et al. (2012) assume that the order of operations on T is the same as the order on v, i.e. Merge before Agree in a morphologically ergative language. This order has the following consequences on the TP-cycle: If DP_{ext} is extracted, it is internally merged to SpecTP *before* T triggers Agree (assignment of the internal case), cf. (5-b). Due to the Spec-Head bias, DP_{ext} in SpecTP will be assigned the external case by T, although it already bears the internal case assigned by v (recall that case-stacking is not prohibited), cf. (5-c). But this means that there is no case value left for DP_{int}, violating (iii). Consequently, the derivation crashes because early movement of DP_{ext} bleeds case assignment to DP_{int}.

---

[1] In an accusative language, the order is Agree before Merge, both on v and on T. I will not talk about accusative languages and the absence of a corresponding ban on accusative movement; the reader is referred to the detailed derivations in Assmann et al. (2012).

(5)   *Illegitimate movement of DP_{erg}:*

    a. Case assignment to DP_{ext} in the vP

```
                        TP
              ┌──────────┴──────────┐
              T                     vP
        ┌              ┐      ┌──────┴──────┐
        │ [•EF•]       │    DP_ext          v'
        │ [*CASE:EXT*] │  [CASE:INT]    ┌────┴────┐
        └              ┘                v        VP
                                   [*CASE:INT*] ┌─┴──┐
                                                V   DP_int
                                                   [CASE:☐]
                                          ❶
```
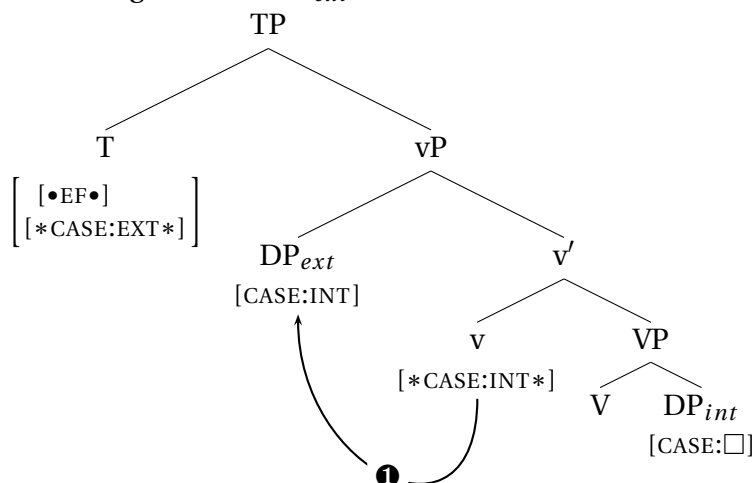
    b. Merge before Agree on T triggers movement of DP_{erg} first

```
                        TP
              ┌──────────┴──────────┐
            DP_ext                  T'
         [CASE:INT]           ┌─────┴─────┐
                              T           vP
                        [*CASE:EXT*]  ┌───┴───┐
                                      t       v'
                                          ┌───┴───┐
                                          v       VP
                                     [*CASE:INT*] ┌─┴──┐
                                                  V  DP_int
                                    ❷                [CASE:☐]
```
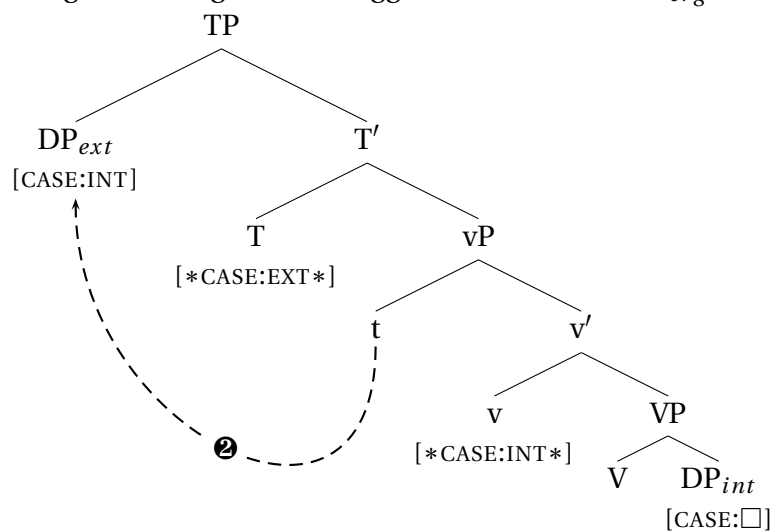
    c. Specifier-head bias triggers case assignment to DP_{ext} in SpecTP

```
                        TP
              ┌──────────┴──────────┐
            DP_ext                  T'
         [CASE:INT]           ┌─────┴─────┐
                              T           vP
                        [*CASE:EXT*]  ┌───┴───┐
                                      t       v'
                           ❸              ┌───┴───┐
                                          v       VP
                                     [*CASE:INT*] ┌─┴──┐
                                                  V  DP_int
                                                    [CASE:☐]
```
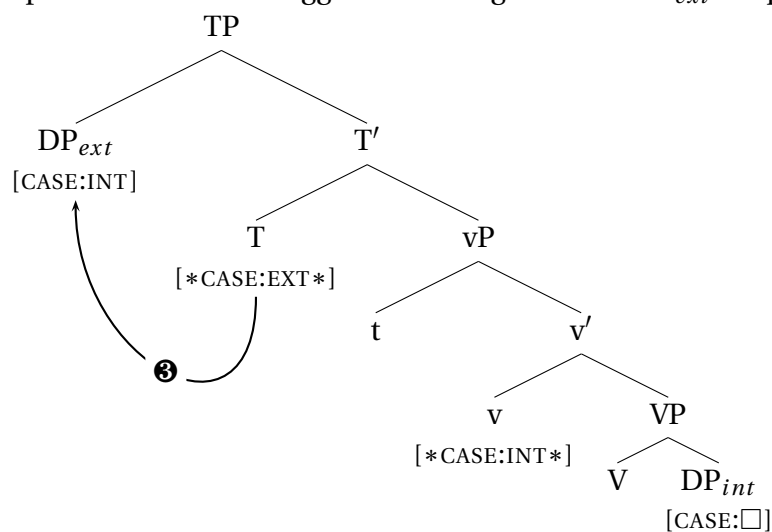
When DP_{int} is extracted, no problem arises because there is a derivation where both arguments receive a case value, cf. (6): DP_{int} first moves to its intermediate landing site

SpecvP; this happens before v triggers Agree, due to the ergative order Merge before Agree. Hence, both DP_{ext} and DP_{int} are in SpecvP. Since there is no minimality condition on Agree, v can assign the internal case to any of the two DPs. The derivation will converge if v assigns case to DP_{ext}, cf. (6-a). Afterwards, T merges with vP. T first triggers Merge of DP_{int} to SpecTP due to the order Merge before Agree; thereby T discharges its edge feature, cf. (6-b). T assigns case to DP_{int} in SpecTP in a subsequent step, as required by the Spec-head bias, cf. (6-c). DP_{int} has not been assigned a case value before. This derivation converges because both arguments receive a case value: DP_{ext} from v and DP_{int} from T.[2]

(6)   *Legitimate movement of DP_{abs}:*

   a.   Structure after T is merged
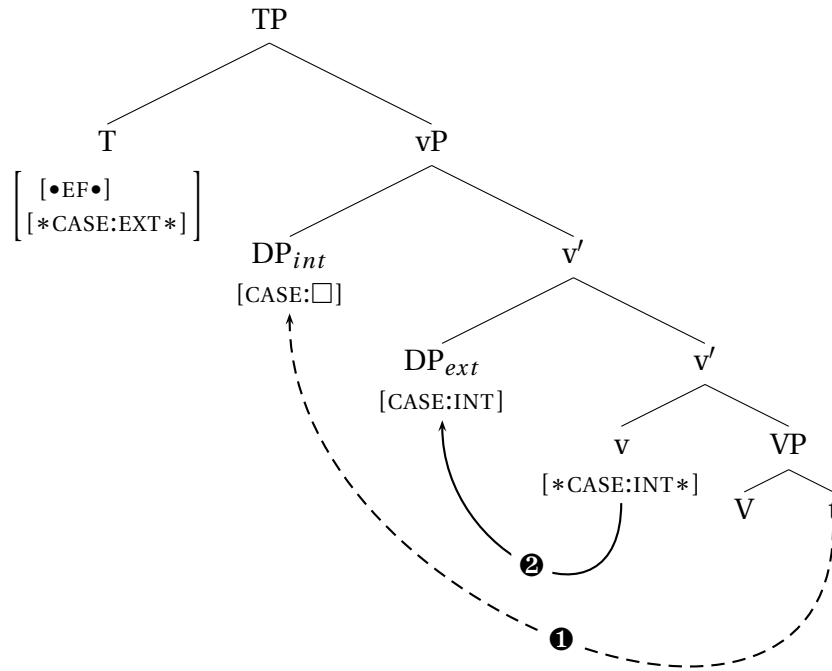
   b.   Merge before Agree on T triggers movement of DP_{int} first

---
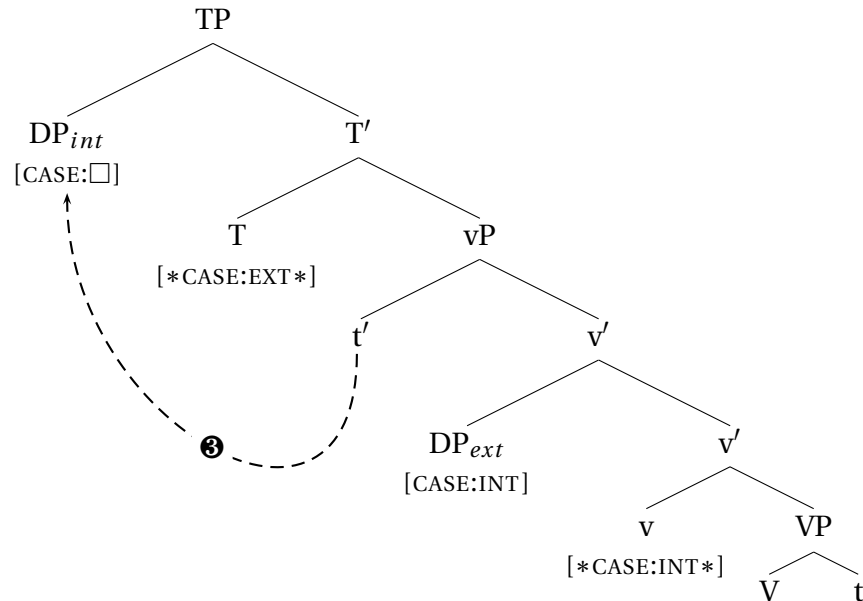
[2]If DP_{int} was assigned the external case by v after it moved to SpecvP (given the absence of a minimality constraint), the derivation crashes because DP_{int} will receive the case from T, as well, after it moved to SpecTP. This would leave DP_{ext} without a case value, parallel to the bleeding effect we get for DP_{int} with extraction of DP_{ext}. The crucial point is that there is a converging derivation when DP_{int} is extracted.

c. T assigns the ensures external to DP$_{int}$

```
                              TP
                ┌──────────────┴──────────────┐
            DP$_{int}$                        T′
         [CASE:EXT]              ┌────────────┴────────────┐
                                 T                        vP
                            [∗CASE:EXT∗]          ┌────────┴────────┐
                              ❹               t′              v′
                                                      ┌────────┴────────┐
                                                  DP$_{ext}$            v′
                                                 [CASE:INT]      ┌──────┴──────┐
                                                                 v            VP
                                                            [∗CASE:INT∗]    ┌──┴──┐
                                                                           V     t
```
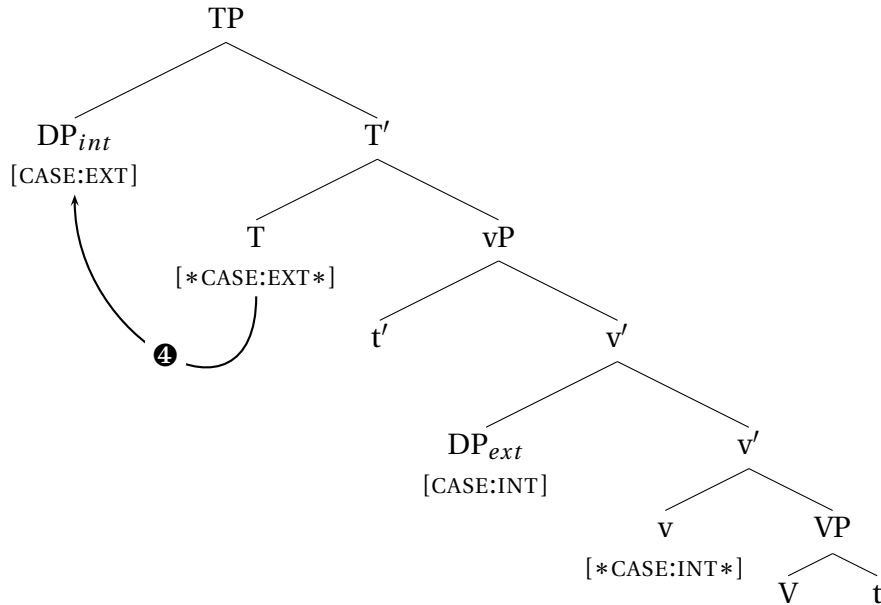
To summarize, Ā-movement of the ergative is impossible because it moves early and thereby bleeds case assignment to its co-argument. Early movement of the internal argument does not have this effect because its co-argument already has a case value when T attracts DP$_{int}$, so there is no bleeding. There is also no bleeding in intransitive contexts because there is no co-argument that could "steal" the case provided for the absolutive argument in the first place.

Let me now come to the crucial data that suggest a split between external and internal Merge. The approach by Assmann et al. (2012) makes an interesting prediction: Ā-movement of the ergative DP should be possible if the absolutive DP is Ā-moved as well. There are actually two converging derivations for this case. The first is illustrated in (7): Due to Merge before Agree, v first triggers Merge of DP$_{int}$ to SpecvP (vP is a phase) and external Merge of DP$_{ext}$ (order irrelevant). Afterwards, v assigns the internal case to DP$_{ext}$ (the Spec-head bias makes no predictions if both goals are in SpecvP and there is no minimality requirement). Then, both DPs undergo edge-feature-driven movement to SpecTP. Afterwards, T assigns external case to DP$_{int}$ in its specifier (again, this is possible because there is no minimality condition on Agree). Since both DPs receive a case value, the derivation converges.

(7)   *Legitimate movement of DP_{ext} and DP_{int}, morphological ergativity*
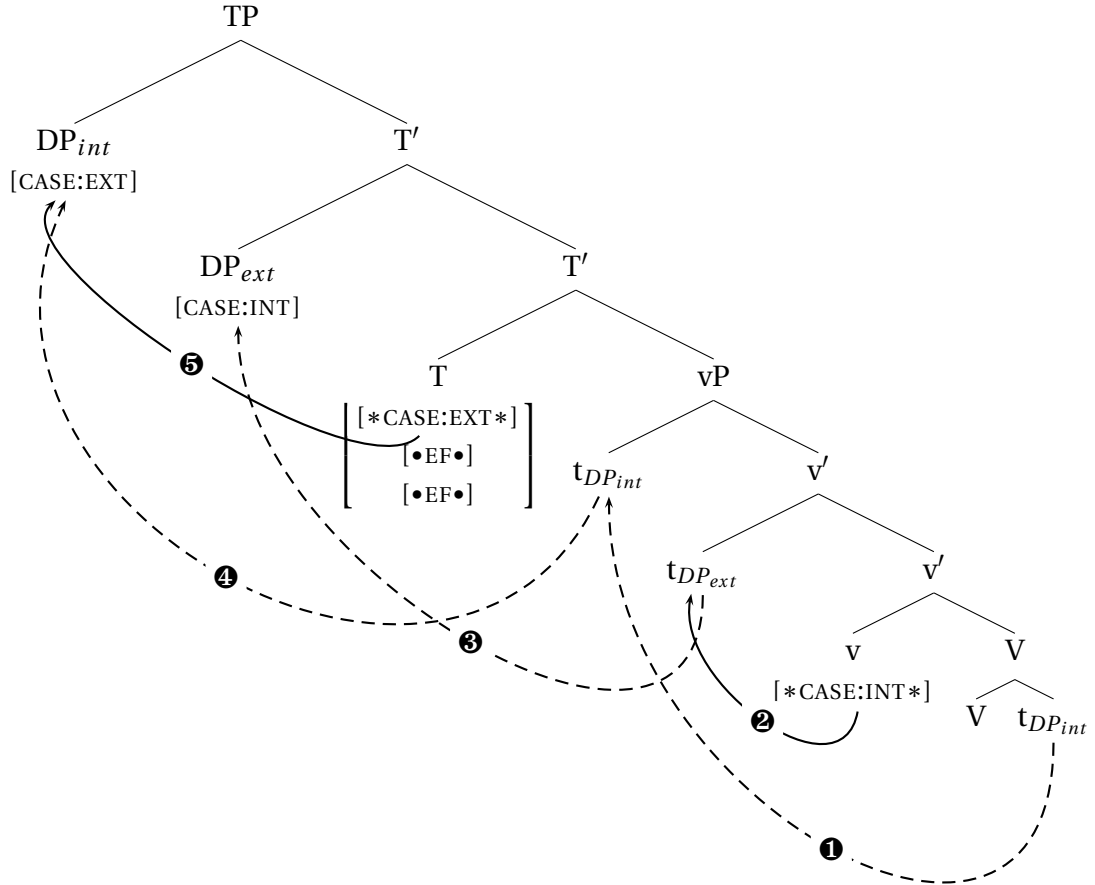
Extraction of DP_{ext} is indeed possible in (at least some) languages with the ban on ergative movement if DP_{int} is extracted as well (see Assmann et al. 2012: 36 for data).

The second converging derivation with extraction of DP_{ext} and DP_{int} is shown in (8): At the point of the derivation where both DP_{ext} and DP_{int} are in SpecvP, v can assign the internal case to DP_{int} because there is no minimality condition. Afterwards, both DPs are moved to SpecTP and T assigns the external case to DP_{ext}. Hence, both arguments get a case value and the derivation converges. However, note that this derivation leads to an accusative pattern of case assignment: DP_{int} ends up with the marked external case from v; DP_{ext} receives the unmarked external case from T, just as the sole argument of an intransitive verb. Hence, the account predicts that double extraction in a language with the order Merge before Agree, which produces morphological ergativity without extraction and if only a single argument is extracted, can lead to morphological accusativity under double extraction.

(8)   *Legitimate movement of DP_{ext} and DP_{int}, morphological accusativity*
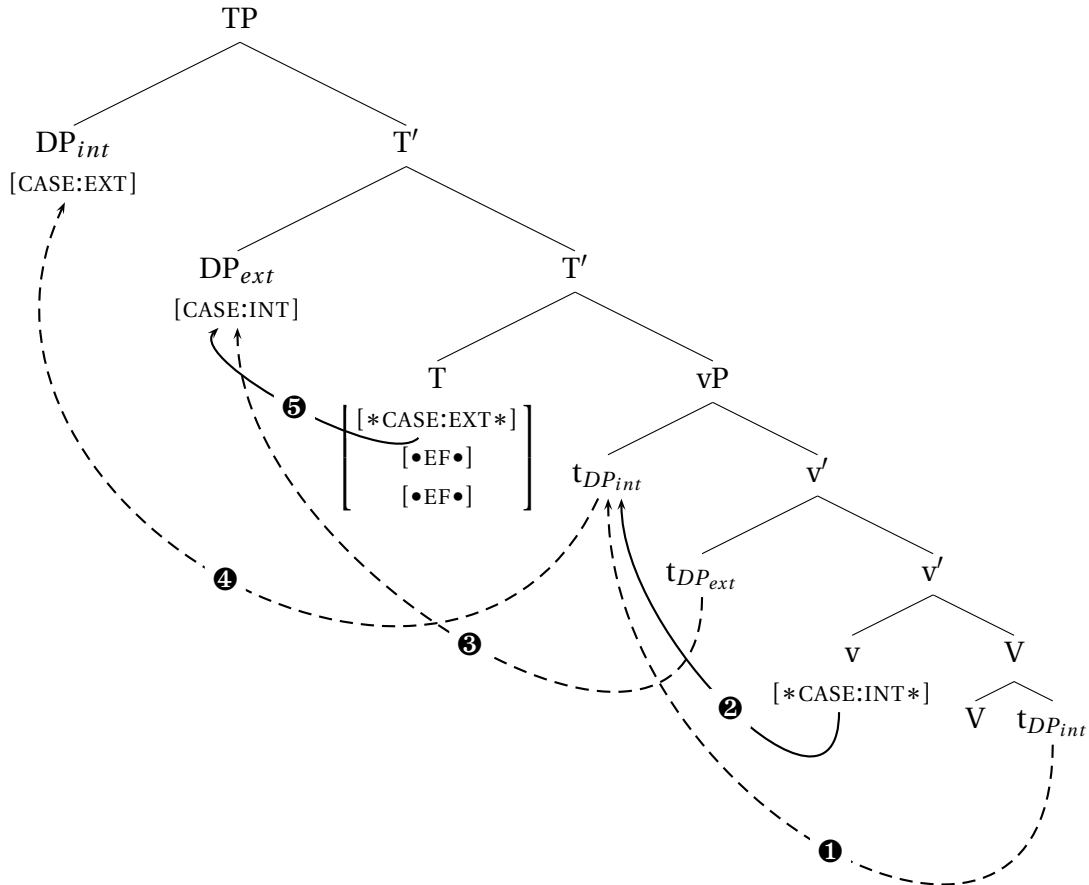
The only difference to the preceding derivation is the argument to which v and T assign case: If v assigns case to DP_{ext} (and T to DP_{int}), we get an ergative pattern; if, however, v assigns case to DP_{int} (and T to DP_{ext}), we get an accusative pattern. To the best of my knowledge, there is no language that changes its basic argument encoding pattern from ergative to accusative if both arguments of a transitive verb are extracted. With the present set of assumptions, the analysis in Assmann et al. (2012) thus overgenerates. They propose a solution to this problem that is based on the Merge over Move condition (Chomsky 2001) plus a reintroduction of a minimality condition on Agree for contexts in which a head projects more than one specifier (p.36): Agree is only possible with the innermost specifier of a head. Since external Merge applies before internal Merge (due to Merge before Move), DP_{ext} is in the innermost specifier of v when both arguments are extracted and will hence receive the internal case. This excludes the derivation in (8).

I want to propose a different solution that does not require minimality (which is desirable because the analysis does not needs this additional condition elsewhere). The crucial difference between the two derivations is which DP gets which case in the vP. We need to enforce that v assigns [∗CASE:INT∗] to DP_{ext} (as in (7)) in SpecvP but not to DP_{int} (as in (8)). This can be done when we make reference to types of Merge: DP_{ext} is externally merged in SpecvP whereas DP_{int} is internally merged in SpecvP. The required asymmetry follows if external Merge applies before case-Agree and internal Merge applies after Agree (the external Merge trigger is represented as [•D•]):

(9)   *Ordering statement in languages with the ban on ergative movement*:
       [•F•], [•D•] ≻ [∗CASE∗] ≻ [•EF•]

As a consequence, only DP_{ext} can be assigned case by v. At the point where v triggers Agree, only DP_{ext} is in SpecvP; DP_{int} is still in the complement of v. The Spec-head bias thus predicts case assignment to DP_{ext} in SpecvP. Movement of DP_{int} to SpecvP comes too late to bring the internal argument into the structural position to be a potential competitor for DP_{ext} with respect to case assignment from v. Thus, external Merge feeds upward Agree with v, whereas internal Merge counter-feeds Agree. The order in (9) correctly rules out the derivation in (8), which leads to morphological accusativity, because v assigns case to DP_{int} in SpecvP. The derivation in (7) is still possible, as desired.

The ordering explanation presupposes that external and internal Merge can apply at different points of the derivation, which in turn requires that they are of a different type. If Merge was a uniform operation, we could not order it both before and after Agree on a single head; but if subtypes are distinguished, Agree can apply after subtype 1 and before subtype 2. I take the approach that simply orders external vs. internal Merge on v to be superior to the minimality approach proposed by Assmann et al. (2012) for the following reasons: (i) Merge over Move is dubious because there is evidence that the opposite order of external and internal Merge is required as well (cf.section 4.1.1); the constraint cannot hold universally (see section 5.2 for a more general discussion of the role of such constraints); (ii) Assmann et al. (2012) need an additional assumption on top of those introduced at the beginning of this subsection, viz. minimality, although it is not necessary for any other derivation in their system. In the ordering approach, there is no minimality condition on top of the other assumptions. All effects follow from the order of elementary operations, which the analysis by Assmann et al. (2012) makes use of anyway; it is just not fine-grained enough to exclude the derivation in (8). If external and internal Merge can be ordered differently relative to Agree (and not just Merge in general), ordering is all that is needed to rule this derivation out. Hence, there is evidence for the order external Merge ≻ Agree ≻ internal Merge.

However, there is a problem if case assignment applies before intermediate movement steps, as demanded by the ordering statement in (9): If this order also holds on the TP-cycle, movement of the external argument of a transitive verb (cf. the derivation in (5)) should not bleed absolutive assignment to the internal argument: The intermediate movement step of DP_{ext} to SpecTP would come too late for DP_{ext} to be assigned the external case from T. DP_{ext} could be assigned this case because at the point where T probes, both arguments are still in the c-command domain of T, and since there is no minimality condition on Agree, DP_{ext} could be a potential goal. But there is also a converging derivation where the T-case is assigned to DP_{int}. Consequently, extraction of the ergative DP alone should not result in ungrammaticality under the order in (9). There are two solutions to this problem: First, the order in (9) may only hold for the v head (as a lexical specification); in the general ordering statement that holds for all other heads, including T, final and intermediate movement steps apply before case assignment. Consequently, intermediate movement steps to SpecTP apply early enough to bleed subsequent assignment of the external case to DP_{int}, which derives the ban on ergative movement. An alternative would be to say that the ordering statement in (9) holds for all heads, but Ā-movement terminates in SpecTP. If this was the case, movement of the external argument to SpecTP in (5) qualifies as a final movement step, which applies before case assignment even in (9). Hence, movement of DP_{ext} to SpecTP will still bleed case assignment to DP_{int}. Indeed, Aissen (1992) argues for Tzutujil (Mayan) that focus movement, one of the move-

ment types that triggers the ban on ergative movement, ends in SpecTP rather than in SpecCP. Evidence for this low position of focussed DPs come from the observation that material that would be associated with C or the TP precedes focussed DPs (but follows topics, which are assumed to be in SpecCP):

  (i) The negative particle occurs between topic and focus:
       top – neg – foc

 (ii) The interrogative particle (realizing C) occurs between topic and focus:
       top – Q – foc

(iii) Sentential adverbs (adjoined to TP) occur between topic and focus:
       top – adverb – foc

If movement to SpecTP is final, then the order of operation-inducing features in (9), which distinguishes between external Merge / final internal Merge and intermediate internal Merge, correctly predicts both the ban on ergative movement for extraction of a single argument, as well as the grammaticality of double extraction (with morphological ergativity only).[3] It remains to be seen, however, whether Ā-movement types that trigger the ban on ergative movement end in SpecTP in all languages with this extraction asymmetry. If this is not the case, I would have to adopt the first solution with a lexically specified order of features on v.

4.1.3   The external / internal Merge split and the Cycle

I have argued in section 4.1.1 that internal and external Merge can be interleaved with Agree such that internal Merge applies before and external Merge after Agree:

(10) $[\bullet\text{EF}\bullet] \succ [\ast\text{F}{:}\square\ast] \succ [\bullet\text{D}\bullet]$

As a consequence, external Merge to SpecvP comes too late to feed upward Agree with v. But in principle, there is another way to avoid that external Merge of DP_{ext} feeds Agree with v: The external argument is merged in a projection, viz. GP, *above* the vP, cf. (11):

(11) $[_{\text{GP}} DP_{ext} [_{\text{G}'} G [_{\text{vP}} DP_{int} [_{\text{v}'} v_{[\ast\text{F}{:}\square\ast]} [_{\text{VP}} V \underline{\quad}_{DP_{int}}]]]]]$

v bears the relevant probe feature; movement of DP_{int} to SpecvP can feed upward Agree with the probe on v. However, DP_{ext} cannot feed Agree with v because it is not in SpecGP at any point of the derivation (recall that upward Agree reduces to Spec-head agreement under the current set of assumptions, see section 2.5.3.1 for discussion). Hence, the subject-object asymmetry follows from the Cycle: If every XP is a cyclic domain, internal Merge applies to a domain (vP) which is included in the domain where external Merge applies (GP). The attempt of v to Agree with an element in its specifier applies at a point where DP_{ext} is not even part of the structure. Indeed, it has been argued in the literature on argument structure that the head that introduces the external argument must be distinct from the head v (see among others Pylkkänen 2002; Alexiadou et al. 2006; Merchant 2008; Harley 2009; 2013). The head that introduces DP_{ext}, often called Voice, takes vP as

---

[3]The order would also predict that topicalization of DP_{ext} in Tzutujil is grammatical: Since it targets SpecCP, it makes an intermediate movement step in SpecTP and should thus apply too late to bleed case assignment to its co-argument. Indeed, this is the case in Tzutujil, as Aissen (1992) shows.

its complement.[4] If DP_{ext} is introduced in SpecVoiceP, it cannot feed Agree with v. There-fore, one might object that the different behaviour of external and internal Merge with respect to Agree is not a convincing argument for the need of a language-specific order-ing of these operations on a single head; it may simply be the result of the Cycle.

The challenge for this alternative analysis is, however, that there is also evidence for the reverse order of operations in (12) (cf. section 4.1.2):

(12)　　*Opaque interaction of external / internal Merge with upward Agree*:
　　　　$[\bullet D\bullet] \succ [*F:\square*] \succ [\bullet EF\bullet]$

Given this order, external Merge feeds upwards Agree but internal Merge comes too late to do so. If this was to follow from the Cycle as well, we would have to assume that the projection that introduces DP_{ext} (GP) is dominated by the projection that triggers internal Merge of DP_{int} (vP),and that the head of the lower projection G bears the probe feature, cf. (13).

(13)　　$[_{vP} DP_{int} [_{v'} v [_{GP} DP_{ext} [_{G'} G_{[*F:\square*]} [_{VP} V \underline{\quad}_{DP_{int}}]]]]]$

There are several problems with this analysis. First of all, if we want to derive both orders in (11) and (12) from the Cycle, this would amount to say that languages can differ in the order of projections in the functional sequence, which I take to be undesirable. But even if we would allow for such variation, locating GP below vP as in (13) does not help if GP is a phase (as it would be in a very local system like the present one where every XP is a phase): DP_{int} must make a stop-over in the specifier of every phrase along the path of movement; hence, it would also have to go through SpecGP, the projection where DP_{ext} is externally merged. As a consequence, we end up with the same indeterminacy that we started with: External and internal Merge apply within the same phrase (here: GP) and thus in the same cycle; therefore, they need to be ordered in this cycle. The only way out would be to assume that in languages with the order in (12) DP_{int} is base-merged above DP_{ext}, in contrast to languages with the order in (11). Given the pervasive evidence for subject-object-asymmetries cross-linguistically, I take this solution to be highly implau-sible (but see Marantz 1984 for a proposal along these lines). To conclude, the fact that we find the opposite orders of external and internal Merge relative to Agree provides evidence that the two types of Merge operations need to be distinguished (here, by the features that trigger the respective operations).

4.2　Movement-type specific final and intermediate Merge

So far, I have argued on the basis of the configuration in (14) that we need to distinguish between external and internal Merge as well as between final and intermediate internal Merge because they apply at different points in the derivation.

---

[4]A terminological note: Kratzer (1996) argues that the external argument is introduced not in the VP but in the Spec of a head that takes VP as its complement. She calls this head Voice. It is also responsible for accusative assignment to the internal argument of a transitive verb in nominative-accusative languages. In subsequent literature, this head has usually been labeled v instead of Voice, meaning one and the same head with the same functions. In the literature on argument structure, it is argued that there is another functional head in addition to Kratzer's Voice above the VP, called v. Crucially, this v is not to be confused with the use of the label v sometimes used for Kratzer's Voice head. The v from the argument structure literature is a head that categorizes (verbalizes) category-less root (projections), cf. Marantz (1997); in contrast to Kratzer's Voice head (often called v), it does *not* introduce DP_{ext}.

(14)   *Order of operations triggered by a single head H*:
         Merge > Agree > Merge

The question arises whether there is evidence from interactions of the type in (14) for even more subtypes of Merge. In this section I summarize and extend an argument from the literature for the need of different types of final and intermediate internal Merge triggers.

4.2.1   Evidence

den Dikken (2009*b*: 92-93) provides an argument against the postulation of a generic edge feature as the trigger for intermediate movement steps. The argument is based on the following data from Belfast English as described in Henry (1995):

(15)   *Inversion in Belfast English (Henry 1995: 108-109, 120)*:

    a.   [$_{CP}$ Who$_k$ did John say [$_{CP}$ did Mary claim [$_{CP}$ had John feared [$_{CP}$ would Bill attack ___$_k$]]]] ?

    b.   *[$_{CP}$ Who$_k$ do you think [$_{CP}$ did John convince ___$_k$ [$_{CP}$ did Mary go]]] ?

    c.   [$_{CP}$ Who$_k$ do you think [$_{CP}$ did John convince ___$_k$ [$_{CP}$ that Mary went]]] ?

    d.   *[$_{CP}$ Who$_k$ did John hope [$_{CP}$ that could he help ___$_k$]] ?

    e.   [$_{CP}$ Who$_k$ did John hope [$_{CP}$ that he could help ___$_k$]] ?

    f.   [$_{CP}$ This is the [$_{NP}$ man$_i$ [$_{CP}$ [who$_i$]$_k$ John claimed [$_{CP}$ that I saw ___$_k$]]]].

    g.   *[$_{CP}$ This is the [$_{NP}$ man$_i$ [$_{CP}$ [who$_i$]$_k$ John claimed [$_{CP}$ did I see ___$_k$]]]].

Belfast English has a syntactic reflex of successive-cyclic movement: Subject-auxiliary-inversion (optionally) applies in all clauses along the path of wh-movement (pattern I), cf. (15-a). That this is indeed a reflex of movement is clear because inversion can only apply in clauses crossed by movement but not below the base position of the wh-phrase, cf. (15-b) vs (15-c). Inversion involves movement of the finite verb from T to C: It is impossible if the complementizer is overt in the embedded clause, cf. (15-d) vs. (15-e). Crucially, inversion cannot apply under relativization, cf. (15-f) vs. (15-g).

      On the basis of these data, den Dikken argues that inversion is not a *general* property of successive-cyclic movement.[5] Assuming that wh-movement and movement of the relative operator are subject to the same locality restrictions (making stop-overs in the same positions), we would expect them to pattern alike in triggering a reflex of movement in intermediate positions, but they do not in Belfast English. Abels (2012: 53-59), who elaborates on den Dikken's argument, concludes that inversion must be sensitive to the type of movement. Hence, we need *movement-type specific features* on intermediate C heads. (Abels 2012: 53) comments that "[...] it is difficult to make the embedded complementizer sensitive to the content of the moving phrase under accounts that assume a generic (EPP) feature on the embedded complementizer to trigger successive-cyclic movement (or no feature at all)." (where the EPP feature corresponds to what I have been calling edge feature). Let me illustrate the argument on the basis of the present ordering approach. Descriptively, wh-movement feeds inversion in both final and intermediate positions (pattern I), whereas relative operator movement counter-feeds it in both types of positions (pattern IV). Assume that inversion is a reflex of upward Agree between the C head and an

---

    [5]In fact, den Dikken (2009*b*) concludes that inversion is not a reflex of successive-cyclic movement at all, but of terminal movement steps. He denies the existence of successive-cyclic movement through intermediate SpecCPs in general; movement to this position is always terminal. 'Long movement' actually involves clause-bounded movements to the local SpecCP plus scope-marking and concord between the scope marker and the wh-phrase (full-concordial scope marking). See the discussion in section 2.4.1.2.

XP moved to SpecCP (see section 2.6.4.1 for a detailed analysis of inversion in Spanish). If there was only a single type of edge feature, we would need the two orders in (16) to derive patterns I and IV:

(16) *An ordering paradox in Belfast English*:
   a. [•WH•], [•EF•] $\succ$ [∗F:□∗]   *pattern I*
   b. [∗F:□∗] $\succ$ [•REL•], [•EF•]   *pattern IV*

When comparing (16-a) and (16-b), we run into a paradox: The generic edge feature would have to be ordered both before and after the Agree relation that ultimately results in inversion. Hence, there would not be a single ordering statement that determines the required orders for both wh-movement and relativization. I thus follow Abels (2012) in concluding that we need movement-type specific triggers for intermediate movement steps (see Branigan 2011: ch. 3.2, 3.3 for the same conclusion based on inversion in English, whose availability also depends on the movement type). In what follows, I indicate the movement-type as an index to the edge feature, but this is only a notational convention. As a consequence of postulating different types of edge features, we only need a single ordering statement for Belfast English that covers both the wh-movement and the relativization orders in (16), cf. (17-a):

(17) *Ordering statement in Belfast English*:
   a. [•WH•], [•EF$_{wh}$•] $\succ$ [∗F:□∗] $\succ$ [•REL•], [•EF$_{rel}$•]
   b. C head that triggers final wh-movement:
       C $\begin{bmatrix} [•WH•] \\ [∗F:□∗] \end{bmatrix}$
   c. C head that triggers intermediate wh-movement:
       C $\begin{bmatrix} [•EF_{wh}•] \\ [∗F:□∗] \end{bmatrix}$
   d. C head that triggers final operator-movement:
       C $\begin{bmatrix} [∗F:□∗] \\ [•REL•] \end{bmatrix}$
   e. C head that triggers intermediate operator-movement:
       C $\begin{bmatrix} [∗F:□∗] \\ [•EF_{rel}•] \end{bmatrix}$

(17-b-e) illustrate the suborders that are consequently instantiated on heads that trigger final and intermediate wh- and operator movement, respectively.[6]

   More evidence for movement-type specific edge features comes from defective intervention in Italian: When discussing this effect in section 3.3, I sowed that Romance languages exhibit pattern I: EPP-driven movement of a full DP to SpecTP is blocked by an experiencer that intervenes between T and this DP. If the experiencer moves to the T domain (cliticization) or undergoes Ā-movement to SpecCP, this movement feeds EPP-movement of the lower DP to SpecTP. In Italian, topicalization is an instance of Ā-movement that has this effect, cf. (18) (we have seen the same effect for wh-movement of the experiencer in French, cf. section 3.3.4):

---

[6]The same abstract pattern as in Belfast English holds in Spanish (Torrego 1984): Overt wh-movement triggers inversion along the path of movement, but inversion cannot apply if a relative operator is moved. Neither Henry (1995), nor Torrego (1984), nor Pesetsky and Torrego (2001) (who develop an analysis of inversion under operator movement) have an account for the split between the wh-movement and relativization, as the latter two explicitly state. Postulating movement-type specific edge features is a way to derive the effect.

(18)   *Defective intervention in Italian (Rizzi 1986b: 75)*:
    a. ?*Gianni sembra a Piero non fare   il   suo devere.
        Gianni seems   to Piero neg to.do def his duty
        'Gianni doesn't seem to Piero to do his duties.'
    b.   Gianni non gli   sembra fare   il   suo devere.
        Gianni neg him seems   to.do def his duty
        'Gianni doesn't seem to him to do his duties.'   *cliticization*
    c.   A Piero, Gianni non sembra fare   il   suo devere.
        to Piero Gianni neg seem   to.do def his duty
        'To Piero, Gianni does't seem to to his duties.'   *topicalization*

Interestingly, Italian is special among the Romance languages in that not all types of Ā-movement have a feeding effect: As pointed out by Holmberg and Hróarsdóttir (2003: 1015) (attributing the data to Paolo Aquaviva), relativization of the experiencer does *not* feed EPP-movement of the lower DP, the experiencer still blocks movement:

(19)   *Relativization of a defective intervenor in Italian*:
    a. ?*l'uomo   a cui   Gianni sembra non fare   il   suo dovere
        DEF.man to whom Gianni seems   NEG to.do DEF his duty
        'the man to whom Gianni doesn't seem to do his duties'
    b. ?*L'unico   osservatore a cui   la   situazione sembra (non) essere
        DEF.only observer   to whom DEF situation   seems NEG   to.have
        migliorata e' …
        improved is
        'the only observer to whom the situation doesn't seem to have improved is …'

With respect to relativization, Italian exhibits opacity of the Icelandic B-type: A final movement step of the experiencer to the T-domain (cliticization) cancels the intervention effect, whereas an intermediate movement step of a relative operator to SpecTP does not. However, there is feeding even for intermediate movement steps if the moving XP undergoes topicalization. Hence, different types of edge feature must be distinguished in Italian as well: intermediate topicalization vs. intermediate relativization. The effects follow from the order in (20). The EPP is represented as [•D•]; I assume that cliticization is triggered by a structure-building feature [•CL•] on T, but nothing hinges on the exact nature of this feature, only its order is relevant.

(20)   *Order of features on T in Italian*:
    [•CL•], [•EF$_{top}$•] > [•D•] > [•EF$_{rel}$•]

Early movement of the experiencer to the T-domain (like cliticization to T or intermediate movement of a topic XP to SpecTP) feed EPP-driven movement of a lower DP because the experiencer then no longer intervenes. Late intermediate movement of an experiencer that undergoes relativization counter-feeds EPP-driven movement to SpecTP, because the experiencer still intervenes at the point where the EPP tries to attract a lower DP.

    So far for intermediate movement triggers; but what about final movement triggers? Since Ā-dependencies in Belfast English have the same effect on final clauses as on non-final clauses (inversion possible under wh-movement, impossible under under relativization), we also need to postulate *movement-type specific features* that trigger final movement steps. In current terms, we need a [•WH•]- and a [•REL•]-feature for Belfast English because only the former triggers a movement step that feeds inversion; this has

been tacitly assumed in the ordering statement above. Of course, there are probably more distinct features for other Ā-movement types.

Note that the distinctions between types of internal Merge triggers I postulate are not identical to those that Abels (2012) assumes: In fact Abels concludes that though there are movement-type specific features, the specific feature used for movement of type X is identical on heads that trigger intermediate movement steps and final movement steps. In a wh-chain, for example, there is a [•WH•]-feature (using the current notational system) on *all* C heads between the base and the final landing site of the wh-phrase, and a [•REL•]-feature on all C-heads; viz. there is no distinction between final and intermediate movement step triggers. However, I have extensively argued for the latter distinction in chapters 2 and 3. Combining den Dikken and Abels' insight with mine, we actually need to distinguish between movement-type specific edge features *and* movement-type specific final features, as done in the ordering statement in (20-a). Thus, I argue for an even more fine-grained system of internal Merge triggers.

That different types of final movement steps must be distinguished has independently been proposed by Doliana (2013). The argument is based on the same configuration that I am using to detect the nature of Merge operations: Merge operations can apply at different points relative to Agree when triggered by a single head. Doliana (2013) presents an argument based on variation in PCC effects in West Germanic languages for a distinction between the two final movement steps scrambling and movement to the Wackernagel position.[7] The PCC is a constraint on the combination of phonologically weak objects depending on their person values. In Germanic varieties with a more variable order of weak object pronouns, less PCC effects occur. Doliana suggests the following derivation of the PCC: PCC effects with ditransitive verbs arise if both weak objects of a ditransitive verb Agree with v; in the basic structure, the IO asymmetrically c-commands the DO. v probes downwards and Agrees with the two objects, cf. (21-a).

(21)   *Interaction of downward Agree and Wackernagel movement*:
    a. Step 1: Agree with IO and DO

$$[_{vP} \text{ v} \quad [_{VP} \text{ IO} \quad [_{V'} \text{ V} \quad \text{DO }]]]$$
Agree

    b. Step 2: Wackernagel movement of both weak objects

$$[_{vP} \text{ IO} \quad [_{v'} \text{ DO} \quad [_{v'} \text{ v} \quad [_{VP} \text{ t} \quad [_{V'} \text{ V t }]]]]]$$
Move
Move

The combination of person features thereby valued on v is subject to constraints; certain combinations are filtered out, leading to the PCC effect (the technical details of the analysis need not concern us here). Objects that have entered into Agree with v undergo movement to the Wackernagel position in the respective Germanic languages. If both objects undergo movement, the second object tucks-in under the first moved object, resulting in the pre-movement order IO > DO on the surface (parallel movement, cf. (21-b)). Some varieties are less strict and also allow for the surface order DO > IO (Swiss German, Dutch). Interestingly, there is no PCC effect with this order, unlike with the reverse order IO > DO. Doliana derives this effect of word order on the PCC as follows: To generate the DO > IO order, only the DO must undergo Wackernagel movement (if both were Wackernagel-

---

[7]See also Müller (1998: ch. 5.3) for the claim that scrambling and Wackernagel movement must be distinguished in German.

moved, they would end up in the pre-movement order IO > DO). Recall that Wackernagel movement of a DP depends on previous Agree with v; thus, in order to escape Wackernagel movement, the IO must not Agree with v. This follows if it undergoes a different type of movement, viz. scrambling to SpecvP, which applies before Agree, cf. (22-a).

(22)   *Interaction of scrambling, downward Agree and Wackernagel movement*:
   a.   Step 1: scrambling of IO
         $[_{vP}$ IO $[_{v'}$ v $[_{VP}$ t $[_{V'}$ V DO $]]]]$
                   └──── Move ────┘

   b.   Step 2: Agree with DO
         $[_{vP}$ IO$_k$ $[_{v'}$ v $[_{VP}$ t$_k$ $[_{V'}$ V DO $]]]]$
                          └──── Agree ────┘
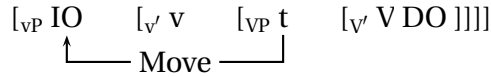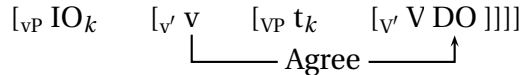
   c.   Wackernagel movement of DO
         $[_{vP}$ DO $[_{v'}$ IO $[_{v'}$ v $[_{VP}$ t$_{IO}$ $[_{V'}$ t $]]]]]$
                   └──────────── Move ────────────┘
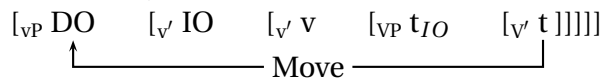
Since v probes downwards, early scrambling of the IO bleeds Agree between v and the IO because the IO is no longer in v's c-command domain at this point of the derivation, cf. (22-b). Since v can only Agree with the DO, only this object undergoes subsequent Wackernagel movement to an outer specifier of vP, resulting in the DO > IO order, cf. (22-c). Since the PCC can only arise if an offending combination of person feature of both the IO and the DO is present on v (due to Agree), there can be no PCC effect if v only Agrees with one of the weak object pronouns. In this way, Doliana derives the dependency of the PCC effect on word order. It is crucial for this analysis that scrambling and Wackernagel movement apply at different points relative to Agree:

(23)   Order of operations on v (Doliana 2013):
        scrambling > Agree > Wackernagel movement

Scrambling applies before Agree and can thus bleed Agree; Wackernagel movement applies after Agree and thus counter-bleeds Agree. We have thus another piece of evidence for the split between final movement steps.

4.2.2   Movement-type specific edge features and *Phase Balance*

If there are movement-type specific edge features, the question arises how we can ensure that the correct type of edge feature is inserted on a head: If the final feature is a [•WH•]-feature, for instance, the corresponding edge feature inserted on structurally lower heads must be [•X$_{wh}$•]. In fact, *Phase Balance,* which regulates edge feature insertion in the present approach, is well-suited for this purpose. The definition I proposed in section 2.5.2.3 is repeated in (24):

(24)   *Phase Balance*:
        An edge feature [•EF•] is inserted on the selected phase head H for every feature [•F•] on a head Y in the numeration if:
   a.   Y ≠ H,
   b.   there is no accessible matching feature [F].

Since *Phase Balance* makes reference to final features anyway when it checks whether an edge feature needs to be inserted, we can easily formulate a requirement that the type of final feature that triggers edge feature insertion and the edge feature inserted match in movement type. The extended version is given in (25); it is formulated as an instruction:

(25)   *Phase Balance*:
       For every feature [•α•] (where α is a variable over movement types) on a head Y in the numeration, insert an edge feature [•EF$_\alpha$•] on the selected phase head H if:
       a.   Y $\neq$ H,
       b.   there is no accessible matching feature [F].

In this way, the movement type of the edge feature is linked to the movement type of the final Merge trigger.

4.3   Pattern V: More types of internal Merge?

In chapter 2 I have presented four different patterns of reflexes of successive-cyclic movement (upward Agree): The reflex occurs in every clause crossed by movement (pattern I); the reflex occurs only in the clause in which the moved phrase surfaces (pattern II); the reflex occurs in all clauses except for the one where the moved phrase surfaces (pattern III); there is no reflex at all (pattern IV). However, there is another pattern, excluded from the discussion so far:

(26)   *Pattern V*:
       The reflex occurs only in the clause in which the moved phrase originates, but not in higher clauses crossed by movement.

A well-known instance of pattern V involving upward Agree is past participle agreement in French: The participle agrees in gender and number with a phrase that is base-generated in the VP (direct object of a transitive verb, sole argument of an unaccusative verb) and undergoes movement out of the VP. This is illustrated in (27) (inherent features of a DP are set in brackets). (27-a) provides the baseline sentence without extraction; participle agreement is excluded in this case. (27-b) illustrates cliticization, (27-c) relativization, (27-d) wh-movement, and (27-e) A-movement of the internal argument of a passivized verb:[8]

(27)   *Participle agreement in French, Kayne (1989: 85-86), Branigan (1992: 36)*:
       a.   Paul a   repeint   / *repeint-es   les chaises
            Paul has repainted / repainted-FEM.PL the chairs(FEM.PL)
            'Paul has repainted the chairs.'
       b.   Paul les$_k$ a   *repeint   / repeint-**es**   ___$_k$
            Paul them(FEM.PL) has repainted / repainted-FEM.PL
            'Paul has repainted them (the chairs).'
       c.   les chaises$_i$ [$_{CP}$ [que$_i$]$_k$ Paul a   *repeint   / repeint-**es**   ___$_k$]
            the chairs(FEM.PL)   that   Paul has repainted / repainted-FEM.PL
            'the chairs that Paul has repainted'

---

[8]See e.g. Kayne and Pollock (1978); Obenauer (1994); Déprez (1998); Branigan (1992) for more data and variation in the distribution of French participle agreement, as well D'Alessandro and Roberts (2008) and references cited there for participle agreement in Italian.

    d.   Je me demande [$_{CP}$ [combien de tables   ]$_k$ Paul a *repeint /
        I me ask   how.many of tables(FEM.PL) Paul has repainted /
        repeint-**es**   ___$_k$]
        repainted-FEM.PL
        'I wonder how many tables Paul has repainted.'

    e.   [Les livres de Jules Verne]$_k$ ont tous été *imprimeé / imprimeé-**s** ___$_k$
         the books by Jules Verne   have all been printed   / printed-PL
         'The books by Jules Verne have all benn printed.'

Since participle agreement does not occur with VP-internal arguments (in-situ), I conclude that it involves upward Agree between v and a DP internally merged to SpecvP. Furthermore, a DP that is externally merged in SpecvP never triggers participle agreement. In light of these facts, Kayne (1989) develops an analysis where participle agreement is the result of an upward agreement relation between a functional head in the extended projection of the verb (AgrOP) and the VP-internal argument moved to the specifier of that projection (the external argument is base-merged in a structurally higher projection).

    Crucially, under long Ā-movement, participle agreement is only triggered in the clause in which the moving XP originates (Branigan 1992). This is illustrated in (28-a) for long relativization and in (28-b) for long wh-movement ((28-b) is taken from Grohmann 2003: 287 who attributes it to Ian Roberts, p.c.):

(28)   *Participle agreement under long extraction*:

    a.   la lettre$_i$ [$_{CP}$ [qu$_i$]$_k$ 'il a   dit / *dit-e   [$_{CP}$ que Pierre lui a
        the letter(FEM)   that   he has said / said-FEM   that Pierre him has
        *envoyé / envoyé-**e** ___$_k$]]
        sent   / sent-FEM
        'the letter that he said that Pierre sent to him' (Chomsky 1995: 325)

    b.   [$_{CP}$ [Quelles   chaises   ]$_k$ as-tu   dit / *dit-es   [$_{CP}$ qu' il
         which.FEM.PL chairs(FEM.PL) have-you said / said-FEM.PL   that he
        a   *repeint / repeint-**es**   ___$_k$]]
        has repainted / repainted-FEM.PL
        'Which chairs did you say that he repainted?'

The central question is why there is no agreement on higher participles under long extraction. The standard answer in the literature is that there is simply no movement through (the equivalent of) SpecvP in clauses where the moving XP does not originate, cf. Branigan (1992); Grohmann (2003). Rather, cross-clausal movement is such that it can only proceed from SpecCP to SpecCP. The moving phrase can move through various projections in the clause where it originates; but once it has reached SpecCP of the minimal clause, it moves on directly to SpecCP of the higher clause(s) without a stop-over in other positions in these higher clauses. Branigan (1992: 42) makes this follow from binding theory: SpecAgrOP must not be filled by an element that binds a variable, i.e. it cannot be occupied by an Ā-moved phrase and cross-clausal movement is always Ā-movement; movement to SpecvP in the clause where the XP originates is A-movement and thus possible. In Grohmann (2003: 287f.) movement from SpecCP to SpecCP is enforced by the *Interclausal Movement Generalization* according to which cross-clausal movement is required not to target a lower domain on a hierarchy of domains than the domain of the launching site of the phrase in the lower clause (this is a version of the *Williams Cycle*, see section 5.2.1 for details). For the present purposes, it suffices to know that CP and vP constitute different

domains and that the former is higher on a hierarchy of domains than the latter. Since Ā-movement makes a stop-over in the embedded SpecCP, it must directly target the matrix SpecCP; a stop-over in the matrix SpecvP would be movement to a domain that is lower on the hierarchy of domains than the previously targeted CP. These analyses are an instance of hypotheses (HI) / (HII) introduced in section 2.4.1: If there is no reflex of movement on H, there has been no movement to SpecHP. I have argued in the aforementioned section that this correlation cannot in general explain the absence of a reflex in a certain position. In addition, I do not want to assume that long-distance movement has different locality restrictions across languages, viz. that it makes a stop-over in all SpecvP positions along the path of movement in languages with a pattern I reflex in the vP-domain (e.g. in Indonesian), but that it can skip higher SpecvP positions in languages like French. Chomsky (1995: 325f.) provides a different account of the French facts that does not deny an intermediate stop-over of the moving XP in SpecAgrOP of higher clauses. He claims, following a suggestion by Kayne (1989), that in cross-clausal movement dependencies, the moving phrase adjoins to the vP in higher clause (since it is Ā-movement), whereas is moved to the specifier of the vP in the lower clause (an instance of A-movement). By assumption, adjuncts to a projection of a head H are not in the checking domain of H, i.e. there is no agreement with adjuncts (Chomsky 1995: 119), but only with specifiers (and complements). This accounts for the pattern V distribution of participle agreement. That adjuncts are different in this respect does not follow from anything, it is a stipulation. Furthermore, it is questionable whether this account is translatable into more recent versions of Minimalism where the distinction between subjects and adjuncts is blurred (cf. Kayne 1994 as well as Chomsky 2013; in these systems there are no specifiers, they are more like adjuncts in creating XP-XP configurations when integrated into the structure). I conclude that there is no satisfying analysis of French participle agreement that assumes successive-cyclic movement through SpecvP of every clause and that can account for the absence of the morphological reflex in higher clauses.

Pattern V also does not automatically fall out from the analysis developed so far in this thesis because there is no type of Merge that exclusively refers to the first movement to the specifier of a certain projection (vP in French). Assume there is such a designated trigger; I will call it [•P•]. If this trigger is ordered before Agree on the head of the projection under question, abstractly called HP in what follows, the first movement step to SpecHP would feed upward Agree with H. If all other movement operations on H, call them [•P′•], are triggered after Agree, they would counter-feed Agree on all structurally higher H-heads. Thus, the order in (29) would give the desired result.

(29) $[•P•] > [*F:\square*] > [•P′•]$

However, there are serious doubts about this analysis. First of all, all subtypes of Merge I have argued for in this thesis have been proposed independently, on the basis of phenomena that do not involve the order of syntactic operations. Final and intermediate movement steps need to be distinguished when it comes to linearization, for instance; in a language like English, only the occurrence in the final position is pronounced. In addition, it is standardly assumed (in an Attract-based theory of movement) that it is the feature on the head that projects the final landing site that triggers movement in the first place; intermediate movement steps only apply for locality reasons. External and internal Merge can be distinguished e.g. in terms of θ-marking: External Merge can involve θ-marking of the merged XP, whereas internal Merge cannot (if the θ-Criterion holds, but see Hornstein 2001; Boeckx et al. 2010 for a different view). Different types of final

Merge have been postulated because Ā-dependencies (wh-movement, relativization, focus movement, etc.) have different properties. But I am not aware of any other area where the first movement step to a head of a certain category is relevant (in French, this category is v). Hence, there is no independent evidence that would support the postulation of a designated trigger for such a movement operation. Note that it is not sufficient to say that it is the very first movement step in a movement chain that is special in French: If every phrase is a phase, the XP base-merged as a complement of V first moves to SpecVP, and only afterwards does it move to SpecvP, the phrase where we find the reflex of movement. What is crucial is that it is the first step to the specifier *of a certain category*. In my opinion, it is highly implausible that a feature [•P•] that triggers such an internal Merge operation exists.

Therefore, I conclude that pattern V is different from patterns I to IV in that it does not result from the order of movement and Agree operations. But how does pattern V arise? A solution might be gained from the following observation: A similar effect is found with resumption, although the morphological reflex is not marked on a head, as in French, but in the base position of the moved XP. It has been repeatedly observed that resumptive pronouns only occur in the base position of their associated XP, but not in any other clause between the resumptive pronoun and its antecedent (see e.g. Salzmann 2011: 198). If resumption involves movement of the XP (Pesetsky 1998; Aoun et al. 2001; Boeckx 2003), the resumptive pronoun is a reflex of movement that has a pattern V distribution. A prominent analysis of resumption is the stranding / peeling analysis, cf. Aoun et al. (2001); Boeckx (2003); Klein (2013): XP is base-generated with a functional shell above XP, call this shell GP, cf. (30-a). Resumption results if XP is subextracted from GP, viz. movement of XP strands the functional shell GP, cf. (30-b). The stranded structure is realized as the resumptive pronoun.

(30)   *Stranding analysis*:

    a. $[_{GP} \text{ G XP }]$
    b. $[_{HP} \text{XP}_k \ [_{H'} \text{ H } ... \ [_{GP} \text{ G } \underline{\quad}_k \ ]]]$
    c. $[_v \ [_V \ [_V \text{ G V }] \text{ v }]]$

This derives pattern V because of (i) and (ii): (i) XP can strand the functional projection above it only once, of course. (ii) subextraction of XP from GP can only happen in the base position of XP. If the whole GP is moved to an intermediate landing site first, subsequent subextraction of XP is blocked because moved phrases are islands (*Freezing effect*, cf. Wexler and Culicover 1980; Browning 1991; Collins 1994; Müller 2010). What I would like to propose is that pattern V, as found with French participle agreement, arises in the same way, i.e. the movement-related morphology we see is the spell-out of stranded functional material left by subextraction of an XP. The only difference is that the residue of subextraction is marked in the position of the moved XP under classic resumption (dependent marking), whereas in French it is realized on the head in the vicinity of the stranded material (head-marking). I suggest that head-marking arises through head movement of the stranded functional head (G in the abstract example in (30)) to the next higher functional head in the extended projection of the verb. In French, this would involve the following derivation: GP, including DP, is base-merged as the sister of V. DP subextracts and strands G. Afterwards, G moves to V and subsequently to v, cf. (30-c) (under left adjunction). If only heads of category v have a feature that attracts G in French, this would also explain why we do not get participle agreement with external arguments: This would involve

downward movement of G to v, which is counter-cyclic.[9] In fact, Rocquet (2013: 209f.) develops a peeling analysis of French participle agreement along these lines. More generally, she argues that the French pattern and the definite / indefinite conjugation in Hungarian are head-marking variants of the well-known dependent-marking cases of differential object marking (see e.g. Silverstein 1976; Comrie 1979; Bossong 1985; Aissen 2003). This is not implausible in light of the fact that it has been claimed that participle agreement in French goes hand in hand with a specificity effect, as it is also found with differential object marking. See also Boeckx (2003: 59ff.) for the tight relation between French participle agreement and resumption: both are claimed to have the same semantic effects and to improve island violations.

I will not develop the analysis of pattern V as stranding + head-movement in more detail here. The crucial point is that there is an alternative analysis to the unattractive postulation of a designated trigger [•P•]: stranding; this analysis has been independently proposed for similar effects. In addition, it is not clear how wide-spread pattern V is. First, we would have to investigate whether it is actually a *pattern* that needs to be explained, viz. whether there are more languages with the same abstract distribution of a reflex of movement. I am not aware of any other clear cases. Note that the distribution of Anti-agreement in Fiorentino type languages might seem to follow pattern V as well (cf. section 3.2.1.3): The bleeding effect that Ā-movement of the subject has on subject-verb-agreement only ever occurs in the clause in which the subject is base-merged, but not in higher clauses. However, this distribution follows automatically from the direction of Agree: Since the ϕ-probe on C looks downwards in AAE-languages, movement of the subject of the clause CP_1 that contains this C-head can bleed Agree with C_1 (if movement applies too early); but when the subject of a lower clause Ā-moves to the specifier of a structurally higher clause CP_2, the probe on the C_2-head will always find the subject of the respective clause in its c-command domain, and therefore, there can be no bleeding in higher clauses under long subject extraction. The AAE in Fiorentino-type languages is a pattern I reflex. With upward Agree, things are different because it leads to feeding instead of bleeding: It is not plausible that in French, feeding of Agree between a structurally higher v-head and a DP moving to the specifier of this head is blocked by material in the higher clause; if anything, we would expect that another element from the higher clause feeds Agree with the v-head of this clause, but not that there is no Agree at all in higher clauses.

To summarize, I have argued that what looks like another pattern of a reflex of movement is different from patterns I to IV: Pattern V does not result form an Agree relation that is fed by movement, hence, we do not have to postulate another (rather quirky) subtype of Merge. Instead, the movement-related morphology is the realization of stranded functional material base-generated above the moved XP.[10]

---

[9]Head movement must be blocked when neither GP nor XP move; it seems to be the case that G must somehow be activated to be able to undergo head-movement.

[10]Lahne (2008*b*: 118) also concludes from the peculiar behaviour of French participle agreement that it is not a genuine reflex of successive-cyclic movement like pattern I reflexes. She hints at a different analysis, though: Participle agreement is subject to context-sensitive spell-out: It is the spell-out of agreement between the internal argument of V and the participle; vocabulary insertion into the feature set that the past participle has acquired through Agree is, however, blocked if the direct object is "in a local relation with the participle" (p.119), i.e. when the internal argument is in its base position. Why it is not blocked in higher clauses under long extraction is not clear to me. This would require a more detailed definition of *local relation*.

4.4   Interim conclusions

In this chapter I have argued that we need to distinguish more types of Merge than just final and intermediate movement steps, viz. external vs. internal Merge as well as different types of final and intermediate internal Merge. The evidence for this split of Merge types is the same as in chapter 2: When triggered by the same head H, these Merge operations apply at different points of the derivation relative to Agree; one type feeds Agree, the other type counter-feeds Agree. If Agree was ordered solely to Merge (as a uniform operation), we would expect either feeding or counter-feeding throughout.

Chapter 5

General implications

In this chapter I will investigate the more general implications of the ordering analysis developed in this thesis with respect to the dichotomies introduced in chapter 1: extrinsic vs. intrinsic ordering, sequential vs. simultaneous rule application, derivational vs. representational approaches to interactions, syntactic vs. postsyntactic operations. In section 2.4.4 I argue that Agree is a syntactic and not a postsyntactic operation. In section 5.2 I demonstrate that the variation in the order of movement steps provides evidence for the need of extrinsic ordering, as well as for sequential application of operations within a cyclic domain. In section 5.3 I show that the flexible timing of intermediate movement steps across languages requires that edge feature insertion applies early, and that the timing of edge feature discharge is variable across languages, contra previous claims in the literature. Section 5.4 illustrates that a representational reanalysis of transparent and opaque interactions of Merge and Agree would be of a rather complex type; the derivational ordering approach is much simpler and elegant. Finally, section 5.5 shows that the present Attract-based ordering analysis of interactions, in which the subtypes of operations are distinguished by their triggers, can also be implemented in systems with rather different basic assumptions as long as there is some kind of distinction between the subtypes. This is demonstrated for a Greed-based system and a system where some movement steps are not feature-driven.

5.1   Agree is a syntactic operation

The ordering analysis of patterns of reflexes of movement as outlined and argued for in chapters 2 to 4 provides an argument for Agree as a syntactic operation and against claims that Agree is a postsyntactic operation (see e.g. Bobaljik 2008 on postsyntactic ϕ-agreement and Marantz 1991; McFadden 2004 on postsyntactic case assignment). The argument goes as follows: Postsyntactic operations can be fed or bled by previous, i.e. syntactic operations such as movement because the input to the postsyntactic component is the output of syntax. A postsyntactic operation cannot, however, be counter-fed or counter-bled by a syntactic operation because, by definition, opaque interactions are not surface-true (not retrievable from the output, unless enriched representations are used, but see section 5.4). In the postsyntactic component, an XP either is in the right structural position to be a goal for a probe on a head H or it is not. Syntactic movement of XP cannot come too late to feed or bleed a postsyntactic operation such as Agree since the syntactic operation will necessarily have to apply before the postsyntactic one. Since I have argued that movement can counter-feed or counter-bleed Agree, Agree must have applied before

Agree.

5.2   Extrinsic vs. intrinsic ordering

As pointed out in section 1.2.1.2, a central research question in frameworks that work with sequential rule application has been whether the order of rules must be stipulated in a language-specific fashion (extrinsic ordering) or whether it follows from independent principles of the grammar. There are two major views on the need for extrinsic ordering in the literature. Representative for the first view, Chomsky (1965: 223, fn.6) states that "[t]axonomic linguistics disallows extrinsic ordering, but has not been clear about the status of intrinsic ordering. Generative grammars have ordinarily required both." According to the second view, most explicitly elaborated in Pullum (1979) (but see also Koutsoudas 1973; Keyser and Postal 1976), there is no need for extrinsic ordering; all orderings are determined by universal principles. This is formulated in the *UDRA*-hypothesis:

(1)   *Universally Determined Rule Application (UDRA, Pullum 1979: 18)*:
        All restrictions on the application of rules are determined by universal principles.

From a conceptual point of view, intrinsic ordering is to be preferred because (i) orders do not have to be stored, and (ii) the number of possible grammars is restricted, which in turn facilitates language acquisition. The second point has been emphasized by Pullum (1979: 28f.) (see also McCawley 1988: ch.6): If all orders are intrinsic, there is exactly one possible grammar for a number n of rules; if all orderings are extrinsic, however, the number of possible grammars corresponds to n! (every logically possible order leads to a possible grammar). Since the number of possible grammars under extrinsic ordering is extremely high even with a relatively small number of rules, Pullum concludes that the Strict Ordering Hypothesis with exclusively extrinsic ordering is to be avoided; it would leave the child with the task to choose the target grammar from an enormous number of possible grammars.

In this section, I will show that the data presented in this thesis provide evidence for the first view, viz. that extrinsic ordering is required after all, in addition to intrinsic ordering.

5.2.1   On intrinsic ordering by universal principles

Before we can evaluate whether the orders I have proposed in this thesis require extrinsic or intrinsic ordering, I will summarize what principles have been proposed to predict orderings. We have seen in section 1.2.1 that the order of rules may follow from the way the rules are formulated, and – depending on one's model of the architecture of grammar – it can follow from the level / stratum at which rules apply: If a rule R_1 applies to a level L_1 and another rule R_2 applies to a level L_2 that is ordered after L_1, R_1 precedes R_2; consequently, R_1 can feed or bleed R_2, but not the other way around. As for the the Y-model of grammar in GB and Minimalism (cf. figure 1.7), this means that, for instance, postsyntactic PF rules cannot feed or bleed syntactic rules. In addition, it has been proposed that there are universal principles that predict orderings. The following list gives an overview of the candidates for such principles, suggested in the literature on rule interaction in phonology, morphology and syntax:

(i)   SPECIFICITY: The more specific rule applies before the less specific rule.

(ii) COMPLEXITY: The less complex rule applies before the more complex rule.

(iii) OBLIGATORINESS: Obligatory rules apply before optional rules.

(iv) HIERARCHY-GOVERNED ORDERING: Domains of syntactic structure are ordered on a hierarchy such that a rule in a domain D_1 precedes a rule in a domain D_2 if D_1 is higher on the hierarchy than D_2.

(v) CYCLICITY: Rules affecting a cyclic domain α apply before rules in a cyclic domain β if α is contained in β.

Specificity is the basis of the *Elsewhere Principle*, already introduced in section 1.2.1.3: If two rules compete for application, the more specific rule blocks the application of the less specific rule. Thus, the Elsewhere Principle is usually interpreted as disjunctive: It is not the case that both rules can apply and the principle predicts their order; rather, only a single rule is chosen for application. An example for its application in syntax is the interaction of Extraposition and It-deletion, introduced in section 1.2.1.3 (cf. example (37)). The importance of specificity for syntactic derivations has recently been highlighted by van Koppen (2005) and Lahne (2012).

The second concept that has been argued to determine rule ordering is complexity. An instance of this concept in syntax is the constraint *Merge before Move* (Chomsky 1995; 2000). If a structure-building feature can be discharged either by external Merge or by internal Merge (movement), the former is chosen because internal Merge is more complex then Merge: Under the copy theory of movement and the assumption that Agree is a precondition for movement, internal Merge involves Merge plus Agree plus the creation of a copy. *Merge before Move* becomes relevant e.g. when the subject position in English (SpecTP) is to be filled: T bears an EPP-feature which can in principle be discharged by external Merge of an expletive or by movement of a DP already contained in the structure. This constraint is also disjunctive because the structure-building feature can be discharged only once.

That obligatoriness determines rule ordering has been proposed by Ringen (1972) in the *Obligatory Precedence Principle* according to which obligatory rules apply before optional rules. As exemplified in section 1.2.1.3, this principle determines the counterbleeding order Reflexivization before Imperative Subject Deletion in English.[1]

Hierarchy-governed ordering is introduced in Williams (1974): Williams proposes the Inclusion hierarchy in (2-a) which refers to nodes in the syntactic structure, cf. (2-b):

(2) a. Inclusion Hierarchy: S̄ ⊃ S ⊃ Pred ⊃ VP
     b. [$_\text{S̄}$ Comp [$_\text{S}$ NP [$_\text{Pred}$ Pred [$_\text{VP}$ V NP]]]]

A rule with the domain D_1 (the minimal XP in the Inclusion Hierarchy that contains all elements affected by the rule) applies before rules in a domain D_2 which is to the left of D_1 on the hierarchy; this is known as the Williams Cycle (see also Williams 2003 for further developments of this idea and Müller 2013*a* for a local reformulation). This approach predicts, for instance, the counter-bleeding order of Topicalization and Reflexivization in (3), discussed in section 1.2.1.1:

---

[1]As Pullum (1979: 51) notes, the order Extraposition before It-deletion, determined by the Elsewhere Principle, is the reverse of what the Obligatory Precedence Principle predicts: Extraposition is optional whereas It-deletion is obligatory; according to obligatoriness, they should apply in the order It-deletion before Extraposition, contrary to what is required to derive (37-b). Pullum solves this puzzle by assuming that the universal principles can be ranked in case of conflict (reminiscent of constraint-ranking in optimality theory): The Elsewhere Principle is more important than the Obligatory Precedence Principle when they make conflicting predictions.

(3)   Myself, I never buy anything for.

Topicalization of the anaphor should bleed Reflexivization because on the surface, the reflexive pronoun is no longer in the c-command domain of its antecedent. Reflexivization applies in the S-domain, whereas Topicalization applies in the S̄–domain. Hence, Reflexivization applies before Topicalization. Related hierarchies are proposed by Grohmann (2003) and Abels (2008).[2]

The Cyclic Principle has been introduced and discussed in section 1.2.1.2. The definition is repeated in (4):

(4)   *Cyclic Principle (definition based on McCawley 1988: 28)*:
         When one domain to which rules can apply is contained in another, the applications of a rule to the smaller domain precede the applications of this rule to the larger domain.

Recall that the applicability of the Cyclic Principle has been maximized by increasing the number of domains. In Chomsky's (1973) *Strict Cycle Condition* every XP is a cyclic node; and in the strongest version proposed by McCawley (1984; 1988) every projection is a cyclic node. The Cycle and hierarchy-governed approaches to ordering both demand that rules apply "from bottom to top" in the syntactic structure. Nevertheless, they make different predictions, in particular when it comes to cross-clausal movement. Let us assume that there is wh-movement out of an embedded clause that makes a stop-over (at least) in the embedded SpecCP and the matrix SpecvP on its way to the final landing site in matrix SpecCP. The movement step from the embedded SpecCP to the matrix SpecvP is in accordance with the Cyclic Principle because the domain in which the wh-phrase lands (matrix vP) includes the domain from which it has moved (embedded CP). However, this movement step is not allowed by hierarchy-governed approaches because movement must not only target a position that is higher in the structure, but one that is higher on the abstract hierarchy. But the S̄-domain from which movement starts (embedded SpecCP) is lower on the Inclusion Hierarchy than the Pred-domain in which the wh-phrase lands (matrix SpecvP).

5.2.2   On the status of universal principles

What is the status of the universal principles introduced in the preceding subsection? Note first that specificity and complexity can contradict each other. If a rule R_1 is less

---

[2]Abels proposes the hierarchy in (i) where θ-related operations outrank case-related operations (A-movement) which in turn outrank Ā-movement (there are even more-fine grained versions):

(i)   *The Universal Constraint on Operational Ordering in Language (UCOOL, Abels 2008: 60)*:
         Op ≻ A-mvt. ≻ θ

The following constraint restricts the sequence of movement operations:

(ii)   *Generalized Prohibition against Improper Movement (GenPIM, Abels 2008: 66)*
         No constituent may undergo movement of type τ if it has been affected by movement of type σ, where $\tau \prec \sigma$ under UCOOL.

Grohmann (2003) subdivides the relevant parts of the structure into what he calls *prolific domains*. These form the following hierarchy: Ω-domain (≈ CP) ≻ Θ-domain (≈ IP) ≻ Φ-domain (≈ vP). According to the Intraclausal Movement Generalization movement of an XP must always target the next higher prolific domain (the prohibition is not absolute, though; violations may occur but require a special repair).

complex than a rule R_2 in terms of its application context because the structural description of R_1 is contained in the structural description of R_2, then complexity requires the order $R_1 \succ R_2$ but specificity demands the order $R_2 \succ R_1$. If the corresponding principles are assumed to be universal, as Pullum (1979) did, there is an inherent conflict between specificity and complexity that needs to be resolved. This might be done by ranking the principles (cf. footnote 1); the ranking would need to be extrinsic if languages differ in which principle has priority.

What this also shows us is that the predicted order heavily depends on how specificity or complexity are measured. This becomes evident when have a look at what Chomsky (2013) says about the constraint *Merge before Move*. As outlined above, Merge before Move was originally assumed to follow from complexity in terms of the number of basic operations that are involved in Merge and Move (internal Merge): Move = Merge+ Agree + copy. However, Chomsky (2013: 41) claims that the opposite holds (Move before Merge) because external Merge (EM) is less complex in terms of the search space of an operation-inducing feature than internal Merge: "A residue of that error is the belief that EM is somehow simpler and preferable to IM. There is no basis for that belief. If anything, IM is simpler, since it requires vastly less search than EM (which must access the workspace of already generated objects and the lexicon)." Hence, depending on how complexity is measured, the predicted order of operations might be different.

As for Obligatory Precedence, this constraint cannot be easily applied anymore in Minimalism. As outlined in section 2.5.3.2, operations cannot be optional if they are triggered by features; only the *features* that trigger them could be optionally present of absent. But if there is an operation-inducing feature, it *must be* discharged. In this respect, all operations for which there is a trigger are obligatory and the Obligatory Precedence Principle cannot determine the order between them.[3]

To conclude, specificity, complexity, and obligatoriness are more or less problematic. What has been and still is considered to be a highly relevant principle in syntax is the Cyclic Principle in the version of the *Strict Cycle Condition*.

5.2.3 An argument for extrinsic ordering from reflexes of movement

What all the aforementioned principles have in common is the following: If there are two rules (or operations) R_1 and R_2 and they differ in some respect (complexity, specificity, etc.), then the principles predict either the order R_1 before R_2 or R_2 before R_1, viz. they predict an *asymmetric* order; if the rules have the same property (both are equally specific / complex, etc.), the principles do not make any predictions about their order. In the latter case, the order of the rules must thus be extrinsic (if there is evidence for the need of ordering in the first place, e.g. from transparent interactions).

What do the data presented in this thesis tell us about the extrinsic / intrinsic dichotomy? To see this, I repeat the orderings of Agree and Merge I have argued for in this thesis in (5) (exemplified with the order of final vs. intermediate movement steps):

(5)   Attested orders of final and intermediate movement triggers relative to a probe:
      a.   pattern I order   [•F•], [•EF•] $\succ$ [∗F∗]
      b.   pattern II order   [•F•] $\succ$ [∗F∗] $\succ$ [•EF•]
      c.   pattern III order [•EF•] $\succ$ [∗F∗] $\succ$ [•F•]
      d.   pattern IV order [∗F∗] $\succ$ [•F•], [•EF•]

---

[3]See section 2.5.3.2 for a proposal of how to deal with optionality in Minimalism.

Crucially, patterns I and IV as well as patterns II and III are the exact opposite of one another with respect to the order of the Merge operations. For pattern II, for example, we need the order Agree before intermediate internal Merge; for pattern III, we need the reverse order intermediate internal Merge before Agree. Hence, the principles under discussion would have to be able to predict both orders (symmetric order), but this is impossible because they can only produce asymmetric orders (viz. either (final / intermediate) Merge before Agree or Agree before (final / intermediate) Merge), regardless of how specificity, complexity, etc. are defined. Thus, the principles cannot predict the whole range of orderings required to account for cross-linguistic variation. Consequently, extrinsic ordering (parochial ordering in Pullum's 1979 sense) is needed after all: Two languages can differ solely in the order of rules / operations.

This conclusion also holds for the Cyclic Principle, the principle which plays a major role in determining rule ordering in syntax. Recall that I assume that the cyclic domain is the phrase (See section 3.3 for evidence). Since I investigate a configuration in which all interacting operations are triggered by the same head H and all operations cause changes within HP (Merge creates specifiers of H and upward Agree is restricted to a Spec-head relation), they necessarily apply within the same domain. Consequently, the Cycle (or its current version, the Strict Cycle Condition) does not predict an order between the operations; it only predicts the order of operations that apply on different cycles. But since an order is required because we get transparent interactions between operations triggered by a single head, the order must be extrinsic. Importantly, nothing changes if we adopt the stricter definition of cyclic domain by McCawley (1984; 1988), according to which every projection is a cyclic domain. Let me discuss what the Cycle would predict under this condition to see that the argument for extrinsic ordering is not simply an effect of having too large cyclic domains. The configuration I have investigated in this thesis is shown in (6): A head triggers both a Merge and an Agree operation.

(6)   H$'$

                H   XP
        [•F•], [∗F:□∗]
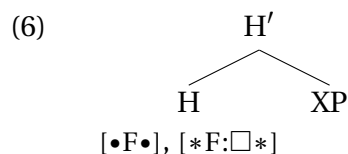
What is the domain of these two operations? This depends on the direction of Agree and at what point the domain is determined. First, assume that Agree applies downwards. Its domain is H$'$. The domain of Merge is less clear: At the point where Merge is triggered, the domain is H$'$ because no specifier is yet projected; but after the application of Merge, the domain is HP because the minimal projection that contains all nodes affected by the operation is the phrase in whose specifier an XP is merged.[4] If the domain of Merge is H$'$, it has the same domain as downward Agree. Consequently, the Cycle predicts no ordering between the two operations, contrary to what is required. If the domain of Merge is HP, the strictest version of the Cycle (with every projection a cyclic domain) predicts that downward Agree applies before Merge because HP includes H$'$. However, we have seen in chapter 3 that downward Agree can also apply after Merge. Next, assume that Agree applies upwards. Its domain is HP because the target is in SpecHP. If the domain of Merge is HP as well, the operations apply within the same domain and the Cycle predicts no ordering. If, however, the domain of Merge is H$'$, then the Cycle predicts that Merge precedes upward Agree because H$'$ is included in HP. To summarize, when Merge and Agree,

---

[4]I thank Gereon Müller for pointing out the alternative view to me, according to which Merge applies within HP when looking at the output of the operation.

triggered by a single head, apply within the same cyclic domain (which they always do if the domain is the phrase and which they sometimes do if every projection is a cyclic domain), then the Cyclic Principle is *too weak*: It predicts no ordering. If Merge and Agree, triggered by the same head, apply within different domains (which they may, depending on the direction of Agree and the domain of Merge), then the Cyclic Principle is *too strong*: It either predicts Merge before Agree or Agree before Merge, but we have seen evidence from variation that the same type of Merge applies before Agree in some, and after Agree in other languages. Under the assumption I have made throughout this thesis, viz. that every phrase is a cyclic domain, the Cycle is too weak. But given the range of attested orderings, this is actually a good result: If it made a prediction, say Merge-type X applies before Agree, it would wrongly rule out languages with the opposite order, Agree before Merge-type X. Consequently, extrinsic ordering is required.

Crucially, I do not claim that all orderings are extrinsic. The principles introduced above predict a large number of orderings, especially the Cycle, and this is desirable because the relevant orders need not be stored. Nevertheless, in the cases where the principles are too weak, extrinsic ordering is necessary. Hence, I agree with the first view on the extrinsic / extrinsic distinction presented in the introduction to the present section 5.2: The grammar requires both intrinsic and extrinsic ordering.

This insight is also important for the debate about sequential vs. simultaneous application of rules / operations, cf. chapter 1. Recall that simultaneous application can derive opaque interactions (because no intermediate representations are created), but not transparent interactions. Sequential application can derive both. This has led many researchers to conclude that the *Strict Ordering Hypothesis* is superior to the *Simultaneous Application Hypothesis*. However, as mentioned in section 1.2.1.2, mixtures of both simultaneous and sequential application have been proposed as well: Pullum (1979), for instance, assumes that all rules within a cyclic domain apply simultaneously by default; this derives all opaque interactions. In order to capture transparent interactions, Pullum proposes that an order is established between rules in a cycle if and only if a universal principle predicts an order between them. Furthermore, he claims that all orders that lead to transparent interactions are predicted by a universal principle, and hence, there is no need for extrinsic (parochial) ordering. However, the conclusion of the previous discussion was that there can be no principle that predicts all the orders in (5) because they are symmetric. Hence, the types of transparent interactions between Agree and Merge triggered by a single head cannot be derived in Pullum's (1979) mixed system; the operations would be predicted to apply simultaneously if there is no principle that requires an order between them, and therefore, feeding or bleeding interactions within a cyclic domain cannot be derived.

Finally, a historical remark: McCawley had hoped that under his very strict definition of cyclic domains there would be no cases where two rules apply within the same cycle. This was actually the case with the kind of rules / transformations he worked with, because they induced large-scale changes in various parts of the structure (see e.g. McCawley's definition of the Passive transformation in section 1.2.1.1, example (9)). But with the rise of elementary operations in Minimalism that induce only minimal changes in a very small domain, such a configuration arises whenever a head triggers more than one operation. To my knowledge, Řezáč (2004: 7), who investigates transparent interactions, was the first to point out that the Cycle does not make any predictions in this context: "Cycles are *ordered* with respect to each other; operations on a cycle, if there can be more than one, are not, by definition." He continues as follows: "All the operations triggered by features of a lexical item while it projects are on the same cycle. [...] So featural cyclicity

and other versions of the Locus Principle do not say anything about when features on a locus have to trigger operations." (p. 10). Thus, although the extrinsic / intrinsic distinction is the subject of a very old debate of which many people might think that there is nothing new to say about, we can indeed gain new insights about it from interactions of Minimalist operations.

To summarize, the universal principles that have been proposed in the literature to determine the order of rules / operations are not sufficient to predict all attested orderings of (different types of) Merge and Agree when the two operations are triggered by a single head. Either they make no predictions at all or they predict one order but then fail to predict the reverse order, which is, however, attested. Thus, extrinsic ordering is needed in addition to intrinsic ordering.

5.2.4   Extrinsic ordering and variation

If extrinsic ordering is adopted, the empirical and conceptual problems that Pullum (1979) raises emerge again (see the introduction to the present section 5.2). In particular, he worries about the large amount of variation that such an approach predicts; this variation is probably not attested and complicates language acquisition. Let me briefly elaborate on this objection to extrinsic ordering.

First of all, since I do not assume that all orderings are extrinsic, the variation is limited: There are not n! grammars for n rules / operations; most of the orderings still follow from independent principles of the grammar (intrinsic ordering). Only those orders required for interactions of operations triggered by a single head need to be ordered extrinsically.

Second, the variation is also restricted by the number of feature types. It remains to be seen how many Merge and Agree operations are required to derive all interactions (I have argued for about ten), but I expect that their number is considerably smaller than the number of transformations that have been proposed in the transformational grammar literature which the conceptual argument against extrinsic ordering is based on (Pullum takes as a reference point a grammar of English with 27 transformations). Furthermore, it may well be the case that not all the subtypes of Merge (and Agree) triggers I have argued for are present in the grammar of the child from the beginning. Rather, the child may start with just Merge and Agree and split them up into subtypes when evidence from (phonological, morphological, syntactic) reflexes of Merge operations suggests a distinction, viz. when a symmetric order of operations as in (7) is required to capture the data:

(7)   Merge > Agree > Merge

If the child only needs to order the subtypes of Merge for which it has independent evidence, then the number of grammars to choose from is much smaller as if all subtypes of Merge and Agree that may play a role in other languages are taken into account.

Third, I have shown that the variation through reordering of the different subtypes of Merge and Agree is indeed attested and it would thus be fatal to exclude one of them by a principle. McCawley (1988) discusses the variation issue on the basis of the counterbleeding interaction of Reflexivization and Imperative Subject Deletion (ISD) in English. If the order of the rules was extrinsic, he argues, we would expect there to be a dialect of English with the order ISD before Reflexivization, which would lead to the grammatical output *Defend you!*, meaning 'Defend yourself!' McCawley (1988: 157) supposes that "there is little evidence that any such variation actually exists. [...] no such dialect is known to exist, and it is very doubtful that such a dialect is even possible." As for the inter-

actions of operations I investigate, it was insightful to look at unrelated languages instead of dialects of the same language to see that the expected variation is indeed attested.

Note finally that extrinsic ordering is standardly used in Optimality theory (OT) (Prince and Smolensky 2004): The ranking of constraints is purely extrinsic.[5] Variation emerges through reordering; all reorderings instantiate possible grammars of natural languages (factorial typology). The present analysis with ordering of operation-inducing feature can easily be reformulated within OT if the operations are triggered by the need to fulfill an OT-constraint; the order of operations is implemented by the ranking of the corresponding constraints. See section 5.5.4 for details.

To summarize, both the conceptual and the empirical concerns with the postulation of extrinsic ordering are removed: The number of grammars predicted for n rules/operations is much smaller then n! because there is also intrinsic ordering, and because there is only a rather small set of subtypes of Merge in the grammar of an individual grammar (and not all possible subtypes). Moreover, the variation we expect from reordering of the subtypes of Merge I have argued for is indeed attested.

5.3   On the timing of edge-feature discharge

Opaque interactions of internal Merge and Agree can be derived by the order of operations if final and intermediate movement have designated triggers. In the present system where all operations are feature-driven and movement is based on Attract, this is implemented by postulating different structure-building features for these two types of internal Merge. Intermediate movement steps are triggered by edge features. Furthermore, I assume that edge features are not freely available on a phase head; rather, they are inserted if required (regulated by *Phase Balance*). Importantly, the ordering analysis of transparent and opaque interactions I have presented in this thesis sheds new light on the timing of edge feature insertion and edge feature discharge.

There are two explicit proposals in the literature on when edge features are inserted on a phase head, and consequently, when they can be discharged. On the one hand, Chomsky (20010: 109; 2001: 34) proposes that edge features can be inserted on the head of a phase if the phase is otherwise complete, i.e. if the phase head has discharged all of its inherent operation-inducing features. Consequently, edge features are the last features that a head discharges (intermediate movement steps are the last operation triggered by a phase head). On the other hand, Müller (2010: 37) suggests that the opposite holds (to derive a version of the CED from the PIC, see also Müller 2011): An edge feature can be inserted on a phase head as long as it is not complete, i.e. as long as it still has at least one operation-inducing feature. Müller assumes that features are ordered on a stack of which only the topmost feature is accessible. Thus, the edge feature is put on top of the stack. Being on top of the stack, it will be discharged directly after its insertion. The operation-inducing feature that is below the edge feature on the stack (and which keeps the head active) is discharged afterwards. It follows that intermediate movement steps cannot be the last operation triggered by a phase head; there will also be another feature on the stack below the edge feature that needs to be discharged after the edge feature.

Crucially, both of these proposals for the timing of edge feature insertion and discharge are in conflict with some of the orders I argued for in this thesis: We have seen that

---

[5]The only exception are rankings between basic constraints and constraints derived from them by Harmonic Alignment (Prince and Smolensky 2004; Aissen 1999; 2003) or Local Conjunction (Smolensky 1995); for these, there is a fixed order.

intermediate movement steps can apply at various points of the derivation: They can be the first or the last operation in a phase, or they can be interleaved with other operations in the phase. The variable order of operation-inducing features derives cross-linguistic variation (bleeding / feeding vs. counter-feeding / counter-bleeding). The attested orders that are ruled out by the two aforementioned approaches to edge feature insertion are listed in (8):

(8)   *Orderings that conflict with the timing of edge feature insertion*:
       a.   Intermediate movement steps are the last operation in a phase
            (Chomsky 2000; 2001):
            $\not\downarrow$ [•F•], [•EF•] ≻ [∗F:□∗]   *pattern I order*
            $\not\downarrow$ [•EF•] ≻ [∗F:□∗] ≻ [•F•]   *pattern III order*
       b.   Intermediate movement steps are not the last operation in a phase
            (Müller 2010; 2011):
            $\not\downarrow$ [•F•] ≻ [∗F:□∗] ≻ [•EF•]   *pattern II order*
            $\not\downarrow$ [∗F:□∗] ≻ [•F•] ≻ [•EF•]   *pattern IV order*

As a consequence, these proposals for the timing of edge feature discharge cannot hold universally – they do in some but not in all languages. The timing of edge feature discharge must be more flexible. The present proposal achieves this by assuming that edge features are inserted early (in the numeration), before the phase head has started to discharge any of its operation-inducing features. The features of the phase head, including the edge feature, are then ordered according to language-specific ordering statements; as a result, the edge feature can be on top of the stack (viz. will be discharged first), at the bottom of the stack (viz. will be discharged last) or in the middle (viz. will neither be discharged first nor last). It is only afterwards that the head begins to trigger operations, discharging the features on the stack from top to bottom.

5.4 A representational reanalysis?

I have shown in section 1.2.2 that many instances of opacity can be derived by enriched representations, viz. by postulating empty elements that encode previous stages of the derivation. For instance, the counter-bleeding interaction of Imperative Subject Deletion (ISD) and Reflexivization in (9-a) follows if the subject position is occupied by a silent pronoun as in (9-b):

(9)   *Interaction of ISD and Reflexivization*:
       a.   Wash yourself!
       b.   *pro* wash yourself!

Deletion of the subject should bleed Reflexivization of the coreferent object, but it does not; Reflexivization is possible. This follows given the enriched representation in (9-b) because there is an element (with the relevant features) in the subject position, it is just not visible. Can the counter-feeding and counter-bleeding interactions investigated in this thesis be reconstructed by just looking at the output if it contains empty elements in intermediate landing sites? I will show that this is possible for the counter-bleeding cases involving downward Agree, but it is not possible for counter-feeding with upward Agree.

       Let me begin with the counter-bleeding configuration. On the surface, there is an XP in SpecHP that has successfully entered into Agree with a probe on H. But since the

probe looks downward, it is unclear how XP and H could Agree, given that XP is not in the c-command domain of H. In the derivational ordering approach this follows because at the stage of the derivation where the probe seeks for a goal, XP is still in H's c-command domain; it is moved only afterwards. This information can be recovered in a representational framework if what corresponds to the pre-movement position of XP in the derivational model is occupied by an empty element that bears the features the probe searches for. This empty element is a copy in the structure in (10) (the copy is set in bold):

(10)   *Counter-bleeding configuration, enriched representation*:
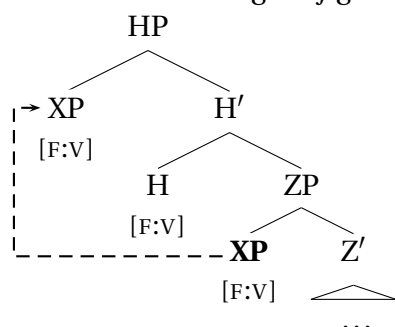
Given this representation, it follows why H is valued with the features of XP: Although XP is realized in SpecHP, there is a (silent) occurrence of XP in H's c-command domain.

Things are different for counter-feeding interactions where the probe on H searches upwards, as in pattern II languages. The configuration looks as follows: There is an XP in SpecHP with the relevant features the probe on H seeks for. Nevertheless, H does not bear the value of XP; it either has a default value or no value at all (default deletion). In the derivational approach this follows from the order of operations: XP has been moved to the search domain of the probe on H, viz. SpecHP, only after H has probed; movement comes too late to feed Agree. This interaction cannot simply be recovered from an enriched representation. To see this, look at the output structure in (11), an abstract pattern II configuration where the final movement step feeds Agree but intermediate movement steps do not:

(11)   *Counter-feeding configuration, enriched representation*:
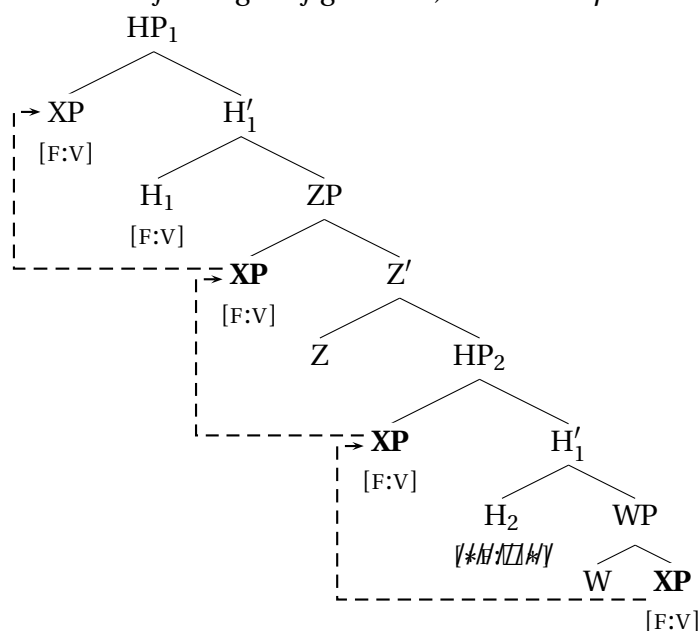
XP is moved successive-cyclically from WP to SpecHP$_1$. The head H bears a probe that can enter into upward Agree with XP, which is obvious in HP$_1$ where XP values the probe on H$_1$. However, the lower head H$_2$ does not Agree with XP; it bears a default value [def]. Why Agree fails is mysterious because there is an occurrence of XP (represented by the trace) in the right structural position SpecHP$_2$ to be a target for the probe on H$_2$. It cannot be read off of this enriched representation that XP reached this position only after the probe on H$_2$ initiated Agree.

Hence, the counter-feeding interactions of Merge and upward Agree investigated in this thesis are such that they cannot be accounted for by simply adding empty elements into positions a moving XP has occupied at what would correspond to earlier stages of the derivation – in contrast to most other instances of opacity.[6] What would be required to capture the opacity in a representational way would be even more complex representations and / or constraints on the empty elements. We have seen proposal along these lines in sections 2.4.4 and 2.4.5: Either different types of empty elements need to be postulated, or empty elements are subject to operations such as deletion or insertion of feature values.

In any case, more complex devices would be required to capture the counter-feeding interactions in a representational way. The derivational approach is much more simple because it neither requires empty elements nor constraints / operations on them. Counter-feeding interactions of Merge and upward Agree thus present a strong argument for a derivational model of syntax. Other opaque interactions that also do not fall out straightforwardly from enriched representations have been presented in Assmann et al. (2012) and Lechner (2010). The abstract configuration in the first paper is identical to the present counter-feeding configuration.

5.5   On the nature of the trigger of internal Merge

The present approach to reflexes of movement is feature-driven and Attract-based: All movement operations are triggered by a feature on a head that attracts an XP to its specifier. The difference between types of Merge is encoded in the triggers for Merge operations. But this is just one way to implement the distinction between movement types. In this section I will show that the analysis of patterns of reflexes of movement can be reformulated in approaches that make rather different basic assumptions, as long as there is *some distinction* between final and intermediate movement steps that the ordering statement can refer to. I will illustrate this for a Greed-based system and for an approach where intermediate movement steps are not feature-driven.

5.5.1   The need for designated triggers of intermediate steps revisited

In the preceding chapters I have argued for a distinction between different types of Merge triggers. Let me recapitulate the argument: The data suggest that Merge operations can apply at different points in the derivation relative to another operation, viz. Agree, on

---

[6]The same holds for the bleeding and the feeding interactions that arise under downward Agree investigated in chapter 3: In the bleeding case, there is an empty element in the search domain of the probe that would be a suitable goal, but valuation fails; in the feeding case, there is an empty element in the search domain of the probe that would be a defective intervenor, but valuation by a structurally lower DP succeeds. Feeding interactions involving upward Agree in intermediate positions, however, fall out straightforwardly with enriched representations: The trace occupies the structural position the probe has access to.

a single head. If Merge were a uniform operation it could be ordered either before or after Agree, as in (12-a) and (12-b). But the symmetric order in (12-c), for which the data provide evidence, could not be established.

(12)   a.   Merge > Agree
         b.   Agree > Merge
         c.   *Merge > Agree > Merge

The paradoxical situation in (12-c) is resolved if the Merge operation that applies before Agree is of a different type than the Merge operation that applies after Agree. I have implemented this distinction by postulating different operation-inducing features for the various types of Merge. With respect to the distinction between final and intermediate movement steps (internal Merge) argued for in chapters 2 and 3, this means that intermediate movement steps must have designated triggers. The present analysis is thus incompatible with approaches that assume that all movement steps in a movement chain are triggered by the *same* features, as proposed e.g. in McCloskey (2002) and Abels (2012). Abels *does* distinguish between different types of movement-inducing features such as triggers for wh-movement, focus movement, relativization etc. to capture asymmetries in the extraction potential of the corresponding phrases (cf. section 4.2), but he assumes that the operation-inducing features for one type of movement, say wh-movement, are the same on all heads that project a landing site for the wh-word, either final or intermediate. McCloskey (2002) states the following about the nature of the trigger of intermediate movement steps of successive-cyclic movement:

> "It is a feature of every such case that I have seen described that the morphosyntax of the intermediate positions is the same as the morphosyntax of the highest C-projection in the dependency (Chung 1998: 234ff.). This pattern is sufficiently general and clear that it would be wrong, in my view, to treat it as accidental or epiphenomenal. The conclusion it suggests is that the mechanisms which are in play in the intermediate positions of Ā-dependencies are the same as the mechanisms which are in play in the highest positions. In other words, if we interpret the morphosyntax as reflecting the presence of a movement-inducing feature on the topmost C of the dependency [...], then it is question-begging not to assume the same for the same morphosyntax in intermediate positions. This implies in turn that the intermediate movements must indeed also be feature-driven. [...] That is, I work from the assumption that since the morphosyntactic consequences of movement look the same in all positions – intermediate positions and topmost positions – we should assume that the same featural mechanisms are at play." (McCloskey 2002: 188)

McCloskey claims that intermediate movement steps must be feature-driven, and crucially, that intermediate and final movement steps in a movement chain are triggered by the same features, as the last sentence of the quote emphasizes. He bases this assumption on the observation that the reflexes in final and intermediate positions are morphologically always identical. The observation seems to be true for languages with a pattern I reflex, which McCloskey (2002) is concerned with. It does not hold, however, for languages with pattern II and pattern III reflexes: Clearly, there is a difference between final and intermediate landing sites in the respective languages. This empirical shortcoming of McCloskey's argumentation has been pointed out by Heck and Müller (2003: fn.20) (but they do not develop an analysis of reflexes of successive-cyclic movement). To conclude,

patterns II and III provide evidence that there must be some kind of distinction between final and intermediate movement steps. To postulate distinct movement triggers is one way to achieve this, but other implementations are possible, as I will show in the remainder of this section.

5.5.2 Feature-driven Attract

There are a number of proposals in the literature on movement which like the present approach assume that (i) movement is Attract-based (the moving XP is attracted by a higher head to its specifier), (ii) that all movement steps are feature-driven, and (iii) that intermediate movement steps have a designated trigger. The following list provides an overview over triggers for intermediate steps that have been proposed (in a wh-movement-chain, hence the terminology):

- pseudo-wh-features (Ferguson and Groat 1994; Collins 1997)
- focus features (Sabel 2000; Sabel and Zeller 2006)
- [-WH]-features (Müller 1999)
- D-features (Fanselow and Mahajan 2000)
- edge features (Chomsky 2000; 2001)

In all of these approaches, the final movement step (of wh-movement) is triggered by a different feature, [+WH], [OP], or [Q]. Hence, these approaches to successive-cyclic movement can easily account for the attested opaque patterns by ordering the final wh-feature and the respective intermediate feature differently relative to Agree, as in the present approach. The only thing we need to do is to replace the edge feature [•EF•] in the ordering statements by one of the features from the list above.

5.5.3 Greed-based theories

In Greed-based theories of movement, movement of an XP is not triggered by a feature on a functional head that requires checking by XP in a local specifier-head-relation; rather, movement of XP is triggered by a feature on the XP itself. A detailed Greed-based analysis of successive-cyclic movement is presented in Bošković (2007*a;b*). Bošković proposes that what basically drives (final) movement of an XP is a feature [UF] of XP. The reason is that [UF] acts as a probe that needs to c-command a matching feature [F] to get discharged (downward probing). If the matching feature is located on a head G that c-commands XP, XP moves to SpecGP; as a result, it c-commands G and can check its [UF]-feature. Intermediate movement steps do not apply in order to enable [UF]-checking. They apply in order to fulfill the PIC: By assumption, every phrase is a phase. If XP cannot check the [UF]-feature in its base position, it must move to the edge of the next higher phrase to remain accessible. If XP stayed in its base position, its [UF]-feature would be trapped in the complement domain of a phase head and could thus never be checked. In general, movement is subject to the version of *Last Resort* in (13):

(13)   *Last Resort (Bošković 2007b: 610)*:
        X undergoes movement iff without the movement, the structure will crash (with crash evaluated locally).

To summarize, intermediate movement steps are driven by the PIC and final movement steps are driven by the need to check [UF] on XP. Let me illustrate this for wh-movement.

Bošković (2007*b*: 631) assumes that the feature on the wh-phrase that drives movement is a focus feature [uFOC]. In the abstract configuration in (14-a) the wh-phrase XP is base-generated within the phase YP, but Y does not provide a matching feature for the uninterpretable feature on XP. Hence, XP moves to SpecYP to be remain accessible, cf. (14-b). [uFOC] is not checked in the intermediate position SpecYP.

(14)   *Intermediate movement steps*:

    a. $[_{YP}$ ...   Y   ...   XP]   *YP = phase*
                                    [uFOC]

    b. $[_{YP}$ XP $[_{Y'} ... Y ... t_{XP}]]$ *YP = phase*
                  [uFOC]

Let me now turn to the final movement step: If XP is in the specifier of a phase head Y, and afterwards the head G bearing a matching feature [FOC] is merged with YP, XP does not undergo PIC-driven movement because it is accessible to G (it is at the edge of the YP-phase), cf. (15-a). Final movement applies in order to bring XP into the closest position where its probe feature [uFOC] c-commands the matching feature on G, viz. SpecGP, cf. (15-b).

(15)   *Final movement step*:

    a. $[_{GP}$ G   [YP   XP $[_{Y'}$ ...   Y   ... $t_{XP}]]]$ *GP = phase*
            [FOC]   [uFOC]

    b. $[_{GP}$ XP $[_{G'}$ G   [YP t_{XP} $[_{Y'}$ $... Y ... t_{XP}]]]]$ *GP = phase*
           [u̶F̶O̶C̶]   [FOC]

To summarize, there is no structure-building feature on intermediate heads or the final head that attracts the wh-phrase.; all movement steps are enforced by [uF] on XP (greedy movement). Crucially, there is a difference between intermediate and final movement steps: The former apply to keep [uF] accessible and the latter to enable it to be checked with a feature in its c-command domain. This distinction between final and intermediate movement steps is enough to reformulate the ordering approach put forward in this thesis in Bošković's Greed system. In what follows, I will sketch such an analysis.

 Even if movement is driven by a requirement of the moving element (greedy movement), Agree is target-driven in Bošković's approach, triggered by a feature on a head. What we need to do in order to derive the patterns of reflexes is to order the Agree-triggering probe on a head H relative to the movement of an XP to SpecHP. To be able to order Agree by H before or after movement to SpecHP, the Agree-probe on H must "know" whether XP movement will be a final or an intermediate movement step. We can locally determine this on the head that bears the Agree-probe by making reference to the presence of a matching feature [F] on the same head. If H bears a feature that matches the feature [uF] on the XP that is to be moved, XP-movement will be a final movement step; if H does not bear a matching feature for XP, XP-movement will be an intermediate movement step.

 Let us assume for the sake of illustration that Agree applies upwards. To derive pattern I, Agree by a head H always applies *after* greedy movement of an XP to SpecHP has applied, regardless of whether H bears a matching feature for XP or not; consequently, movement will always feed Agree. The reverse holds for pattern IV: H always triggers Agree

first, *before* XP is moved to its specifier; thus, movement always counter-feeds Agree. As
for pattern II, a head that has a matching feature [F] for the feature [UF] on XP triggers
Agree *after* XP has moved to SpecHP; but if H does not have a matching feature, H trig-
gers Agree *before* XP-movement. As a result, intermediate movement steps come too late
to feed Agree with H, but final movement steps apply early enough. Pattern III is the
mirror image of pattern II: If H has a matching feature for XP, i.e. if XP movement to
SpecHP is final, Agree initiated by H is triggered *before* XP undergoes movement, result-
ing in counter-feeding. If H has no matching feature, Agree applies *after* XP undergoes
movement; hence, intermediate movement steps feed Agree.[7, 8]

Since Agree is target-driven and movement is greedy (controller-driven), the ap-
proach sketched above requires that operation-inducing features that are located on dif-
ferent nodes are ordered relative to each other: First, the head H needs to look into its c-
command domain to see whether there is an XP with a feature [UF]. Then, H must check
whether it has a matching feature [F]; finally, the ordering statements formulated above

---

[7]Note that Bošković (2007*a;b*) assumes that phase heads that trigger intermediate movement steps can-
not trigger Agree with the moving XP, i.e. there is no feature-checking in intermediate positions. Of course,
this assumption must be given up to translate the present analysis (where Agree is needed to produce a
reflex in intermediate positions) into Bošković's system. Note that this does not affect the main point of this
section which is to show that the crucial ingredient of the ordering approach is that final and intermediate
movement steps have different causes (triggers) so that they can apply at different stages of the derivation.
And this asymmetry is present in Bošković's analysis.

[8]The reader might wonder how Bošković accounts for pattern I reflexes if there is no feature checking
between a moving XP and intermediate phase heads (the same problem arises for pattern III, but he does
not address this pattern). As it stands, his analysis predicts only pattern II reflexes: There are intermediate
movement steps but since there is only Agree with the head that triggers the final movement step we only
see a reflex in the final clause. As for pattern I reflexes, Bošković (2007*a*), following an idea by Boeckx 2008*a*
(manuscript from 2004), proposes an *iterative prolepsis* approach: Instead of long-distance movement of
an operator with several intermediate landing sites and a single final landing site as in (i-a), the apparently
long Ā-dependency is broken down into a sequence of clause-bound dependencies as in (i-b):

(i)   a. $[_{CP} \text{Op}_k \dots [_{CP} \text{t}'_k \dots [_{CP} \dots \text{t}_k]]]$
         b. $[_{CP} \text{Op}_k \dots \text{t}_k \dots [_{CP} \text{OP}_n \dots \text{t}_n \dots [_{CP} \text{Op}_m \dots \text{t}_m]]]$

The argument of the most deeply embedded clause is generated as a dependent of the matrix verb and
undergoes short Ā-movement. It binds a null element that also undergoes short Ā-movement in the em-
bedded clause(s). Since movement to SpecCP is a final movement step in every clause (there is no cross-
clausal movement), every C head can Agree with the XP moved to its specifier and we get pattern I. Hence,
pattern I and pattern II emerge from different ways in which an apparently long Ā-dependency movement
can apply: Pattern II is the result of real long-distance movement with intermediate landing sites in inter-
mediate clauses, whereas pattern I arises from iterative prolepsis with a sequence of short movement steps.
Evidence for the prolepsis approach to pattern I comes from Kinande where there is evidence that there is
no true long-distance Ā-movement although we find agreement with the moving XP on the complemen-
tizer of every clause between the base position and the surface position of the moved XP: In contrast to
short Ā-movement, apparent long Ā-movement is not sensitive to (certain) islands and it does not show
reconstruction effects (cf. Schneider-Zioga 2009). Bošković follows Boeckx (2003; 2008*a*) who argues that it
is not clear whether there are any pattern I languages that have true agreement in intermediate positions,
i.e. languages for the derivation in (i-a) is required: On the one hand, the agreement relation is indirect in
several pattern I languages because the agreeing element does not Agree with the moving XP directly but
rather with the clause containing the XP (the Chamorro pattern). And other pattern I languages where we
do see agreement with the moving XP throughout are of the Kinande type, i.e. there is empirical evidence
against long-distance movement. However, these two options do not cover all languages with pattern I: We
have seen in section 2.3.1 that there are indeed many languages that are not of the Chamorro type and in
which the Ā-dependency that triggers the reflex does exhibit the typical properties of movement (island-
sensitivity, reconstruction, cross-over effects), in contrast to Kinande. For these, we cannot adhere to the
iterative prolepsis analysis; thus, we still need Agree in intermediate landing sites, contra Boeckx (2003;
1999) and Bošković (2007*a;b*).

apply and determine the order of operations. The Greed-based account of movement is slightly more complex in this respect than the purely Attract-based approach presented in this thesis where all operation-inducing features that need to be ordered are located on the same head. Nevertheless, the greedy approach can in principle order final vs. intermediate movement steps relative to Agree because the two types of movement have different causes. Although Bošković (2007*b*) does not address the issue, ordering of this type is required anyway in the Greed-approach: If a phase head H triggers Agree and if an XP needs to undergo movement to SpecHP (for whatever reason), we have the same indeterminacy as with purely target-driven approaches. Once H is merged, both Agree initiated by H and movement to SpecHP could apply, the context for both is given.

5.5.4 Non-feature-driven intermediate steps

Heck and Müller (2000; 2003) develop an analysis of successive-cyclic movement that is radically different from the approaches introduced in the preceding subsections in that they propose that intermediate movement steps are not feature-driven at all. The positive consequence of this assumption is that we can avoid postulating designated triggers for intermediate movement steps like pseudo-wh-features whose nature is far from clear. In contrast to intermediate movement steps, final movement steps are feature-driven. Hence, there is a difference in the trigger between the two types of movement that we can refer to when they need to be ordered differently relative to Agree. With this distinction, the approach by Heck and Müller (2000; 2003) can also capture the opacity effects that arise with patterns II and III, as I will illustrate in what follows.

Let me first introduce how (non-)feature-driven movement is implemented in Heck and Müller (2000; 2003). The analysis is formulated within the framework of Optimality Theory (OT, Prince and Smolensky 2004) where constraints are ranked and violable. Crucially, they adopt the Harmonic Serialism model (see McCarthy 2010 for an introduction): Optimization is not global, but rather local. Instead of evaluating the whole derivation, only small parts of the structure are evaluated; the optimal candidate of an evaluation E_n is sent back to Gen and the outputs of Gen based on the optimal input serve as the input for a subsequent evaluation. The generator and H-Eval alternate until there is no material left in the workspace. The relevant optimization domain is the phrase. As in this thesis, Merge is triggered by structure-building features [•F•]. The constraint LAST RESORT demands that internal Merge is feature-driven:

(16)   LAST RESORT (LR):
         Every Movement requires matching [F] and [•F•] at an edge.
         (the edge comprises a head H and SpecHP)

LR is satisfied for final movement steps where internal Merge of, say, a wh-phrase with the feature [WH] in SpecCP leads to matching and checking of the Merge-triggering feature [•WH•] on C (a final feature in my terminology). But in intermediate landing sites LR is violated because movement does not result in checking of a final feature; it is by assumption non-feature-driven. But how is an intermediate movement step enforced then? Heck & Müller propose that LR can be minimally violated by intermediate movement steps in favor of the satisfaction of a higher ranked constraint. The constraint in question is PHASE BALANCE (PB); Heck & Müller's version of it is given in (17):

(17)   PHASE BALANCE (PB): For every feature [•F•] in the numeration there must be an accessible feature [F] at the phase level.

(18)   *Accessibility*:
        A feature [F] is accessible if (i) or (ii) holds:

        (i)   [F] is on X or edgeX of the present root of the derivation.
        (ii) [F] is part of the workspace of the derivation.

Sticking to our wh-movement example, PB requires an XP with the feature [WH] to be accessible at the phase level for the corresponding structure-building feature [•WH•] in the numeration. Assume for the sake of illustration that vP and CP are phases. XP is accessible if it is either at the edge (the specifier) of the phase or if it is in the workspace (comprising the numeration and previously generated, but unconnected structures). Let us see what happens at the vP-level when the internal argument of V is a wh-phrase that will move to SpecCP: There is a C head in the numeration with the feature [•WH•]; for this feature there must be an accessible feature [WH]. If the internal argument is the only wh-phrase, i.e. there is no other wh-phrase in the numeration, the only chance to fulfill PB in the vP-phase is to move the wh-phrase to SpecvP, which is accessible from outside. At the same time, this movement step violates LR since v does not have a trigger for internal Merge. The violation of LR is non-fatal under the ranking PB $\gg$ LR. Hence, an intermediate movement step is a repair operation that applies in order to fulfill the high-ranked constraint PB. This is illustrated in (19) for the intermediate movement step of the internal argument of a transitive verb to SpecvP (wh-movement).

(19)   *Optimization of the vP*:

Input: [$_{VP}$ V DP$_{wh}$], DP$_{nom}$, v Numeration: { C$_{[•WH•]}$, ...}	PB	LR
O$_1$: [$_{vP}$ DP$_{nom}$ [$_{v'}$ v [$_{VP}$ V DP$_{wh}$]]]	*!	
☞O$_2$: [$_{vP}$ DP$_{wh}$ [$_{v'}$ DP$_{nom}$ [$_{v'}$ v [$_{VP}$ V t$_{wh}$]]]]		*

As for the final movement step, another constraint is required that enforces checking of a [•F•]-feature:

(20)   MERGE CONDITION (MC):
        Structure-building features ([•F•]) participate in Merge.

(21) illustrates the final movement step of the wh-phrase to SpecCP, leading to checking of [•WH•] on C. The wh-phrase is at the edge of TP in the input because is has undergone non-feature-driven movement to this position at the TP-cycle, enforced by PB:

(21)   *Optimization of the CP*:

Input: [$_{TP}$ DP$_{wh}$ [$_{T'}$ [$_{vP}$ t$_{wh}$ [$_{vP}$ DP$_{nom}$ [$_{v'}$ v [$_{VP}$ V t$_{wh}$]]]]]], C$_{[•WH•]}$ Numeration: { ... }	MC	PB	LR
O$_1$: [$_{CP}$ C$_{[•WH•]}$ [$_{TP}$ DP$_{wh}$ [$_{T'}$ T [$_{vP}$ t$_{wh}$ [$_{v'}$ DP$_{nom}$ ... [$_{VP}$ V t$_{wh}$]]]]]]	*!		
☞O$_2$: [$_{CP}$ DP$_{wh}$ [$_{C'}$ C [$_{TP}$ t$_{wh}$ [$_{T'}$ T [$_{vP}$ t$_{wh}$ [$_{v'}$ DP$_{nom}$ ... [$_{VP}$ V t$_{wh}$]]]]]]]			

In (21), MC forces movement of the wh-phrase to SpecCP where the [•WH•]-feature can be checked. PB is trivially fulfilled because there is no head with a structure-building feature left in the numeration. To summarize, MC causes final, feature-driven movement steps whereas PB $\gg$ LR causes intermediate non-feature-driven movement steps.

With the analysis for non-feature-driven intermediate movement steps in place, we need to let PB, LR, and MC, responsible for internal Merge, interact with Agree to derive the patterns of reflexes of movement. The interaction of Merge and Agree actually lends

itself to an optimality-theoretic treatment: Given the *Earliness Principle*, a conflict arises if a single head triggers both Agree and Merge. The two operations cannot apply simultaneously (otherwise there could be no transparent interactions), therefore, one operation must take priority over the other. Conflict resolution is a core property of OT; it is implemented via the ranking of constraints. Instead of saying that Merge applies before Agree (or vice versa), we can say that the constraint enforcing Merge is ranked above the constraint enforcing Agree (or vice versa). This also means that we need to evaluate every derivational step rather than phrases. Such an approach to the interaction of Merge and Agree is developed in Heck and Müller (2007; 2013). In addition to the constraint MC, which triggers Merge, they propose the constraint AC that triggers Agree (upward Agree for present purposes):[9]

(22) AGREE CONDITION (AC):
      Probes ([∗F∗]) participate in Agree.

Crucially, to be compatible with the analysis presented in this thesis where Agree can fail to result in valuation, I take AC to be satisfied if a probe is discharged, even if this involves default discharge (Agree must apply, but the derivation does not crash if it fails to result in valuation, cf. Preminger 2011*b*). If every step of the derivation is subject to optimization, the ranking of MC and AC predicts which of the two operations Merge and Agree applies first in case a head triggers both of them. This is illustrated in (23) for the ranking MC ≫ AC and in (24) for the ranking AC ≫ MC. The context is the following: v is merged with VP; v triggers both (external) Merge of DP_{ext} and Agree.

(23) MC ≫ AC (feeding):
      a. <u>Evaluation of v′</u>

Input: [$_{v′}$ v$_{[•D•],[∗F:□∗]}$ [$_{VP}$ V DP$_{int}$]]		
Numeration: { DP, ... }	MC	AC
O$_1$: [$_{v′}$ v$_{[•D•],}$ [$_{VP}$ V DP$_{int}$]]	*!	
☞O$_2$: [$_{vP}$ DP$_{ext}$ [$_{v′}$ v$_{,[∗F:□∗]}$ [$_{VP}$ V DP$_{int}$]]]		*

      b. <u>Evaluation of vP</u>

Input: [$_{vP}$ DP$_{ext}$ [$_{v′}$ v$_{[∗F:□∗]}$ [$_{VP}$ V DP$_{int}$]]] Numeration: { ... }	MC	AC
O$_{21}$: [$_{vP}$ DP$_{ext}$ [$_{v′}$ v$_{[∗F:□∗]}$ [$_{VP}$ V DP$_{int}$]]]		*!
☞O$_{22}$: [$_{vP}$ DP$_{ext}$ [$_{v′}$ v$_{[F:V]}$ [$_{VP}$ V DP$_{int}$]]]		

In (23-a) the ranking MC ≫ AC enforces that v first triggers Merge of DP_{ext} (as for O$_1$, AC is fulfilled because there has been an attempt to Agree, although no goal is found). In the subsequent evaluation (23-b), based on the optimal candidate O$_2$ of the previous evaluation, AC enforces Agree with DP_{ext} which results in valuation of the probe on v; MC is trivially fulfilled because v has no structure-building feature anymore. The result is feeding of Agree by early Merge. Under the reverse ranking AC ≫ MC, upward Agree has priority over Merge:

---

[9]The direction of Agree can be be encoded in a high-ranked constraint (see e.g. the constraint SPEC-HEAD BIAS in Heck and Müller 2013 that prefers Agree with the specifier of a head H instead of an element in the complement domain of H) or in the definition of Agree; I do not include it in the tableaux because it will not be violated in any of the derivations.

(24)   AC ≫ MC (counter-feeding):

    a. <u>Evaluation of v$'$</u>

Input: $[_{v'} \, v_{[\bullet D\bullet],[*F:\square*]} \, [_{VP} \, V \, DP_{int} \,]]$ Numeration: { DP, ... }		AC	MC
☞O$_1$: $[_{v'} \, v_{[\bullet D\bullet], \cancel{[*F:\square*]}} \, [_{VP} \, V \, DP_{int} \,]]$			*
O$_2$: $[_{vP} \, DP_{ext} \, [_{v'} \, v_{\cancel{[\bullet D\bullet]}, [*F:\square*]} \, [_{VP} \, V \, DP_{int} \,]]]$		*!	

    b. <u>Evaluation of vP</u>

Input: $[_{v'} \, v_{[\bullet D\bullet]} \, [_{VP} \, V \, DP_{int} \,]]$ Numeration: { DP, ... }		AC	MC
O$_{11}$: $[_{v'} \, v_{[\bullet D\bullet]} \, [_{VP} \, V \, DP_{int} \,]]$			*!
☞O$_{12}$: $[_{vP} \, DP_{ext} \, [_{v'} \, v_{\cancel{[\bullet D\bullet]}} \, [_{VP} \, V \, DP_{int} \,]]]$			

In (24-a), Agree applies but does not lead to valuation because there is no DP in SpecvP. Nevertheless, AC is fulfilled, and therefore, O$_1$ is optimal. In the subsequent optimization in (24-b), Merge of DP$_{ext}$ can apply because v does not have a probe feature anymore. The result is that Merge counter-feeds Agree because it applies too late.

       Thus, the order of Merge and Agree is derived by constraint ranking in OT with local optimization of every derivational step. Next, we have to combine this analysis with the one that is responsible for the distinction of final and intermediate movement steps presented in Heck and Müller (2000; 2003). Recall that intermediate movement steps are not feature-driven; rather they are enforced by the interaction of PB and LR. Final movement steps are triggered by MC. Hence, in order to derive patterns I – IV we simply need to order PB ≫ LR and MC relative to AC and evaluate every derivational step. Since final and intermediate movement steps have different causes (MC and PB ≫ LR, respectively), they can apply at different points relative to Agree, which gives rise to opacity. (25) lists all possible rankings of the relevant constraints PB, LR, MC and AC.[10]

(25) *Rankings of PB, LR, MC and AC*:

    a. MC ≫ PB ≫ LR ≫ AC
        Agree after final and intermediate movement steps
        ↪ both Merge types feed Agree, pattern I

    b. PB ≫ LR ≫ MC ≫ AC
        Agree after final and intermediate movement steps
        ↪ both movement types feed Agree, pattern I

    c. MC ≫ AC ≫ PB ≫ LR
        Agree after the final movement step and before intermediate movement steps
        ↪ only final steps feed Agree, pattern II

    d. PB ≫ LR ≫ AC ≫ MC
        Agree before the final movement step and after intermediate movement steps
        ↪ only intermediate steps feed Agree, pattern III

    e. AC ≫ MC ≫ PB ≫ LR
        Agree before the final movement step and before intermediate movement steps
        ↪ neither Merge type feeds Agree, pattern IV

---

[10]Further re-rankings arise if AC or MC are interleaved between PB and LR. But these orders will not lead to the selection of a different optimal candidate: What matters is the order of PB relative to AC and MC since it is PB that enforces intermediate steps. So both PB ≫ LR ≫ AC and PB ≫ AC ≫ LR result in feeding of Agree by intermediate movement steps.

    f.   AC ≫ PB ≫ LR ≫ MC
        Agree before the final movement step and before intermediate movement steps
        ↪ neither Merge type feeds Agree, pattern IV

Re-ranking of the constraints leads to the attested cross-linguistic variation. Note that some orders produce the same patterns because the order of MC and PB ≫ LR) relative to Agree is irrelevant for the morphological reflex if they both apply before of after Agree.

    To summarize, the transparent and opaque interactions of internal Merge and Agree can be derived in any system in which final and intermediate movement steps have different causes that can be ordered differently relative to Agree. This has been illustrated for an approach where movement is greedy and for an optimality-theoretic analysis where some movement steps are not feature-driven at all. In both, final movement steps apply in order to check a (final) feature, either on the moving XP or on a functional head; intermediate steps apply to fulfill the PIC, viz. to keep a feature on the moving XP accessible, but there is no feature-checking in intermediate positions.

5.6   Interim conclusions

In this chapter I have investigated the general theoretical implications of the ordering approach to reflexes of Merge. First, I have shown that the orderings that are needed to derive cross-linguistic variation must be extrinsic because a feature F_1 can be ordered *before* another feature F_2 in one language, but *after* F_2 in another language. Universal principles can, however, only predict one of the orders. Second, I have argued that the variable timing of intermediate movement steps suggests that they cannot generally be the last or the non-last operation in a phase. Furthermore, I have shown that counter-feeding opacity cannot simply be derived by enriched representations with empty elements in intermediate landing site; more complex representations would be required to do so. The ordering approach provides a much simpler analysis. Finally, I have illustrated that the ordering analysis can be reformulated in other approach as long as they make some distinction between the types of Merge I have argued for in this thesis. Therefore, the account does not provide an argument for a feature-driven Attract-based approach to movement, nor for the existence of edge features (or related features). Rather, it argues more generally for a split of Merge into subtypes and the need to order them relative to other operations.

Chapter 6

Conclusions

In this thesis I have investigated the nature of the syntactic operations Merge and Agree and their mode of application. I have been concerned with the configuration in (1), where a single head triggers both Merge ([•F•]) and Agree [∗F:□∗]:

(1)   H′
                      ⁄‾‾‾‾‾‾⁀
                  H   XP
              [•F•], [∗F:□∗]
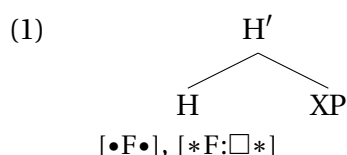
The existence of transparent interactions between these two operations requires that they apply sequentially rather than simultaneously: The application of Merge can feed or bleed a subsequent Agree operation. This can be derived if the operations are ordered relative to each other. The ordering is language-specific, viz. according to the literature there is a parameter that is set either to Merge > Agree or Agree > Merge. I have argued that we actually need a more fine-grained typology of Merge operations (as well as Agree operations). There is empirical evidence which suggests that Merge can apply both before and after Agree in a single language, cf. (2), because some Merge operations triggered by a head H feed or bleed Agree with H, whereas others have the opposite effect, viz. they counter-feed or counter-bleed Agree with H.

(2)   Merge > Agree > Merge

This symmetric order cannot be produced if Merge were a uniform operation; the parameter would order it either before or after Agree. This paradox is resolved if the Merge operations are of distinct types:

(3)   Merge$_1$ > Agree > Merge$_2$

Consequently, the parameter does not only order Merge and Agree, but rather subtypes of Merge. Therefore, the order of the Merge types relative to Agree can be different, leading to opacity.

I have used the configuration in (2) where Merge is interleaved with Agree to determine which types of Merge need to be distinguished. I have argued for a split between final and intermediate internal Merge, external vs. internal Merge, as well as for movement-type specific final and intermediate internal Merge. I have implemented this difference by postulating distinct triggers for these types that are ordered relative to the Agree trigger.

The primary empirical evidence for a split between final and intermediate internal Merge comes from the comparison of patterns of reflexes of movement found cross-linguistically: Under long-distance movement, a reflex of movement occurs in all clauses crossed by movement (pattern I), in others it occurs either only in the final clause (pattern II) or only in non-final clauses along the path of movement (pattern III). Patterns II and III strongly suggest a distinction between final and intermediate movement triggers. The patterns simply follow from reordering of the operation-inducing features, cf. the table (4) for feeding relations that arise under upward Agree (under downward Agree, we get either feeding or bleeding relations):

(4)   *Orderings of final vs. intermediate internal Merge and Agree*

	order of features	**interactions**		**pattern**
		final steps	intermediate steps	
a.	$[\bullet F\bullet]$, $[\bullet EF\bullet] \succ [*F:\square*]$	feed Agree		P I
b.	$[\bullet F\bullet] \succ [*F:\square*] \succ [\bullet EF\bullet]$	feed Agree	counter-feed Agree	P II
c.	$[\bullet EF\bullet] \succ [*F:\square*] \succ [\bullet F\bullet]$	counter-feed Agree	feed Agree	P III
d.	$[*F:\square*] \succ [\bullet F\bullet]$, $[\bullet EF\bullet]$	counter-feed Agree		P IV

I have shown that opacity effects of the type in (2) are found on virtually all functional heads in the clausal spine: v, T, C and n (a DP-internal functional head); I did not find any opaque interactions of Merge and Agree on P, but transparent interactions are attested (Baker 2008*b*: 192ff.).

In addition to a split of Merge types, we have also seen evidence for an equivalent split of Agree operations (person vs. number Agree, ϕ- vs. case Agree, class and definiteness vs. category Agree) because Agree can be interleaved with Merge:

(5)   Agree $\succ$ Merge $\succ$ Agree

I have argued that the opacity that arises through the interaction of Merge and Agree cannot straightforwardly be reconstructed in a representational way with enriched representations. Furthermore, previous approaches based on enriched representations require operations and constraints in addition to the postulation of (different types of) empty elements, and they cannot account for languages with mixed patterns.

The ordering approach is compatible with all Minimalist implementations of movement, in contrast to previous approaches. The insights of the analysis proposed in this thesis can also be integrated into approaches where intermediate movement steps are implemented differently, as long as they are distinguished in some way from final movement steps. Finally, I have shown that the valuation-based analysis of reflexes of movement is able to derive all types of reflexes: addition, deletion, replacement, syntactic reflexes.

The present approach has the following general implications: Since the present account relies on the order of operations on a single head, it argues for a strictly derivational model of grammar. Furthermore, the attested orderings provide evidence that Agree is a syntactic operation and for the need of extrinsic ordering in addition to intrinsic ordering: Some patterns involve the opposite order of operation-inducing features A and B, viz. we find both A $\succ$ B and B $\succ$ A; but universal principles, proposed to determine the order of operations (intrinsically), can only predict one of the orders. The attested variation in the timing of edge feature discharge suggests that intermediate movement steps are not universally the last or the non-last operation in a phrase, contrary to pre-

vious claims in the literature. Moreover, the timing of intermediate movement steps in Dinka, Hungarian, and Icelandic B also shows that edge features are inserted very early, before the head that will bear the edge feature has triggered any other Merge operations (other Merge triggering features counter-bleed edge feature insertion); hence, *Phase Balance* is not an output-oriented constraint. In addition, the ordering approach to reflexes of Merge provides a simple way to capture optionality: Optionality is the result of partial ordering, viz. some features have a variable order. As a consequence, we do not have to assume that operations or the features that induce them are optional (which is problematic in some cases). Rather, the operations always apply, what varies is their timing. As for the locality of movement, the present approach to patterns of reflexes of movement is compatible with the assumption that movement is subject to the same locality restrictions in all languages; long-distance movement does not have to apply in one-fell swoop in some languages to produce pattern II, but successive-cyclically in languages with pattern I. Rather, movement is always successive-cyclic, what differs is simply the timing of elementary operations.

Bibliography

Abels, Klaus (2003): Successive Cyclicity, Anti-Locality, and Adposition Stranding. PhD thesis, University of Connecticut, Storrs, Connecticut.

Abels, Klaus (2004): Right Node Raising: Ellipsis or Across the Board Movement?. *In:* K. Moulton and M. Wolf, eds, *Proceedings of NELS 34*. GLSA, Amherst, Mass, pp. 45–60.

Abels, Klaus (2008): 'Towards A Restrictive Rheory of (Remnant) Movement', *Linguistic Variation Yearbook* **7**, 53–120.

Abels, Klaus (2012): *Phases: An essay on cyclicity in syntax*. Linguistische Arbeiten 543, de Gruyter, Berlin.

Abels, Klaus and Kristine Bentzen (2011): Are Movement Paths Punctuated or Uniform?. *In:* A. Alexiadou, T. Kiss and G. Müller, eds, *Local Modelling of Non-Local Dependencies in Syntax*. Linguistische Arbeiten, De Gruyter, Berlin.

Abels, Klaus and Peter Muriungi (2008): 'The focus marker in Kîîtharaka: Syntax and semantics', *Lingua* **118**, 687–731.

Ackema, Peter and Ad Neeleman (2004): *Beyond Morphology: Interface Conditions on Word Formation*. Vol. 6 of *Oxford Studies in Theoretical Linguistics*, Oxford University Press, Oxford.

Adger, David, Susana Béjar and Daniel Harbour (2003): 'Directionality of Allomorphy: A Reply to Carstairs-McCarthy', *Transactions of the Philological Society* **101**, 109–115.

Aelbrecht, Lobke (2010): *The Syntactic Licensing of Ellipsis*. John Benjamins, Amsterdam.

Agüero-Bautista, Calixto (2007): 'Diagnosing Cyclicity in Sluicing', *Linguistic Inquiry* **38**, 413–443.

Aissen, Judith (1992): 'Topic and Focus in Mayan', *Language* **68**, 43–80.

Aissen, Judith (1999): 'Markedness and Subject Choice in Optimality Theory', *Natural Language and Linguistic Theory* **17**, 673–711.

Aissen, Judith (2003): 'Differential Object Marking: Iconicity vs. Economy', *Natural Language and Linguistic Theory* **21**, 435–483.

Aldridge, Edith (2008): 'Phase-based account of extraction in Indonesian', *Lingua* **118**, 1440–1469.

Alexiadou, Artemis, Elena Anagnostopoulou and Florian Schäfer (2006): The properties of anticausatives crosslinguistically. *In:* M. Frascarelli, ed., *Phases of Interpretation*. Mouton, Berlin, pp. 187–212.

Anagnostopoulou, Elena (2003): *The Syntax of Ditransitives: Evidence from Clitics*. Studies in Generative Grammar 54, Mouton de Gruyter, Berlin, New York.

Anand, Pranav and Andrew Nevins (2005): The locus of ergative case assignment: Evi-

dence from scope. *In:* A. Johns, D. Massam and J. Ndayiragije, eds, *Ergativity: Emerging issues.* Kluwer, Dordrecht, pp. 3–27.

Andersen, Torben (1991): 'Subject and topic in Dinka', *Studies in Language* **15**, 265–294.

Andersen, Torben (1995): 'Morphological stratification in Dinka: On the alternations of voice quality, vowel length and tone in the morphology of transitive verbal roots in a monosyllabic language', *Studies in African Linguistics* **23**, 1–63.

Andersen, Torben (2002): 'Case inflection and nominal head marking in Dinka', *Journal of African Languages and Linguistics* **23**, 1–30.

Anderson, Stephen R. (1969): West Scandinavian vowel systems and the ordering of phonological rules. PhD thesis, MIT, Cambridge, Mass.

Anderson, Stephen R. (1974): *The Organization of Phonology.* Academic Press, New York.

Anderson, Stephen R. (1992): *A-Morphous Morphology.* Cambridge University Press, Cambridge, Mass.

Aoun, Joseph (1985): *A grammar of anaphora.* Linguistic Inquiry Monographs 11, MIT Press, Cambridge, Mass.

Aoun, Joseph (1986): *Generalized binding.* Foris, Dordrecht.

Aoun, Joseph and Audrey Li (1989): 'Scope and constituency', *Linguistic Inquiry* **20**, 219–251.

Aoun, Joseph, Lina Choueiri and Norbert Hornstein (2001): 'Resumption, Movement, and Derivational Economy', *Linguistic Inquiry* **32**, 371–403.

Aronoff, Mark (1994): *Morphology by Itself.* MIT Press, Cambridge, Mass.

Arregi, Karlos and Andrew Nevins (2012): *Morphotactics: Basque Auxiliaries and the Structure of Spellout.* Springer, Dordrecht.

Asarina, Alya (2011): Case in Uyghur and beyond. PhD thesis, MIT, Cambridge, Mass.

Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller and Philipp Weisser (2012): Ergatives Move Too Early: On an Instance of Opacity in Syntax. Ms., Universität Leipzig, available at LingBuzz: lingbuzz/001646. To appear in *Syntax.*

Assmann, Anke, Fabian Heck, Johannes Hein, Stefan Keine and Gereon Müller (2010): Does Chain Hybridization in Irish Support Movement-Based Approaches to Long-Distance Dependencies. *In:* S. Müller, ed., *Proceedings of the HPSG 10 Conference.* CSLI, Stanford, pp. 27–46.

Assmann, Anke, Svetlana Edygarova, Doreen Georgi, Timo Klein and Philipp Weisser (2013): Case stacking below the surface: On the possessor case alternation in Udmurt. Ms., Universität Leipzig, to appear in *The Linguistic Review.*

Baker, Mark C. (1988): *Incorporation: A Theory of Grammatical Function Changing.* University of Chicago Press, Chicago.

Baker, Mark C. (2008*a*): 'On the Nature of the Antiagreement Effect: Evidence from *Wh*-in-Situ in Ibibio', *Linguistic Inquiry* **39**, 615–632.

Baker, Mark C. (2008*b*): *The Syntax of Agreement and Concord.* Vol. 115 of *Cambridge Studies in Linguistics*, Cambridge University Press, Cambridge.

Baker, Mark and Nadya Vinokurova (2010): 'Two Modalities of Case Assignment: Case in Sakha', *Natural Language and Linguistic Theory* **28**, 593–642.

Baković, Eric (1998): Optimality and Inversion in Spanish. *In:* P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky, eds, *Is the Best Good Enough?*. MIT Press, Cambridge, Mass., pp. 35–58.

Baković, Eric (2011): Opacity and Ordering. *In:* J. Goldsmith, J. Riggle and A. Yu, eds, *The Handbook of Phonological Theory (2nd ed)*. Wiley-Blackwell, Oxford, chapter 2, pp. 40–67.

Barbiers, Sjef (2002): Remnant Stranding and the Theory of Movement. *In:* A. Alexiadou, E. Anagnostopoulou, S. Barbiers and H.-M. Gärtner, eds, *Dimensions of Movement: From features to remnants*. Benjamins, Amsterdam, pp. 47–67. Linguistik aktuell (Linguistics today) Vol. 48.

Barbiers, Sjef, Olaf Koeneman and Marika Lekakou (2009): 'Syntactic doubling and the structure of wh-chains', *Journal of Linguistics* **46**, 1–46.

Barss, Andrew (1986): Chains and Anaphoric Dependence: On Reconstruction and Its Implications. PhD thesis, MIT, Cambridge, Mass.

Bartos, Huba (2001): Object agreement in Hungarian: a case for minimalism. *In:* G. Alexandrova and O. Arnaudova, eds, *The Minimalist Parameter*. John Benjamins, Amsterdam, pp. 311–324. Current issues in linguistic theory 192.

Béjar, Susana (2003): Phi-Syntax: A Theory of Agreement. PhD thesis, University of Toronto, Toronto.

Bejár, Susana and Diane Massam (1999): 'Multiple case checking', *Syntax* **2**, 65–79.

Béjar, Susana and Milan Řezáč (2009): 'Cyclic Agree', *Linguistic Inquiry* **40**, 35–73.

Bermúdez-Otero, Ricardo (2003): The acquisition of phonological opacity. *In:* J. Spenader, A. Eriksson and Ö. Dahl, eds, *Variation within Optimality Theory: Proceedings of the Stockholm Workshop on 'Variation within Optimality Theory'*. Stockholm University, pp. 25–36.

Bermúdez-Otero, Ricardo (to appear): *Stratal Optimality Theory*. Oxford Studies in Theoretical Linguistics, Oxford University Press, Oxford. Book Ms., University of Manchester.

Bianchi, Valentina (2004): Resumptive Relatives and LF Chains. *In:* L. Rizzi, ed., *The Cartography of Syntactic Structures. Volume 2.*. Oxford University Press, Oxford, pp. 76–114.

Bickel, Balthasar and Johanna Nichols (2009): Case marking and alignment. *In:* A. Malchukov and A. Spencer, eds, *The Oxford Handbook of Case*. Oxford University Press, Oxford, pp. 304–321.

Biloa, Edmond (1993): Clitic Climbing in Bantu. *In:* S. Mufwene and L. Moshi, eds, *Topics in African Linguistics. Papers from the XXI Annual Conference on African Linguistics, University of Georgia, April 1990*. John Benjamins, Amsterdam, pp. 67–77.

Biskup, Petr (2011): *Adverbials and the Phase Model*. John Benjamins, Amsterdam.

Bittner, Maria and Ken Hale (1996): 'The Structural Determination of Case and Agreement', *Linguistic Inquiry* **27**, 1–68.

Bobaljik, Jonathan (1993): Ergativity and Ergative Unergatives. *In:* C. Phillips, ed., *Papers on Case and Agreement II*. MITWPL, Cambridge, Mass, pp. 45–88.

Bobaljik, Jonathan (2000): The Ins and Outs of Contextual Allomorphy. *In:* K. Grohmann and C. Struijke, eds, *University of Maryland Working Papers in Linguistics*. Vol. 10, University of Maryland, College Park, pp. 35–71.

Bobaljik, Jonathan (2002): Syncretism without Paradigms: Remarks on Williams 1981, 1994. *In:* G. Booij and J. van Marle, eds, *Yearbook of Morphology 2001*. Kluwer, Dordrecht, pp. 53–85.

Bobaljik, Jonathan (2008): Where's Phi? Agreement as a Post-Syntactic Operation. *In:* D. Harbour, D. Adger and S. Béjar, eds, *Phi-Theory: Phi Features Across Interfaces and Modules*. Oxford University Press, Oxford, pp. 295–328.

Bobaljik, Jonathan and Susanne Wurmbrand (2005): 'The Domain of Agreement', *Natural Language and Linguistic Theory* **23**, 809–865.

Boeckx, Cedric (1999): 'Conflicting C-Command Requirements', *Studia Linguistica* **53**, 227–250.

Boeckx, Cedric (2003): *Islands and chains: resumption as stranding*. Linguistik aktuell (Linguistics today) Vol. 63, Benjamins, Amsterdam.

Boeckx, Cedric (2008*a*): *Bare syntax*. Oxford University Press, Oxford.

Boeckx, Cedric (2008*b*): *Understanding Minimalist Syntax: Lessons from Locality in Long-Distance Dependencies*. Vol. 9 of *Generative Syntax*, Blackwell Publishing, London.

Boeckx, Cedric and Howard Lasnik (2006): 'Intervention and Repair', *Linguistic Inquiry* **37**, 150–155.

Boeckx, Cedric and Kleanthes K. Grohmann (2007): 'Putting Phases in Perspective', *Syntax* **10**, 204–222.

Boeckx, Cedric, Norbert Hornstein and Jairo Nunes (2010): *Control as Movement*. Cambridge Studies in Linguistics 126, Cambridge University Press, New York.

Bonet, Eulália (1991): Morphology after Syntax. PhD thesis, MIT, Cambridge, Mass.

Bošković, Željko (2002): 'A-movement and the EPP', *Syntax* **5**, 167–218.

Bošković, Željko (2007*a*): On Successive Cyclic Movement and the Freezing Effect of Feature Checking. *In:* J. Hartmann, V. Hegedus and H. van Riemsdijk, eds, *Sounds of Silence: Empty Elements in Syntax and Phonology*. Elsevier, Amsterdam, pp. 195–233.

Bošković, Željko (2007*b*): 'On the Locality and Motivation of Move and Agree: An Even More Minimal Theory', *Linguistic Inquiry* **38**, 589–644.

Bošković, Željko and Howard Lasnik (1999): 'How Strict Is The Cycle', *Linguistic Inquiry* **30**, 691–703.

Bossong, Georg (1985): *Differenzielle Objektmarkierung in den neuiranischen Sprachen*. Narr, Tübingen.

Brandi, Luciana and Patrizia Cordin (1998): Two Italian dialects and the Null Subject Parameter. *In:* O. Jaeggli and K. Safir, eds, *The Null Subject Parameter*. Kluwer Academic Publishers, Dordrecht, pp. 111–142.

Branigan, Phil (1992): Subjects and Complementizers. PhD thesis, MIT, Cambridge, Mass.

Branigan, Phil (2011): *Provocative Syntax*. Linguistic Inquiry Monographs 61, MIT Press.

Bresnan, Joan (1976): 'On the Form and Functioning of Transformations', *Linguistic Inquiry* **7**, 3–40.

Bresnan, Joan (1978): *Contraction and the Transformational Cycle in English*. Indiana University Linguistics Club, Bloomington.

Brody, Michael (2002): On the Status of Representations and Derivations. *In:* S. Epstein

and T. Seely, eds, *Derivation and Explanation in the Minimalist Program.* Blackwell, Oxford, pp. 19–41.

Browning, Marguerite (1991): 'Bounding Conditions on Representations', *Linguistic Inquiry* **22**, 541–562.

Bruening, Benjamin (2005): The Algonquian Inverse is Syntactic: Binding in Passamaquoddy. Ms., University of Delaware.

Bruening, Benjamin (2006): 'Differences Between the Wh-Scope-Marking and Wh-Copy Constructions in Passamaquoddy', *Linguistic Inquiry* **37**, 25–49.

Burzio, Luigi (1986): *Italian Syntax: A government-binding approach.* Reidel, Dordrecht.

Bye, Patrick (2009): Virtual Phonology: Rule Sandwiching and Multiple Opacity in North Saami. PhD thesis, Universitetet i Tromsø, Tromsø.

Campos, Héctor (1997): 'On Subject Extraction and the Antiagreement Effect in Romance', *Linguistic Inquiry* **28**, 92–119.

Carstairs-McCarthy, Andrew (2001b): 'Grammatically Conditioned Allomorphy, Paradigm Structure, and the Ancestry Constraint', *Transactions of the Philological Society* **99**, 223–245.

Carstens, Vicki (2003): 'Rethinking Complementizer Agreement: Agree with a Case-Checked Goal', *Linguistic Inquiry* **34**, 393–412.

Carstens, Vicki (2005): 'Agree and EPP in Bantu', *Natural Language and Linguistic Theory* **23**, 219–279.

Carstens, Vicki (2012): Delayed Valuation: a reanalysis of "upwards" complementizer agreement and the mechanics of Case. To appear in *Syntax.*

Castillo, Juan Carlos, John E. Drury and Kleanthes K. Grohmann (2009): 'Merge Over Move and the Extended Projection Principle: MOM and the EPP Revisited', *Iberia* **1**, 53–114. Previously published in 1999 as *Merge over Move and the Extended Projection Principle* in University of Maryland Working Papers in Linguistics 8.

Cheng, Lisa Lai-Shen (1991): On the Typology of Wh-Questions. PhD thesis, MIT, Cambridge, Mass.

Cheng, Lisa Lai-Shen (2006): Decomposing Bantu Relatives. *In:* C. Davis, A. R. Deal and Y. Zabbal, eds, *Proceedings of NELS 36.* GLSA, Amherst, Mass, pp. 197–215.

Chomsky, Noam (1951): Morphophonemics of Modern Hebrew. Master's thesis, University of Pennsylvania.

Chomsky, Noam (1957): *Syntactic Structures.* Mouton, The Hague and Paris.

Chomsky, Noam (1964): *Current Issues in Linguistic Theory.* Mouton & Co., The Hague.

Chomsky, Noam (1965): *Aspects of the Theory of Syntax.* MIT Press, Cambridge, Mass.

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 (1975): *The Logical Structure of Linguistic Theory.* Plenum Press, New York.

Chomsky, Noam (1976): 'Conditions on Rules of Grammar', *Linguistic Analysis* **2**, 303–351.

Chomsky, Noam (1977): On Wh-Movement. *In:* P. Culicover, T. Wasow and A. Akmajian, eds, *Formal Syntax.* Academic Press, New York, pp. 71–132.

Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, Noam (1982): *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press, Cambridge, Mass.

Chomsky, Noam (1986*a*): *Barriers*. Linguistic Inquiry Monographs 13, MIT Press, Cambridge, Mass.

Chomsky, Noam (1986*b*): *Knowledge of language: its nature, origin and use*. Praeger, New York.

Chomsky, Noam (1991): Some Notes on Economy of Derivation and Representation. *In:* R. Freidin, ed., *Principles and Parameters in Comparative Grammar*. MIT Press, Cambridge, Mass., pp. 417–454.

Chomsky, Noam (1993): A Minimalist Program for Syntactic Theory. *In:* K. Hale and S. J. Keyser, eds, *The View from Building 20*. MIT Press, Cambridge, Mass, pp. 1–52.

Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Mass.

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. *In:* M. Roger, M. David and J. Uriagereka, eds, *Step by Step*. MIT Press, Cambridge, Mass., 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 (2004): Beyond Explanatory Adequacy. *In:* A. Belletti, ed., *Structures and Beyond. The Cartography of Syntactic Structures Vol. 3*. Oxford University Press, Oxford.

Chomsky, Noam (2005): 'Three Factors in Language Design', *Linguistic Inquiry* **36**, 1–22.

Chomsky, Noam (2008): On Phases. *In:* R. Freidin, C. Peregrín Otero and M. L. Zubizarreta, eds, *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*. MIT Press, pp. 133–166.

Chomsky, Noam (2013): 'Problems of projection', *Lingua* **130**, 33–49.

Chomsky, Noam and Morris Halle (1968): *The sound pattern of English*. Harper & Row, New York.

Chomsky, Noam, Morris Halle and Fred Lukoff (1956): On Accent and Juncture in English. *In:* M. Halle, H. Lunt and C. McLean, H. van Schooneveld, eds, *For Roman Jakobson. Essays on the occasion of his sixtieth birthday*. Mouton, The Hague, pp. 65–80.

Chung, Sandra (1982): 'Unbounded Dependencies in Chamorro Grammar', *Linguistic Inquiry* **13**, 39–77.

Chung, Sandra (1983): 'The ECP and government in Chamorro', *Natural Language and Linguistic Theory* **1**, 207–244.

Chung, Sandra (1991): Sentential Subjects and Proper Government in Chamorro. *In:* C. Georgopoulos and R. Ishihara, eds, *Interdisciplinary approaches to language: essays in honor of S.-Y. Kuroda*. Kluwer, Dordrecht, pp. 75–99.

Chung, Sandra (1994): 'Wh-Agreement and "Referentiality" in Chamorro', *Linguistic Inquiry* **25**, 1–44.

Chung, Sandra (1998): *The Design of Agreement: Evidence from Chamorro*. University of Chicago Press, Chicago.

Chung, Sandra and James McCloskey (1987): 'Government, barriers and small clauses in Modern Irish', *Linguistic Inquiry* **18**, 173–237.

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

Clements, George (1984*a*): 'Binding Domains in Kikuyu', *Studies in the Linguistic Sciences* **14**, 37–56.

Clements, George (1984*b*): Principles of Tone Assignment in Kikuyu. *In:* G. Clements and J. Goldsmith, eds, *Autosegmental Studies in Bantu Tone*. Foris, Dordrecht, pp. 281–339. Publications in African Languages and Linguistics.

Clements, George (2000): 'In defense of serialism', *The Linguistic Review* **17**, 181–198.

Clements, George, James McCloskey, Joan Maling and Annie Zaenen (1983): 'String-Vacuous Rule Application', *Linguistic Inquiry* **14**, 1–17.

Clements, George and Kevin Ford (1979): 'Kikuyu Tone Shift and its Synchronic Consequences', *Linguistic Inquiry* **10**, 179–210.

Cole, Peter (1982): 'Subjacency and Successive Cyclicity: Evidence from Ancash Quechua', *Journal of Linguistic Research* **2**, 35–58.

Cole, Peter and Gabriella Hermon (1998): 'The typology of wh-movement and wh-questions in Malay', *Syntax* **1**, 221–258.

Cole, Peter and Gabriella Hermon (2000): Partial Wh-Movement: Evidence from Malay. *In:* U. Lutz, G. Müller and A. von Stechow, eds, *Wh-Scope Marking*. John Benjamins, Amsterdam, pp. 101–130.

Cole, Peter, Gabriella Hermon and April Yanti (2006): Voice in Malay/Indonesian. Paper Presented at The Tenth International Symposium on Malay/Indonesian (ISMIL 10), University of Delaware.

Collins, Chris (1997): *Local Economy*. MIT Press, Cambridge, Mass.

Collins, Chris (2004): The agreement parameter. *In:* A. Breitbarth and H. van Riemsdijk, eds, *Triggers*. Mouton de Gruyter, Berlin, pp. 115–136.

Collins, Christopher (1993): Topics in Ewe Syntax. PhD thesis, MIT, Cambridge, Mass.

Collins, Christopher (1994): 'Economy of Derivation and the Generalized Proper Binding Condition', *Linguistic Inquiry* **25**, 45–61.

Comrie, Bernard (1979): 'Definite and animate direct objects: A natural class', *Linguistica silesiana* **3**, 13–21.

Corbett, Greville and Norman Fraser (1993): 'Network Morphology: A DATR Account of Russian Nominal Inflection', *Journal of Linguistics* **29**, 113–142.

D'Alessandro, Roberta and Ian Roberts (2008): 'Movement and Agreement in Italian Past Participles and Defective Phases', *Linguistic Inquiry* **39**, 477–491.

Davis, Stuart (2000): 'Some analytical issues in Optimality Theory', *The Linguistic Review* **17**, 117–134.

de Vries, Marc (2006): 'Possessive Relatives and (Heavy) Pied Piping', *The Journal of Comparative Germanic Linguistics* **9**, 1–52.

de Vries, Marc (2013): 'Multidominance and Locality', *Lingua* **134**, 149–169.

Dékány, Éva Katalin (2011): A profile of the Hungarian DP. The interaction of lexicalization, agreement and linearization with the functional sequence. PhD thesis, University of Tromsø, CASTL, Tromsø.

Delsing, Lars-Olof (1998): Possession in Germanic. *In:* A. Alexiadou and C. Wilder, eds, *Possessors, Predicates and Movement in the Determiner Phrase.* Benjamins, Amsterdam, pp. 87–109.

Demirdache, Hamida (1991): Resumptive Chains in Restrictive Relatives, Appositives and Dislocation Structures. PhD thesis, MIT.

den Besten, Hans (1977): On the interaction of root transformations and lexical deletive rules. *In:* W. Abraham, ed., *On the Formal Syntax of the Westgermania.* John Benjamins, Amsterdam, pp. 47–131.

den Dikken, Marcel (1999): On the structural representation of possession and agreement: The case of (anti-)agreement in Hungarian possessed nominal phrases. *In:* I. Kenesei, ed., *Crossing boundaries: Advances in the theory of Central and Eastern European languages.* Benjamins, Amsterdam, pp. 137–178.

den Dikken, Marcel (2007): 'Phase Extension: Contours of a theory of the role of head movement in phrasal extraction', *Theoretical Linguistics* **33**, 1–41.

den Dikken, Marcel (2009*a*): 'Arguments for successive-cyclic movement through SpecCP: A critical review', *Linguistic Variation Yearbook* **9**, 89–126.

den Dikken, Marcel (2009*b*): On the nature and distribution of successive cyclicity. Adjunction, resumption, and scope marking as the roads to success in long-distance relation building. Ms., CUNY.

den Dikken, Marcel (2012): On the strategies for forming long Ā-dependencies: Evidence from Hungarian. Ms., CUNY.

Déprez, Viviane (1998): 'Semantic effects of agreement: the case of French past participle agreement', *Probus* **10**, 1–65.

Diercks, Michael (2010): Agreement with subjects in Lubukusu. PhD thesis, Georgetown University, Washington, DC.

Diesing, Molly (1992): *Indefinites.* MIT Press, Cambridge, Mass.

DiSciullo, Anna-Maria and Edwin S. Williams (1987): *On the Definition of Word.* MIT Press, Cambridge, Mass.

Doliana, Aaron (2013): 'On the Person-Case Constraint: From the Giga to the Zero Version with Copy Impoverishment and Check', BA thesis, Universität Leipzig.

du Plessis, Hans (1977): 'Wh Movement in Afrikaans', *Linguistic Inquiry* **8**, 723–726.

Dukes, Michael (1992): On the Status of Chamorro Wh-Agreement. *In:* J. Maed, ed., *Proceedings of the Eleventh West Coast Conference on Formal Linguistic.* CSLI Publications, Stanford, pp. 177–190.

Embick, David (2000): 'Features, Syntax, and Categories in the Latin Perfect', *Linguistic Inquiry* **31**, 185–230.

Embick, David (2010): *Localism versus Globalism in Morphology and Phonology.* Linguistic Inquiry Monographs 60, MIT Press, Cambridge, Mass.

Embick, David and Rolf Noyer (2001): 'Movement Operations after Syntax', *Linguistic Inquiry* **32**, 555–595.

Epée, Roger (1975): The Case for a Focus Position in Duala. *In:* R. Herbert, ed., *Proceedings of the Sixth Conference on African Linguistics, Ohio State University, April 1975.* Ohio State University, Columbus, pp. 210–226.

Epée, Roger (1976*a*): Generative Syntactic Studies in Duala. PhD thesis, Cornell University, NY.

Epée, Roger (1976*b*): 'On Some Rules That Are Not Successive Cyclic in Duala', *Linguistic Inquiry* **7**, 193–198.

Epstein, Samuel D. (2007): 'On I(nternalist) Functional Explanation in Minimalism', *Linguistic Analysis* **33**, 20–53.

Epstein, Samuel D. and Daniel T. Seely (2002): Rule Applications as Cycles in a Level-Free Syntax. *In:* S. Epstein and T. Seely, eds, *Derivation and Explanation in the Minimalist Program.* Blackwell, Oxford, pp. 65–89.

Epstein, Samuel D., Erich Groat, Ruriko Kawashima and Hisatsugu Kitahara (1998): *A Derivational Approach to Syntactic Relations.* Oxford University Press, Oxford and New York.

Fanselow, Gisbert (1991): 'Minimale Syntax', Habilitation thesis, Universität Passau.

Fanselow, Gisbert (2006): Partial wh-movement. *In:* M. Everaert and H. van Riemsdijk, eds, *The Blackwell Companion to Syntax.* Blackwell, Oxford, pp. 437–492.

Fanselow, Gisbert and Anoop Mahajan (2000): Towards a Minimalist Theory of Wh-Expletives, Wh-Copying, and Successive Cyclicity. *In:* U. Lutz, G. Müller and A. von Stechow, eds, *Wh-Scope Marking.* John Benjamins, Amsterdam, pp. 195–230.

Felser, Claudia (2004): 'Wh-copying, Phases and Successive Cyclicity', *Lingua* **114**, 543–574.

Ferguson, Scott and Erich Groat (1994): Defining "Shortest move". Ms., Harvard University. Paper presented at GLOW 17 in Vienna.

Fiengo, Robert (1974): Semantic Conditions on Surface Structure. PhD thesis, MIT, Cambridge, Mass.

Fiengo, Robert (1977): 'On Trace Theory', *Linguistic Inquiry* **8**, 35–61.

Finer, Daniel L. (1985): 'The Syntax of Switch-Reference', *Linguistic Inquiry* **16**, 35–55.

Fischer, Silke (2004): Towards an Optimal Theory of Reflexivization. PhD thesis, University of Tübingen.

Fox, Danny (1999): 'Reconstruction, Binding Theory, and the Interpretation of Chains', *Linguistic Inquiry* **30**, 157–196.

Fox, Danny (2000): *Economy and Semantic Interpretation.* MIT Press, Cambridge, Mass.

Fox, Danny (2002): 'Antecedent-contained deletion and the copy theory of movement', *Linguistic Inquiry* **33**, 63–96.

Frampton, John (2002): Syncretism, Impoverishment, and the Structure of Person Features. *In:* M. Andronis, E. Debenport, A. Pycha and K. Yoshimura, eds, *Papers from the Chicago Linguistics Society Meeting.* Vol. 38, Chicago, pp. 207–222.

Frampton, John (2004): Copies, Traces, Occurrences, and All That. Ms., Northeastern University.

Frampton, John (2009): *Distributed Reduplication.* Linguistic Inquiry Monographs 52, MIT Press, Cambridge, Mass.

Freidin, Robert (1978): 'Cyclicity and the Theory of Grammar', *Linguistic Inquiry* **9**, 519–549.

Freidin, Robert (1986): Fundamental issues in the theory of binding. *In:* B. Lust, ed., *Studies in the Acquisition of Anaphora, Vol. I*. Reidel, Dordrecht, pp. 151–188.

Freidin, Robert (1994): Generative Grammar: Principles and Parameters Framework. *In:* R. Asher and J. Simpson, eds, *The Encyclopedia of Language and Linguistics, vol. III*. Pergamon Press, Oxford, pp. 1370–1385.

Freidin, Robert (1999): Cyclicity and Minimalism. *In:* S. Epstein and N. Hornstein, eds, *Working Minimalism*. MIT Press, Cambridge, Mass., pp. 95–126.

Fuß, Eric (2007): Cyclic Spell-Out and the domain of post-syntactic operations: Evidence from complementizer agreement. *In:* K. Grohmann, ed., *Dynamic interface*. Special issue of Linguistic Analysis Vol. 33.3-4, pp. 267–302.

Gallego, Ángel J. (2005): Phase sliding. Ms., University of Barcelona and University of Maryland.

Gallego, Ángel J. (2010): 'Binding through Agree', *Linguistic Analysis* **34**, 163–192.

García Mayo, María Del P. (1995): 'The licensing of parasitic gaps', *Journal of English and American Studies* **16**, 125–140.

Gärtner, Hans-Martin (2002): *Generalized Transformations and Beyond*. Akademie-Verlag, Berlin.

Gavruseva, Elena (2000): 'On the syntax of possessor extraction', *Lingua* **110**, 743–772.

Gazdar, Gerald (1981): 'Unbounded Dependencies and Coordinate Structure', *Linguistic Inquiry* **12**, 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum and Ivan Sag (1985): *Generalized Phrase Structure Grammar*. Blackwell, Oxford.

Georgi, Doreen (2013): Opaque Interaction of Internal Merge and Agree. *In:* F. Heck and A. Assmann, eds, *Rule Interaction of Grammar*. Universität Leipzig, Institut für Linguistik, pp. 413–461. Vol. 90 of *Linguistische Arbeitsberichte*.

Georgi, Doreen and Martin Salzmann (2011): 'DP-internal double agreement is not double Agree: Consequences of Agree-based case assignment within DP', *Lingua* **121**, 2069–2088.

Georgopoulos, Carol (1985): 'Variables in Palauan Syntax', *Natural Language and Linguistic Theory* **3**, 59–94.

Georgopoulos, Carol (1991): *Syntactic Variables: Resumptive Pronouns and A′ Binding in Palauan*. Studies in Natural Language and Linguistic Theory 24, Kluwer Academic Publishers, Dordrecht.

Goldrick, Matt (2000): Turbid Output Representations and the Unity of Opacity. *In:* M. Hirotani, A. Coetzee, N. Hall and J.-Y. Kim, eds, *Proceedings of NELS 30*. GLSA, Amherst, Mass, pp. 231–245.

Gracanin-Yuksek, Martina (2009): What And Why Can't Be Shared. Ms., Middle East Technical University,.

Grimshaw, Jane (1997): 'Projection, Heads, and Optimality', *Linguistic Inquiry* **28**, 373–422.

Grinder, John (1972): On the Cycle in Syntax. *In:* J. Kimball, ed., *Syntax and Semantics I*. Academic Press, New York, pp. 81–112.

Grohmann, Kleanthes (2003): *Prolific Domains: On the Anti-Locality of Movement Depen-*

dencies. John Benjamins, Amsterdam, Philadelphia.

Haddad, Youssef and Susanne Wurmbrand (2013): Cyclic Spell-Out and Agreement Options in Raising. Ms., University of Florida and University of Connecticut.

Haïk, Isabelle (1990): 'Anaphoric, Pronominal and Referential INFL', *Natural Language and Linguistic Theory* **8**, 347–374.

Haïk, Isabelle, Hilda Koopman and Dominique Sportiche (1985): 'INFL en mooré et le liage dans le système A', Rapport de recherches du groupe de linguistique africaniste, année 1985-1986. Montréal.

Hale, Ken (1972): A New Perspective on American Indian Linguistics. *In:* A. Ortiz, ed., *New Perspectives on the Pueblos.* University of New Mexico Press, Albuquerque, pp. 87–103.

Halle, Morris (1997): Distributed Morphology: Impoverishment and Fission. *In:* B. Bruening, Y. Kang and M. McGinnis, eds, *Papers at the Interface.* Vol. 30, MITWPL, pp. 425–449.

Halle, Morris and Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. *In:* K. Hale and S. J. Keyser, eds, *The View from Building 20.* MIT Press, Cambridge, Mass., pp. 111–176.

Halle, Morris and Alec Marantz (1994): Some Key Features of Distributed Morphology. *In:* A. Carnie, H. Harley and T. Bures, eds, *Papers on Phonology and Morphology.* Vol. 21 of *MIT Working Papers in Linguistics*, MITWPL, Cambridge, Mass., pp. 275–288.

Halpert, Claire (2012): Argument licensing and agreement in Zulu. PhD thesis, MIT, Cambridge, Mass.

Harley, Heidi (2009): The morphology of nominalizations and the syntax of vP. *In:* M. Rathert and A. Giannankidou, eds, *Quantification, Definiteness and Nominalization.* Oxford University Press, Oxford, pp. 320–342.

Harley, Heidi (2013): 'External arguments and the Mirror Principle: On the distinctness of Voice and v', *Lingua* **125**, 34–57.

Harley, Heidi and Rolf Noyer (1999): 'Distributed Morphology. State-of-the-Article', *Glot International* **4**, 3–9.

Hartman, Jeremy (2012): '(Non-)Intervention in A-Movement: Some Cross-Constructional and Cross-Linguistic Considerations', *Linguistic Variation* **11**, 121–148.

Heck, Fabian and Gereon Müller (2000): Successive Cyclicity, Long-Distance Superiority, and Local Optimization. *In:* R. Billerey and B. D. Lillehaugen, eds, *Proceedings of WC-CFL.* Vol. 19, Cascadilla Press, Somerville, MA, pp. 218–231.

Heck, Fabian and Gereon Müller (2003): 'Derivational Optimization of Wh-Movement', *Linguistic Analysis* **33**, 97–148. (Volume appeared 2007).

Heck, Fabian and Gereon Müller (2007): Extremely Local Optimization. *In:* E. Brainbridge and B. Agbayani, eds, *Proceedings of the 26th WECOL.* California State University, Fresno, pp. 170–183.

Heck, Fabian and Gereon Müller (2013): On Accelerating and Decelerating Movement: From Minimalist Preference Principles to Harmonic Serialism. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar.* Institut für Linguistik, Universität Leipzig, Leipzig, pp. 511–558. Vol. 90 of *Linguistische Arbeitsberichte.*

Heck, Fabian and Malte Zimmermann (2004): Argumente für die DP als Phase. Ms, Uni-

versität Tübingen, Humboldt-Universität Berlin.

Heim, Irene (1982): The semantics of definite and indefinite noun phrases. PhD thesis, University of Massachusetts, Amherst, Mass.

Hein, Johannes (2008): Verbflexion im Warembori: Eine Analyse im Rahmen der Distribuierten Morphologie. *In:* F. Heck, G. Müller and J. Trommer, eds, *Varieties of Competition. Vol. 87 of Linguistische Arbeits Berichte*. Institut für Linguistik, Universität Leipzig, pp. 49–63.

Henderson, Brent (2009): 'Anti-Agreement and [person] in Bantu', *Selected Proceedings of the 38th Annual Conference on African Linguistics* **38**, 173–181.

Henderson, Brent (2013): 'Agreement and person in anti-agreement', *Natural Language and Linguistic Theory* **31**, 453–481.

Henry, Alison (1995): *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford Studies in Comparative Syntax, Oxford University Press, New York.

Heusinger, Klaus von and Jaklin Kornfilt (2005): 'The case of the direct object in Turkish: Semantics, syntax and morphology', *Turkic Languages* **9**, 3–44.

Hiraiwa, Ken (2001): Multiple Agree and the Defective Intervention Constraint in Japanese. *In:* O. Matushansky, A. Costa, J. Martin-Gonzalez, L. Nathan and A. Szczegielniak, eds, *Proceedings of the HUMIT 2000. Vol. 40 of MITWPL*. MIT Press, Cambridge, Mass., pp. 67–80.

Hiraiwa, Ken (2005): The Morphosyntax of the EPP and C in Buli. *In:* L. Bateman and C. Ussery, eds, *Proceedings of the NELS 35*. GLSA, Amherst, Mass, pp. 267–278.

Höhle, Tilman N. (2000): The w … w construction: Appositive or scope indicating?. *In:* U. Lutz, G. Müller and A. von Stechow, eds, *Wh-Scope Marking*. John Benjamins, Amsterdam, pp. 249–270.

Holmberg, Anders (2000): 'Scandinavian stylistic fronting: how any category can become an expletive', *Linguistic Inquiry* **31**, 445–484.

Holmberg, Anders and Thorbjörg Hróarsdóttir (2003): 'Agreement and movement in Icelandic raising constructions', *Lingua* **113**, 997–1019.

Hornstein, Norbert (2001): *Move. A Minimalist Theory of Construal*. Blackwell, Oxford.

Horvath, Julia (1997): 'The Status of Wh-Expletives and the Partial Wh-Movement Construction of Hungarian', *Natural Language and Linguistic Theory* **15**, 509–572.

Huang, Cheng-Teh James (1982): Logical Relations in Chinese and the Theory of Grammar. PhD thesis, MIT, Cambridge, Mass.

Huybregts, Riny (2009): The Minimalist Program: Not a Bad Idea. Ms., Universiteit Utrecht.

Idsardi, William J. (2000): 'Clarifying opacity', *The Linguistic Review* **17**, 337–350.

Jacobs, Roderick A. and Peter S. Rosenbaum (1968): *English transformational grammar*. Ginn/Blaisdell, Waltham, Mass.

Johnson, C. D. (1971): 'Unbounded expressions in rules of stress and accent', *Glossa* **4**, 185–196.

Johnson, Kyle (2009): Why Movement. Ms., University of Massachusetts.

Jónsson, Jóhannes Gísli (1996): Clausal Architecture and Case in Icelandic. PhD thesis, University of Massachusetts, Amherst.

Kalin, Laura and Coppe van Urk (2012): Aspect-based agreement reversal in Neo-Aramaic. Handout of a talk at WCCFL 30.

Kayne, Richard (1994): *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.

Kayne, Richard S. (1972): Subject Inversion in French Interrogatives. *In:* J. Casagrande and B. Saciuk, eds, *Generative Studies in Romance Languages*. Newbury House, Rowley, Mass., pp. 70–126.

Kayne, Richard S. (1975): *French Syntax: The Transformational Cycle*. MIT Press, Cambridge, Mass.

Kayne, Richard S. (1989): Facets of Romance past participle agreement. *In:* P. Benincà, ed., *Dialect variation and the theory of grammar*. Foris, Dordrecht, pp. 85–103.

Kayne, Richard S. and Jean Yves Pollock (1978): 'Stylistic Inversion, Successive Cyclicity, and Move NP in French', *Linguistic Inquiry* **9**, 595–621.

Keine, Stefan (2010): *Case and Agreement from Fringe to Core: A Minimalist Approach*. Linguistische Arbeiten 536, de Gruyter, Berlin.

Kenstowicz, Michael and Charles Kisseberth (1977): *Topics in phonological theory*. Academic Press, New York.

Kenstowicz, Michael and Charles Kisseberth (1979): *Generative Phonology*. Academic Press, New York.

Keyser, Samuel J. and Paul Postal (1976): *Beginning English grammar*. Harper & Row, New York.

Kinyalolo, Kasangani K.W. (1991): Syntactic dependencies and the Spec-head agreement hypothesis in Kilega. PhD thesis, UCLA, Los Angeles.

Kiparsky, Paul (1965): Phonological change. PhD thesis, MIT, Cambridge, Mass.

Kiparsky, Paul (1968): Linguistic universals and linguistic change. *In:* E. Bach and R. Harms, eds, *Universals in Linguistic Theory*. Holt, New York, pp. 170–210.

Kiparsky, Paul (1971): Historical linguistics. *In:* W. Dingwall, ed., *A Survey of Linguistic Science*. University of Maryland Linguistics Program, College Park, pp. 576–642.

Kiparsky, Paul (1973*a*): Abstractness, Opacity, and Global Rules. *In:* O. Fujimura, ed., *Three Dimensions in Linguistic Theory*. TEC, Tokyo, pp. 57–86.

Kiparsky, Paul (1973*b*): "Elsewhere" in Phonology. *In:* S. Anderson and P. Kiparsky, eds, *A Festschrift for Morris Halle*. Holt, Rinehart and Winston, pp. 93–106.

Kiparsky, Paul (2000): 'Opacity and Cyclicity', *The Linguistic Review* **17**, 351–367.

Kiss, Katalin É. (2002): *The Syntax of Hungarian*. Cambridge University Press, Cambridge.

Kisseberth, Charles W. (1972a): 'Cyclical Rules in Klamath Phonology', *Linguistic Inquiry* **3**, 3–33.

Kisseberth, Charles W. (1972b): 'An Argument against the Principle of Simultaneous Application of Phonological Rules', *Linguistic Inquiry* **3**, 393–396.

Kitahara, Hisatsugu (1997): *Elementary Operations and Optimal Derivations*. MIT Press, Cambridge, Mass.

Klein, Timo (2013): Agree > Move = Resumption. Ms., Universität Leipzig.

Koopman, Hilda (1992): 'On the Absence of Case Chains in Bambara', *Natural Language and Linguistic Theory* **10**, 555–594.

Koopman, Hilda (2006): Agreement Configurations. *In:* C. Boeckx, ed., *Agreement Systems.* Benjamins, Amsterdam, pp. 159–201.

Koutsoudas, Andreas (1972): 'The Strict Order Fallacy', *Language* **48**, 88–96.

Koutsoudas, Andreas (1973): Unordered rule hypotheses. Duplicated, IULC.

Koutsoudas, Andreas, Gerald Sanders and Craig Noll (1974): 'The application of phonological rules', *Language* **50**, 1–28.

Kramer, Ruth (2010): 'The Amharic Definite Marker and the Syntax-Morphology Interface', *Syntax* **13**, 196–240.

Kratzer, Angelika (1996): Severing the external argument from its verb. *In:* J. Rooryck and L. Zaring, eds, *Phrase Structure and the Lexicon.* Springer, Dordrecht, pp. 109–137.

Lahne, Antje (2008*a*): Excluding SVO in Ergative Languages: A New View on Mahajan's Generalisation. *In:* F. Heck, G. Müller and J. Trommer, eds, *Varieties of Competition.* Universität Leipzig, Institut für Linguistik, pp. 65–80. Vol. 87 of *Linguistische Arbeitsberichte.*

Lahne, Antje (2008*b*): Where There is Fire There is Smoke. Local Modelling of Successive-Cyclic Movement. PhD thesis, Universität Leipzig, Leipzig.

Lahne, Antje (2012): Specificity-driven Syntactic Derivation. *In:* M. Uribe-Etxebarria and V. Valmala, eds, *Ways of Structure Building.* Oxford University Press, Oxford, pp. 271–296.

Laka, Itziar (1993): Unergatives that Assign Ergative, Unaccusatives that Assign Accusative. *In:* J. Bobaljik and C. Phillips, eds, *Papers on Case and Agreement I.* MITWPL, Cambridge, Mass, pp. 149–172.

Lakoff, George (1968): *Deep surface grammar.* Indiana University Linguistics Club, Bloomington, Ind.

Lakoff, George (1970): 'Global rules', *Language* **46**, 627–638.

Lakoff, George (1972): 'The global nature of the Nuclear Stress Rule', *Language* **48**, 285–303.

Lakoff, George and Charles W. Kisseberth (1972): 'Hate Rule Ordering Orgy', *CLS* **8**.

Lasnik, Howard (1999*a*): *Minimalist Analysis.* Blackwell, Oxford.

Lasnik, Howard (1999*b*): 'On Feature Strength: Three Minimalist Approaches to Overt Movement', *Linguistic Inquiry* **30**, 197–217.

Lasnik, Howard (2001): Derivation and Representation in Modern Transformational Syntax. *In:* M. Baltin and C. Collins, eds, *The Handbook of Contemporary Syntactic Theory.* Blackwell, Malden, Mass., pp. 62–88.

Lasnik, Howard and Mamoru Saito (1984): 'On the Nature of Proper Government', *Linguistic Inquiry* **15**, 235–289.

Lebeaux, David (1988): Language Acquisition and the Form of the Grammar. PhD thesis, University of Massachussetts, Amherst.

Lebeaux, David (1990): Relative clauses, licensing, and the nature of derivation. *In:* J. Carter, ed., *Proceedings of NELS 20.* GLSA, Amherst, Mass, pp. 318–332.

Lechner, Winfried (2010): Diagnosing covert movement with the Duke of York and reconstruction. Ms., University of Athens. To appear in L. Cheng and N. Corver (eds.) *Diagnosing Syntax*, Oxford University Press.

Legate, Julie Ann (2002): Phases in "Beyond Explanatory Adequacy". Ms., MIT, Cambridge.

Legate, Julie Ann (2003): 'Some Interface Properties of the Phase', *Linguistic Inquiry* **34**, 506–516.

Lightfoot, David (1976): 'Trace Theory and Twice-Moved NPs', *Linguistic Inquiry* **7**, 559–582.

Lumsden, John (1992): 'Underspecification in Grammatical and Natural Gender', *Linguistic Inquiry* **23**, 469–486.

Macdonald, Roderick Ross and Soenjono Dardjowidjojo (1967): *A student's reference grammar of modern formal Indonesian.* Georgetown University Press, Washington, D.C.

Maling, Joan (1980): 'Inversion in embedded clauses in Modern Icelandic', *Íslenskt mál og almenn málfræði* **2**, 175–193.

Manzini, Rita (1994): 'Locality, Minimalism, and Parasitic Gaps', *Linguistic Inquiry* **25**, 481–508.

Marantz, Alec (1984): *On the Nature of Grammatical Relations.* MIT Press, Cambridge, Mass.

Marantz, Alec (1991): Case and Licensing. *In:* G. Westphal, B. Ao and H.-R. Chae, eds, *Proceedings of the Eighth Eastern States Conference on Linguistics.* University of Maryland, pp. 234–253.

Marantz, Alec (1997): No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *In:* A. Dimitriadis, ed., *University of Pennsylvania Working Papers in Linguistics 4.2,.* University of Pennsylvania, pp. 201–225.

Martinović, Martina (2013): Wh-Morphology and Cyclicity in Wolof. Ms., University of Chicago, available at LingBuzz: lingbuzz/001874.

Maslova, Elena (2003): *A Grammar of Kolyma Yukaghir.* Mouton Grammar Library, de Gruyter, Berlin, New York.

Matushansky, Ora (2006): 'Head Movement in Linguistic Theory', *Linguistic Inquiry* **37**(1), 69–109.

May, Robert (1978): The Grammar of Quantification. PhD thesis, MIT, Cambridge, Mass.

McCarthy, John J. (1999): 'Sympathy and Phonological Opacity', *Phonology* **16**, 331–399.

McCarthy, John J. (2007): *Hidden Generalizations: Phonological Opacity in Optimality Theory.* Equinox, London.

McCarthy, John J. (2010): 'An Introduction to Harmonic Serialism', *Language and Linguistics Compass* **4**, 1001–1018.

McCawley, James D. (1982): 'Parentheticals and Discontinuous Constituent Structure', *Linguistic Inquiry* **13**, 91–106.

McCawley, James D. (1984): Exploitation of the cyclic principle as a research strategy in syntax. *In:* W. de Geest and Y. Putseys, eds, *Sentential complementation.* Foris, Dordrecht, pp. 165–183.

McCawley, James D. (1988): *The Syntactic Phenomena of English.* Vol. 1, University of Chicago Press, Chicago.

McCloskey, James (1979): *Transformational Syntax and Model Theoretic Semantics. A Case Study in Modern Irish.* Reidel, Dordrecht.

McCloskey, James (1990): Resumptive pronouns, Ā-binding and levels of representation in Irish. *In:* R. Hendrick, ed., *Syntax of the Modern Celtic Languages.* Academic Press, New York and San Diego, pp. 199–248. Vol. 23 of *Syntax and Semantics.*

McCloskey, James (2000): 'Quantifier float and wh-movement in an Irish English', *Linguistic Inquiry* **31**, 57–84.

McCloskey, James (2001): 'The morphosyntax of Wh-extraction in Irish', *Journal of Linguistics* **37**, 67–100.

McCloskey, James (2002): Resumption, successive cyclicity, and the locality of operations. *In:* S. Epstein and T. Seely, eds, *Derivation and Explanation in the Minimalist Program.* Blackwell, Oxford, pp. 184–226.

McDaniel, Dana (1989): 'Partial and Multiple Wh-Movement', *Natural Language and Linguistic Theory* **7**, 565–604.

McFadden, Thomas (2004): The Position of Morphological Case in the Derivation: A Study in the Syntax-Morphology Interface. PhD thesis, University of Pennsylvania.

McGinnis, Martha (2001): Variation in the phase structure of applicatives. *In:* P. Pica and J. Rooryck, eds, *Linguistic Variation Yearbook 1.* John Benjamins, Amsterdam, pp. 101–142.

McGinnis, Martha J. (1998): Locality in A-Movement. PhD thesis, MIT, Cambridge, Mass.

McLaughlin, Fiona (1997): 'Noun classification in Wolof: when affixes are not renewed', *Studies in African Linguistics* **26**(1), 1–28.

Merchant, Jason (2006): Polyvalent case, geometric hierarchies, and split ergativity. *In:* J. Bunting, S. Desai, R. Peachey, C. Straughn and Z. Tomkova, eds, *Proceedings of the 42nd annual meeting of the Chicago Linguistics Society.* Chicago Linguistics Society, Chicago, pp. 47–67.

Merchant, Jason (2008): 'An asymmetry in voice mismatches in VP-ellipsis and pseudogapping', *Linguistic Inquiry* **39**, 169–179.

Miyagawa, Shigeru (2010): *Why Agree? Why Move? Unifying Agreement-based and Discourse Configurational Languages.* Linguistic Inquiry Monograph 54, MIT Press.

Moravcsik, Edith A. (1995): Summing up Suffixaufnahme. *In:* F. Plank, ed., *Double Case: Agreement by Suffixaufnahme.* Oxford University Press, New York, pp. 451–484.

Morin, Yves and Joyce Friedman (1971): *Phonological grammar tester. Underlying theory.* Vol. 10 of *Natural Language Studies*, University of Michigan Phonetics Laboratory, Ann Arbor.

Müller, Gereon (1995): *A-bar syntax: A study in movement types.* Vol. 42 of *Studies in Generative Grammar*, Mouton de Gruyter, Berlin.

Müller, Gereon (1998): *Incomplete Category Fronting.* Kluwer, Dordrecht.

Müller, Gereon (1999): 'Imperfect Checking', *The Linguistic Review* **16**, 359–404.

Müller, Gereon (2002): Free Word Order, Morphological Case, and Sympathy Theory. *In:* G. Fanselow and C. Féry, eds, *Resolving Conflicts in Grammars: Optimality Theory in Syntax, Morphology, and Phonology.* Buske, Hamburg, pp. 9–48. Special issue of *Linguistische Berichte.*

Müller, Gereon (2004): Phrase Impenetrability and Wh-Intervention. *In:* A. Stepanov, G. Fanselow and R. Vogel, eds, *Minimality Effects in Syntax.* Mouton de Gruyter, Berlin,

pp. 289–325.

Müller, Gereon (2009): Ergativity, Accusativity, and the Order of Merge and Agree. *In:* K. Grohmann, ed., *Explorations of Phase Theory. Features and Arguments.* Mouton de Gruyter, Berlin, pp. 269–308.

Müller, Gereon (2010): 'On Deriving CED Effects from the PIC', *Linguistic Inquiry* **41**, 35–82.

Müller, Gereon (2011): *Constraints on Displacement: A phase-based approach.* John Benjamins, Amsterdam.

Müller, Gereon (2013*a*): A Local Approach to the Williams Cycle. Ms., Universität Leipzig, to appear in *Lingua.*

Müller, Gereon (2013*b*): On the Order of Syntactic Operations: Brief History of Rule Interaction in Syntax. Handout of a compact course at the University of Cambridge. Available at http://www.uni-leipzig.de/~muellerg/.

Müller, Gereon (2014): Remnant Movement in Local Derivational Grammar. Ms., Universität Leipzig, available at LingBuzz: lingbuzz/002022.

Murasugi, Kumiko (1992): Crossing and Nested Paths. PhD thesis, MIT, Cambridge, Mass.

Muriungi, Peter (2003): Wh-Questions in Kiitharaka. Master's thesis, University of Witwatersrand, Johannesburg.

Muriungi, Peter (2005): 'Wh-Questions in Kitharaka', *Studies in African Linguistics* **34**, 43–104.

Nevins, Andrew (2003): Do Person/Number Syncretisms Refer to Negative Values?. Ms., Harvard University.

Nevins, Andrew (2004): Derivations without the Activity Condition. *In:* M. McGinnis and N. Richards, eds, *Perspectives on Phases.* MITWPL, Cambridge, Mass., pp. 287–310. MIT Working Papers in Linguistics 49.

Nevins, Andrew (2007): 'The representation of third person and its consequences for the person-case constraint', *Natural Language and Linguistic Theory* **25**, 273–313.

Nevins, Andrew and Pranav Anand (2003): Some AGREEment Matters. *In:* G. Gardina and M. Tsujimura, eds, *Proceedings of WCCFL 22.* Cascadilla Press, Sommerville, MA, pp. 101–114.

Newton, Brian E. (1971): 'Ordering paradoxes in phonology', *Journal of Linguistics* **7**, 31–53.

Nikolaeva, Irina (2002): The Hungarian external possessor in a European perspective. *In:* C. Hasselblatt and R. Blokland, eds, *Finno-Ugrians and Indo-Europeans: Linguistic and literary contacts.* Shaker, Maastricht, pp. 272–285.

Noonan, Máire (1992): Case and Syntactic Geometry. PhD thesis, McGill University.

Noonan, Máire (2002): CP-Pied-Piping and Remnant IP Movement in Long Distance Wh-Movement. *In:* A. Alexiadou, E. Anagnostopoulou, S. Barbiers and H.-M. Gärtner, eds, *Dimensions of Movement: From features to remnants.* Benjamins, Amsterdam, p. 269ï£¡295. Linguistik aktuell (Linguistics today) Vol. 48.

Noyer, Rolf (1992): Features, Positions, and Affixes in Autonomous Morphological Structure. PhD thesis, MIT, Cambridge, Mass.

Noyer, Rolf (1998): Impoverishment Theory and Morphosyntactic Markedness. *In:*

S. Lapointe, D. Brentari and P. Farrell, eds, *Morphology and its Relation to Phonology and Syntax*. CSLI, Palo Alto, pp. 264–285.

Nunes, Jairo (2004): *Linearization of chains and sideward movement*. Linguistic Inquiry Monographs 43, MIT Press, Cambridge, Mass.

Obenauer, Hans-Georg (1994): Aspects de la syntaxe A-barre. PhD thesis, Université de Paris VIII, Paris.

Odden, Dave (2000): 'Opacity and ordering: H-deletion in Kikerewe', *The Linguistic Review* **17**, 323–336.

Ormazabal, Javier, Juan Uriagereka and Myriam Uribe-Etxebarria (1994): Word Order and Wh-Movement: Towards a Parametric Account. Ms., University of Connecticut, University of Maryland.

Ott, Dennis (2009): 'Stylistic fronting as remnant movement', *Working Papers in Scandinavian Syntax* **83**, 141–178.

Ouali, Hamid (2008): On C-to-T-Feature Transfer: The Nature of Agreement and Anti-agreement in Berber. *In:* R. D'Alessandro, S. Fischer and G. Hrafnbjargarson, eds, *Agreement Restrictions*. de Gruyter, pp. 159–180.

Ouali, Hamid (2011): *Agreement, Pronominal Clitics and Negation in Tamazight Berber: A Unified Analysis*. Continuum International Publishing, London.

Ouali, Hamid and Acrisio Pires (2005): Complex tenses, Agreement and Wh-Extraction. *In:* R. Cover and Y. Kim, eds, *Proceedings of the 31st Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA, pp. 253–264.

Ouhalla, Jamal (1993): 'Subject-Extraction, Negation and the Antiagreement Effect', *Natural Language and Linguistic Theory* **11**, 477–518.

Ouhalla, Jamal (2005): 'Agreement features, agreement, and anti-agreement', *Natural Language and Linguistic Theory* **23**, 655–686.

Perlmutter, David (1971): *Deep and surface structure constraints in syntax*. Holt Rinehart & Winston, New York.

Perlmutter, David (1978): Impersonal passives and the unaccusative hypothesis. *In:* J. Jaegu and A. Woodbury, eds, *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*. BLS, Berkeley, CA, p. 157Ű189.

Perlmutter, David and Scott Soames (1979): *Syntactic Argumentation and the Structure of English*. The University of California Press, Berkeley.

Pesetsky, David (1989): 'Language-particular Processes and the Earliness Principle', Ms., MIT.

Pesetsky, David (1998): Some Optimality Principles of Sentence Pronunciation. *In:* P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky, eds, *Is the Best Good Enough?*. MIT Press, Cambridge, Mass., pp. 337–383.

Pesetsky, David and Esther Torrego (2001): T-to-C Movement: Causes and Consequences. *In:* M. Kenstowicz, ed., *Ken Hale: A Life in Language*. MIT Press, Cambridge, Mass., pp. 355–426.

Phillips, Colin (2001): Disagreement between Adults and Children. *In:* A. Mendikoetxea and M. Uribe-Etxebarria, eds, *Theoretical Issues on the Morphology-Syntax Interface*. ASJU, San Sebastian, pp. 359–394.

Platzack, Christer (1987): 'The Case of Narrative Inversion in Swedish and Icelandic', *Working Papers in Scandinavian Syntax* **31**, 9–14.

Pollard, Carl J. and Ivan A. Sag (1994): *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Postal, Paul (1964): 'Underlying and superficial linguistic structure', *Harvard Educational Review* **34**, 246–266.

Postal, Paul (1971): *Cross-Over phenomena*. Holt, Rinehart, and Winston, New York.

Postal, Paul (1972): 'On Some Rules That Are Not Successive Cyclic', *Linguistic Inquiry* **3**, 211–222.

Preminger, Omer (2011*a*): Agreement and case: Patterns, interactions, and implications. Handout of a talk presented at the LUCL Syntax Lab, Leiden.

Preminger, Omer (2011*b*): Agreement as a Fallible Operation. PhD thesis, MIT, Cambridge, Mass.

Preminger, Omer (2011*c*): The nature of syntactic computation: evidence from agreement. Handout of a talk presented at the Polinsky Language Lab, Harvard University, Cambridge, Mass.

Preminger, Omer (2013): 'That's not how you agree: A reply to Zeijlstra', *The Linguistic Review* **30**, 491–500.

Prince, Alan and Paul Smolensky (2004): *Optimality Theory. Constraint Interaction in Generative Grammar*. Blackwell, Oxford. (Original version from 1993).

Pullum, Geoffrey K. (1979): *Rule Interaction and the Organization of a Grammar*. Garland, New York & London. (PhD thesis from 1976, University of London).

Pullum, Geoffrey K. (1992): The Origins of the Cyclic Principle. *In:* J. Marshall Denton, G. Chan and C. Canakis, eds, *Papers from the 28th Regional Meeting of the Chicago Linguistic Society. Vol. 2: The Cycle in Linguistic Theory*. Chicago Linguistic Society, pp. 209–236.

Pylkkänen, Liina (2002): Introducing Arguments. PhD thesis, MIT, Cambridge, Mass.

Rackowski, Andrea (2005): The structure of Tagalog: Specificity, voice, and the distribution of arguments. PhD thesis, MIT, Cambridge, Mass.

Rackowski, Andrea and Norvin Richards (2005): 'Phase Edge and Extraction: A Tagalog Case Study', *Linguistic Inquiry* **36**, 565–599.

Reintges, Chris H. and Melanie Green (2004): 'Coptic second tenses and Hausa relative aspects: A comparative view', *Lingua Aegyptia – Journal of Egyptian Language Studies* **12**, 157–177.

Reintges, Chris H., Philip LeSourd and Sandra Chung (2006): Movement, Wh-Agreement, and Apparent Wh-in-Situ. *In:* L. Cheng and N. Corver, eds, *Wh-Movement: Moving On*. MIT Press, Cambridge, Mass., chapter 7, pp. 165–194.

Reuland, Eric (2001): 'Primitives of Binding', *Linguistic Inquiry* **32**, 439–492.

Reuland, Eric (2011): *Anaphora and Language Design*. MIT Press, Cambridge, Mass.

Řezáč, Milan (2003): 'The Fine Structure of Cyclic Agree', *Syntax* **6**, 156–182.

Řezáč, Milan (2004): Elements of Cyclic Agree: Agree and Merge. PhD thesis, University of Toronto.

Řezáč, Milan (2008): Phi-Agree and theta-related Case. *In:* D. Harbour, D. Adger and S. Béjar, eds, *Phi theory*. Oxford University Press, Oxford, pp. 83–129.

Richards, Marc D. (2004): Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax. PhD thesis, University of Cambridge, Cambridge, UK.

Richards, Marc D. (2007): 'On feature-inheritance: an argument from the Phase Impenetrability Condition', *Linguistic Inquiry* **38**, 563–572.

Richards, Marc D. (2009): 'Internal Pair-Merge: the Missing Mode of Movement', *Catalan Journal of Linguistics* **8**, 55–73.

Richards, Marc D. (2011): 'Deriving the Edge: What's in a Phase?', *Syntax* **14**, 74–95.

Richards, Norvin (1997): What Moves Where When in Which Language. PhD thesis, MIT, Cambridge, Mass.

Richards, Norvin (2013): 'Lardil "Case Stacking" and the Timing of Case Assignment', *Syntax* **16**, 42–76.

Ringen, Catherine (1972): 'On Arguments for Rule Ordering', *Foundations of Language* **8**, 266–273.

Rivero, María-Luisa (1978): 'Topicalization and Wh Movement in Spanish', *Linguistic Inquiry* **9**, 513–517.

Rivero, María-Luisa (1980): 'On Left-Dislocation and Topicalization in Spanish', *Linguistic Inquiry* **11**, 363–393.

Rizzi, Luigi (1986*a*): 'Null Objects in Italian and the Theory of 'pro'', *Linguistic Inquiry* **17**, 501–557.

Rizzi, Luigi (1986*b*): On chain formation. *In:* H. Borer, ed., *Syntax and Semantics 19: The Syntax of Pronominal Clitics*. Academic Press, Orlando, pp. 65–96.

Rizzi, Luigi (1990): *Relativized Minimality*. MIT Press, Cambridge, Mass. Linguistic Inquiry Monographs 16.

Rizzi, Luigi (1997): The Fine Structure of the Left Periphery. *In:* L. Haegeman, ed., *Elements of Grammar*. Kluwer, Dordrecht, pp. 281–337.

Rizzi, Luigi (2004): On the Form of Chains: Criterial Positions and ECP Effects. Ms., University of Siena.

Rizzi, Luigi (2007): 'On Some Properties of Criterial Freezing', *CISCL Working Papers on Language and Cognition. STIL* **1**.

Rizzi, Luigi and Ian Roberts (1989): 'Complex Inversion in French', *Probus: International Journal of Latin and Romance Linguistics* **1**, 1–30.

Rizzi, Luigi and Ur Shlonsky (2004): Strategies for subject extraction. *In:* U. Sauerland and H.-M. Gärtner, eds, *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*. Mouton de Gruyter, Berlin, pp. 115–160.

Roberts, Ian (2010): *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. MIT Press, Cambridge, Mass.

Rocquet, Amélie (2013): Splitting Objects: A Nanosyntactic account of direct object marking. PhD thesis, Universiteit Gent, Gent.

Rögnvaldsson, Eiríkur and Höskuldur Thráinsson (1990): Icelandic Word Order Once More. *In:* J. Maling and A. Zaenen, eds, *Syntax and Semantics Vol. 24*. Academic Press,

San Diego, pp. 3–40.

Rooryck, Johan and Guido Van den Wyngaerd (2011): *Dissolving Binding Theory*. Oxford University Press, Oxford.

Ross, John R. (1967): Constraints on Variables in Syntax. PhD thesis, MIT, Cambridge, Mass.

Ross, John R. (1969): 'Guess who?', *CLA* **5**, 252–286.

Sabel, Joachim (2000): Partial Wh-Movement and The Typology of Wh-Questions. *In:* U. Lutz, G. Müller and A. von Stechow, eds, *Wh-Scope Marking*. John Benjamins, Amsterdam, pp. 409–446.

Sabel, Joachim and Jochen Zeller (2006): Wh-Question Formation in Nguni. *In:* J. Mugane, J. Hutchinson and D. Worman, eds, *African Languages and Linguistics in Broad Perspectives*. Cascadilla Press, Sommerville, Mass., pp. 271–283.

Saddy, Douglas (1991): Wh-Scope Mechanisms in Bahasa Indonesia. *In:* L. Cheng and H. Demirdache, eds, *More papers on wh-movement (MITWPL 15)*. MIT Press, Cambridge, Mass., pp. 183–218.

Safir, Ken (1999): 'Vehicle Change and Reconstruction in A-Chains', *Lingustic Inquiry* **30**, 587–620.

Salzmann, Martin (2006): Resumptive Prolepsis. A Study in indirect A′-dependencies. PhD thesis, LOT, Utrecht.

Salzmann, Martin (2011): Silent resumptives in Zurich German possessor relativization. *In:* P. Gallmann and M. Wratil, eds, *Null Pronouns*. Mouton, Berlin, pp. 141–221.

Salzmann, Martin (2012): Deriving reconstruction asymmetries in ATB-movement by means of asymmetric extraction + ellipsis. *In:* P. Ackema, R. Alcorn, C. Heycock, D. Jaspers, J. van Craenenbroeck and G. Vanden Wyngaerd, eds, *Comparative Germanic Syntax: The State of the Art*. John Benjamins, Amsterdam, pp. 353–385.

Salzmann, Martin (2013): Rule ordering in verb cluster formation. On the extraposition paradox and the placement of the infinitival particle *te/zu*. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar*. Universität Leipzig, pp. 65–121. Vol. 90 of *Linguistische Arbeitsberichte*.

Sanders, Gerald (1974): 'Precedence Relations in Language', *Foundations of Language* **11**.

Sato, Yosuke (2008): Minimalist interfaces: Selected issues in Indonesian and Javanese. PhD thesis, University of Arizona, Tucson, AZ.

Sato, Yosuke (2012): 'Syntax-Morphology Interface: Evidence from Standard Indonesian and Kendal Javanese', *Studia Linguistica* **66**(1), 32–57.

Sauerland, Uli (2003): 'Intermediate Adjunction with A-Movement', *Linguistic Inquiry* **34**, 308–314.

Sauerland, Uli and Paul Elbourne (2002): 'Total Reconstruction, PF Movement, and Derivational Order', *Linguistic Inquiry* **33**, 283–319.

Schäfer, Florian (2012): Local Case, Cyclic Agree and the Syntax of Truly Ergative Verbs. *In:* A. Alexiadou, T. Kiss and G. Müller, eds, *Local Modelling of Non-Local Dependencies in Syntax*. de Gruyter, Berlin, pp. 273–304. Linguistische Arbeiten 547.

Schneider-Zioga, Patricia (1995): 'Specifier/Head Agreement in Kinande', *Cahiers Linguistiques d'Ottawa* **23**, 67–96.

Schneider-Zioga, Patricia (2000): 'Anti-agreement and the fine structure of the left edge', *University of California Irvine Working Papers in Linguistics* **6**.

Schneider-Zioga, Patricia (2007): 'Anti-agreement, anti-locality and minimality. The syntax of dislocated subjects', *Natural Language and Linguistic Theory* **25**, 403–446.

Schneider-Zioga, Patricia (2009): Wh-agreement and bounded unbounded movement. *In:* J. Brucart, A. Gavarró and J. Solà, eds, *Merging Features: Computation, Interpretation, and Acquisition.* Oxford University Press, pp. 46–59.

Schütze, Carson (1997): INFL in Child and Adult Language: Agreement, Case and Licensing. PhD thesis, MIT, Cambridge, Mass.

Sells, Peter (1984): Syntax and Semantics of Resumptive Pronouns. PhD thesis, University of Massachusetts, Amherst.

Shamir, E. (1961): 'On sequential grammars', *Technical Report* **7**. O.N.R Information Systems Branch.

Siewierska, Anna (2004): *Person.* Cambridge Textbooks in Linguistics, Cambridge University Press, Cambridge, UK.

Sigurðsson, Halldór (1989): Verbal Syntax and Case in Icelandic. PhD thesis, University of Lund.

Sigurðsson, Halldór Ámann and Anders Holmberg (2008): Icelandic Dative Intervention: Person and Number are Separate Probes. *In:* R. d'Alessandro, G. H. Hrafnbjargarson and S. Fischer, eds, *Agreement Restrictions.* Mouton de Gruyter, Berlin, pp. 251–280.

Silverstein, Michael (1976): Hierarchy of Features and Ergativity. *In:* R. Dixon, ed., *Grammatical Categories in Australian Languages.* Australian Institute of Aboriginal Studies, Canberra, pp. 112–171.

Smolensky, Paul (1995): On the Internal Structure of Con, the Constraint Component of UG. Ms., Johns Hopkins University.

Soh, Hooi-Ling and H. Nomoto (2009): Movement across *meN-* and unaccusatives in Malay. Handout of a talk at the 83rd Annual Meeting of the LSA, 2009, San Francisco.

Sportiche, Dominique (1988): 'A theory of floating quantifiers and its corollaries for constituent structure', *Linguistic Inquiry* **19**, 425–449.

Sportiche, Dominique (1989): 'Le Mouvement Syntaxique: Contraintes et Paramètres', *Langages* **95**, 35–80.

Starke, Michal (2001): Move Dissolves Into Merge: A Theory of Locality. PhD thesis, University of Geneva.

Stechow, Arnim von and Wolfgang Sternefeld (1988): *Bausteine syntaktischen Wissens.* Westdeutscher Verlag, Opladen.

Sternefeld, Wolfgang (2006): *Syntax: Eine morphologisch motivierte generative Beschreibung des Deutschen.* Stauffenburg, Tübingen. Two volumes.

Stump, Gregory (2001): *Inflectional Morphology.* Cambridge University Press, Cambridge.

Sūner, Margarita (1988): 'The role of agreement in clitic-doubled constructions', *Natural Language and Linguistic Theory* **6**, 391–434.

Svenonius, Peter (2001): On Object Shift, Scrambling, and the PIC. *In:* E. Guerzoni and O. Matushansky, eds, *MIT Working Papers in Linguistics 39.* MITWPL, Cambridge, Mass., pp. 267–289.

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

Szabolcsi, Anna (1994): The Noun Phrase. *In:* F. Kiefer, ed., *Syntax and Semantics, vol. 27: The Structure of Hungarian.* Academic Press, San Diego, pp. 179–274.

Takahashi, Daiko (1994): Minimality of Movement. PhD thesis, University of Connecticut, Storrs.

Takahashi, Shoichi and Martina Gracanin Yuksek (2008): 'Morphosyntax of Movement Dependencies in Haitian Creole', *Syntax* **11**, 223–250.

Thráinsson, Höskuldur (1979): *On Complementation in Icelandic.* Garland Publishing, New York.

Thráinsson, Höskuldur (2007): *The Syntax of Icelandic.* Cambridge University Press, Cambridge, UK.

Torrego, Esther (1984): 'On Inversion in Spanish and Some of Its Effects', *Linguistic Inquiry* **15**, 103–129.

Torrego, Esther (1996): Experiencers and Raising Verbs. *In:* R. Freidin, ed., *Current Issues in Comparative Grammar.* Kluwer, Dordrecht, pp. 101–120.

Torrence, Harold (2012): 'The morpho-syntax of silent wh-expressions in Wolof', *Natural Language and Linguistic Theory* **30**, 1147–1184.

Torrence, Harold (2013*a*): *The Clause Structure of Wolof: Insights into the Left Periphery.* John Benjamins, Amsterdam.

Torrence, Harold (2013*b*): The Morphosyntax of Wolof Clefts: Structure and Movement. *In:* K. Hartmann and T. Veenstra, eds, *Cleft Structures.* John Benjamins, Amsterdam, pp. 187–223.

Travis, Lisa (1984): Parameters and Effects of Word Order Variation. PhD thesis, MIT, Cambridge, Mass.

Trommer, Jochen (2011): Aspects of Western Nilotic Mutation Morphology. Habilitation thesis, Universität Leipzig.

Tuller, Laurice Anne (1986): Bijective Relations in Universal Grammar and the Syntax of Hausa. PhD thesis, UCLA, Los Angeles, CA.

van Koppen, Marjo (2005): One Probe – Two Goals: Aspects of agreement in Dutch dialects. PhD thesis, Leiden University., Leiden. LOT-publications number 105.

van Koppen, Marjo (2007): Agreement with (the internal structure of) copies of movement. *In:* N. Corver and J. Nunes, eds, *The copy-theory of movement.* John Benjamins, Amsterdam, pp. 327–350.

van Koppen, Marjo (2011): The directionality of Agree. Handout of a talk at BCGL6: Agreement. Brussels, December 2011.

van Koppen, Marjo, Michael Diercks and Mike Putnam (2011): The directionality of Agree: Can a universal feature value operation be established?. Handout of a talk at Quo Vadis? Newborn, Reborn, or Stillborn? Potsdam, 2011.

van Oostendorp, Marc (2006): A theory of morphosyntactic colours. Ms., Meertens Institute, Amsterdam.

van Riemsdijk, Henk (1982): 'Correspondence effects and the empty category principle', *Tilburg Papers in Language and Literature* **12**.

van Riemsdijk, Henk and Edwin Williams (1981): 'NP structure', *The Linguistic Review* **1**, 171–217.

van Urk, Coppe (2014): The left periphery in Dinka: Intermediate movement is regular movement. Handout of a talk at WCCFL 32.

van Urk, Coppe (in prep.): How to get a case alternation: Voice in Dinka. Ms., MIT.

van Urk, Coppe and Norvin Richards (2013): Two components of long-distance extraction: Successive cyclicity in Dinka. Ms., MIT, available at LingBuzz: lingbuzz/001766.

Vikner, Sten (1995): *Verb movement and expletive subjects in the Germanic languages*. Oxford University Press, Oxford.

Wasow, Thomas (1972): Anaphoric Relations in English. PhD thesis, MIT, Cambridge, Mass.

Watanabe, Akira (1995): Conceptual Basis of Cyclicity. *In:* R. Pensalfini and H. Ura, eds, *Papers on Minimalist Syntax. Vol. 27 of MIT Working Papers in Linguistics*. MIT Press, Cambridge, Mass., pp. 269–291.

Watanabe, Akira (1996): *Case absorption and WH-agreement*. Kluwer, Dordrecht.

Watanabe, Akira (2012): A morphological solution to agreement puzzles in Slavic. *In:* A. Ralli, G. Booij, S. Scalise and A. Karasimos, eds, *Proceedings of the 7th Mediterranean Morphology Meeting: Morphology and Diachrony*. University of Patras, pp. 112–122.

Weisser, Philipp (2014): Derived Coordination: A Minimalist Perspective on Clause-Chains, Converbs and Asymmetric Coordination. PhD thesis, (in preparation), Universität Leipzig, Leipzig.

Wexler, Kenneth and Peter Culicover (1980): *Formal Principles of Language Acquisition*. MIT Press, Cambridge, Mass.

Wiese, Bernd (1999): 'Unterspezifizierte Paradigmen. Form und Funktion in der pronominalen Deklination', *Linguistik Online* **4**. (www.linguistik-online.de/3 99).

Wilder, Chris (2008): Shared Constituents and Linearization. *In:* K. Johnson, ed., *Topics in Ellipsis*. Cambridge University Press, pp. 229–258.

Williams, Edwin S. (1974): Rule ordering in syntax. PhD thesis, MIT, Cambridge, Mass.

Williams, Edwin S. (1994): 'Remarks on Lexical Knowledge', *Lingua* **92**, 7–34.

Williams, Edwin S. (1997): 'Blocking and Anaphora', *Linguistic Inquiry* **28**, 577–628.

Williams, Edwin S. (2003): *Representation Theory*. MIT Press, Cambridge, Mass.

Wurmbrand, Susanne (2012): Agree(ment): Looking up or looking down?. Lecture given in Agreement Seminar, MIT.

Wurmbrand, Susanne (2013): The Merge Condition: A syntactic approach to selection. Ms., University of Connecticut. To appear in *Minimalism and Beyond: Radicalizing the interfaces*, John Benjamins.

Zaenen, Annie, Joan Maling and Höskuldur Thráinsson (1985): 'Case and Grammatical Functions: The Icelandic Passive', *Natural Language and Linguistic Theory* **3**, 441–483.

Zanuttini, Raffaella (1997): *Negation and Clausal Structure: A Comparative Study of Romance Languages*. Oxford University Pres, New York.

Zeijlstra, Hedde (2012): 'There is only one way to agree', *The Linguistic Review* **29**, 491–539.

Zentz, Jason (2013): Bantu wh-agreement and the case against probe impoverishment.

Handout of a talk at GURT, March 2013.

Zwart, Jan-Wouter (1993): Dutch syntax. A Minimalist approach. PhD thesis, University of Groningen.

Selbstständigkeitserklärung

und Erklärung über frühere Promotionsversuche

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Ich habe nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Die Arbeit wurde zuvor weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Institution vorgelegt. Sie ist weder Bestandteil eines ruhenden Verfahrens noch wurde sie in einem gleichartigen Promotionsverfahren als endgültig nicht bestanden erklärt. Die Arbeit ist vorher auch noch nicht veröffentlicht worden.

Ich erkläre, dass ich bisher noch keine Promotionsversuche unternommen habe.

Leipzig, 28.02.2014

Doreen Georgi