Wh-quantifier float in German

New and old insights into the theory of quantifier float

Aaron Doliana

LING 888 PAPER

Department of Linguistics
University of Maryland, College Park

Defense: May 30, 2019

Committee:
Prof. Norbert Hornstein (supervisor)
Prof. Howard Lasnik (supervisor)
Prof. Jeffrey Lidz

last updated April 3, 2020

Contents

1 Introduction

German *wh*-questions can include an extra quantifier, known as "invariant *alles*" 'all':[1]

(1) [Wen alles]$_1$ hat der Peter t_1 zur Party mitgebracht?
who.ACC all have.3SG the.NOM Peter to.the party brought.with
'Who all did Peter bring along to the party?'

The syntax of the quantifier *alles* raises a series of questions. For one, its distribution comprises both positions right-adjacent to the *wh*-phrase, as in (1), but also positions at some distance from the *wh*-phrase, as in (2), with no change in meaning.

(2) Wen$_1$ hat der Peter t_1 alles zur Party mitgebracht?
who.ACC have.3SG the.NOM Peter all to.the party brought.with
'Who all did Peter bring along to the party?'

This property is complicated by the fact that 'distal' (non-adjacent) *alles* can occur in a variety of positions (but not in multiple ones at the same time, see section 3.2):

(3) (Und) wen (alles) hat (alles) der Peter (alles) gestern (alles)
and who.ACC all have.3SG all the.NOM Peter all yesterday all
zur Party mitgebracht?
to.the party brought.with
'(And) who all did Peter bring along to the party yesterday?'

Two questions arise in this context: *what is the configuration in which* alles *can be introduced into the structure?*, and *how are sentences with adjacent and distal* alles *related to each other?*

In addition, *alles* needs to be licensed. Sentences that are similar to (1)-(2), but are not *wh*-questions, are not acceptable. (4a) is a polar question with a *wh*-indefinite; (4b) is a declarative with a topicalized definite object in Spec,C; (4c) is a declarative with a topicalized indefinite subject of a passive.

(4) a. *Hat der Peter welche alles zur Party mitgebracht?
has the.NOM Peter which.PL.WND all to.the party brought.with
Intended: 'Did Peter bring along all of some (e.g. friends) to the party?'

b. *Die (Freunde) hat der Peter alles zur Party mitgebracht.
those friends has the.NOM Peter all to.the party brought.with
Intended: 'Them/those friends all, Peter brought along to the party.'

c. *Einige (Freunde) wurden alles zur Party eingeladen.
some friends be.PST.PL all to.the party invited

[1] Glosses follow Leipzig glossing conventions. Exceptions are: WND='*wh*-indefinite', WQ='operator floating-quantifier, *alles*', MOD='modal verb', DPRT='discourse particle'.

Intended: 'All of some friends were invited to the party.'

This fact raises the question of *what the configuration is in which* alles *is licensed.*

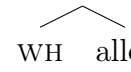
These three questions have been answered in two different ways in the previous literature. Reis (1992), in her seminal paper on invariant *alles*, concludes that *alles* is an $\bar{\text{A}}$-trace clitic, most likely underlyingly right-adjacent to the *wh*-phrase with distal *alles* derived transformationally. *Alles* is then licensed in the position in which it is introduced, in some position forming a constituent with the larger *wh*-phrase. In contrast, Heck and Himmelreich (2017) conclude that *alles* can be one of two things. It can either be an adverbial, introduced in a fixed position of the *v*P-projection, or it can be a *wh*-phrase-internal element. In both cases *alles* is licensed under c-command by a local *wh*-phrase. Adjacent and distal *alles* are then not transformationally related in this analysis.

These two approaches in turn echo the two major families of analysis in the literature on quantifier float in A-movement and $\bar{\text{A}}$-movement dependencies. This paper will not provide, or attempt to provide an exhaustive, or extensive, survey of the vast literature and landscape of analyses (the reader is referred to Bobaljik (2003) and Fitzpatrick (2006) for extensive discussion). Rather a simple question and its consequences are investigated: *is there evidence for a single source for adjacent and distal* alles *in German?* Two competing hypotheses are explored, rejecting one on empirical grounds, and discussing versions of the superior one. In so doing, this paper extends in important ways the empirical understanding of invariant *alles* and the work of Reis (1992) and Zimmermann (2007), and reconnects with results found for West Ulster English by McCloskey (2000); Henry (2012), or Japanese, Korean and Russian in Fitzpatrick (2006).

The two competing hypotheses are the *same-source theory* (SST) and the *different-source theory* (DST).

The SST draws from the *floating approach* (Dougherty, 1970; Kayne, 1975) and the *stranding approach* (Sportiche, 1988; Miyagawa, 1989; Shlonsky, 1991; McCloskey, 2000) to quantifier float, and holds that sentences with an adjacent quantifier and sentences with a floated quantifier have a common source. Specifically, for *alles*, the SST holds that both kinds of *alles* have a stage in the derivation where the *wh*-(containing-)expression and *alles* are sisters, as in (5).[2]

(5) *Same source for adjacent and distal* alles*:*

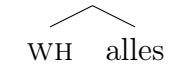


[2] Order is ignored at this stage: (5) is to be considered equivalent to [alles WH] until section 5.

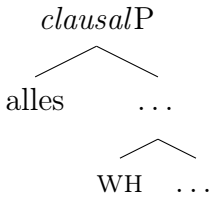
The DST draws from the adverbial approach to quantifier float (Dowty and Brodie, 1984; Bobaljik, 1995; Koopman, 2010; Heck and Himmelreich, 2017). It holds that the floated quantifier has a different source from the quantifier that surfaces in one constituent with its associate, and that the floated quantifier is rather to be analyzed as a clausal/event-level adverbial, e.g. a *v*P-adverb. The floating quantifier is taken to be in a direct or indirect dependency with the associate, yielding (6b) as its characteristic stage in the derivation. ('. . . ' indicates variable content.) The adjacent quantifier has the same source as (5), (6a).

(6) *Different sources for adjacent and distal* alles*:*

a.

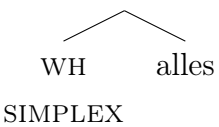


b.



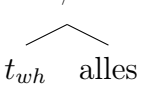
While these two theories are compatible with a variety of analyses, the details won't matter for the first part of the paper. Rather, a range of first-principles predictions is established for the two theories. On the basis of these, the DST will be shown to be incompatible with the full bouquet of empirical generalizations established in section 3, regardless of the details of particular DST-analyses. After rejecting the DST in favor of the SST more broadly, specific analyses in line with the SST are explored, concluding in favor of an analysis where *alles* is a right-branching expression that requires to be the sister of a specific kind of operator at some point in the derivation. In particular, restrictions that apply to complex vs. simplex *wh*-phrase associates suggest that one of the following two representations must hold at some point in derivation, and that therefore narrow syntax, or spell-out/the interfaces, treat simplex *wh*-expressions and traces of *wh*-movement identically:

(7) a.



SIMPLEX

b.



First, the idea that it is narrow syntax that treats them identically is explored. On that route, were there a deep reason for it, it would essentially provide an argument in favor of traces and against the copy theory of movement. The line of inquiry is rejected, turning to the idea that it is spell-out/the interfaces that treat simplex *wh*-expressions (or some effects they can have on the structure) and traces of *wh*-movement identically. Two routes are explored to resolve the disjunction between simplex *wh*-expressions and traces of *wh*-

movement, one an LF condition, the other a spell-out or PF condition. The idea at LF is that *wh*-expressions are themselves variables, and that *alles* must be sister to a variable. The idea at spell-out/PF is that the initial structure formed of [associate+*alles*] is an XP-YP structure and that therefore it either cannot be linearized, or it cannot be labeled. The two options available to reach a linearizable/labelable output create the disjunction: (i) the associate is a simplex *wh* able to project, keeping *alles* adjacent, or (ii) the associate moves, stranding *alles*.

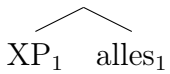
The paper is structured as follows. Section 2.2 outlines some basic properties of *alles*. Section 3 lays out and motivates the empirical generalizations that need to be explained against the backdrop of first-principle predictions made by same-source vs. different-source theories of quantifier float. Consequences of each generalization are discussed along the way. Section 4 recapitulates findings and conclusions concerning the same-source different-source question. Section 5 restricts the analytical space of how the float of *alles* may or may not be created from the shared constituent with the associate. Section 6 concludes, and discusses implications and future directions.

2 Basic facts about *alles*

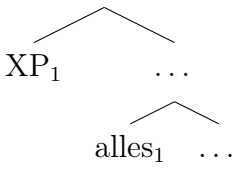
2.1 Note on terminology

Throughout the paper I will make use of the following descriptive terms. I will refer to *alles* and comparable expressions together as *quantifiers*. The XP that the quantifier relates to (or that contains the expression that the quantifier relates to) will be referred to *the/its associate*. The relation will often be indicated through numeric co-indexation for expository purposes. Instances of *alles* occurring in the same constituent with its associate will be referred to as *adjacent alles* (8a). Instances of *alles* occurring somewhere else in the sentence will be referred to as *distal*, or *floated alles*, or more generally as a *distal* or *floated* quantifier (8b).

(8) a.



b.



2.2 Basic facts

The German quantifier *alles* 'all', often referred to as "invariant *alles*", is an expression that affects the discourse properties of a question. *Alles* is morphologically invariant, and is

obligatorily de-stressed, i.e. may not be focused, and it resists surfacing in positions where the intonation cannot be accommodated for this property. Syntactically, *alles* is restricted to associates that are operators, such as *wh*-phrases. Most typically *alles* is found in *wh*-questions, e.g. (9).

(9) a. Wen alles musstest du dafür bestechen?
who.ACC WQ must.PST.2SG you for.that bribe
'Who all did you have to bribe for that?'

b. Wem alles hat Tante Eva ein Geschenk mitgebracht?
who.DAT WQ have.3SG aunt Eva a present brought.with
'Who all did aunt Eva bring a present?'

Alles is also found in *wh*-exclamatives (10a), and, marginally, in restrictive relative clauses (11b) (Reis, 1992: 471).

(10) a. Wen der alles kennt!
who.ACC that.NOM WQ know.3SG
'(All) the people he knows!'

b. ?Das sind die Bücher, die du alles rezensieren wolltest.
that be.3PL the books REL you WQ review wanted
'These are all the (kinds of) books that you wanted to review.'

In *wh*-questions, *alles* can be found in matrix questions (9) as in embedded questions (11a). *Alles* can also relate to echo *wh*-questions, (11b), and to quizmaster-like questions, (11c), (both of which are not true matrix *wh*-questions as they cannot, e.g., license the discourse particle *denn* (on *denn*, see e.g. Bayer et al., 2016)).

(11) a. Ich hab' ja keine Ahnung, wen du alles dafür bestechen
I have DPRT no idea who.ACCSG you WQ for.that bribe
musstest, aber ...
must.PST.2 but
'I mean, I have no idea who all you had to bribe for that, but...'

b. Du hast WEN alles gestern gesehen?
you.NOM have.2SG who.ACC WQ yesterday seen
'You saw WHO all yesterday?'

c. Dossiers über wen alles veröffentlichte WikiLeaks zuletzt vor der
dossiers about who.ACC WQ publish.PST.3SG WikiLeaks last before the
shockierenden Verhaftung Julian Assanges?
shocking arrest Julian Assange.GEN?
'Dossiers about who all has WikiLeaks published last before Julian Assange's shocking arrest?'

More generally, *alles* is restricted to associates that are Ā-operators, and specifically, as Reis notes, *alles* is restricted to occurring with only particular Ā-operator phrases. Appositive relative clauses, as in (12a), or topics in the left periphery, as in (12b), do not license *alles* (example (12a) from Reis (1992: 471)). Reis generalizes this property of *alles* by stating that it is restricted to indefinite associates, i.e. *alles* is related to associates that denote an 'open set', "the variable expression that interrogative, restrictive relative, etc. operator phrases denote" (Reis, 1992: 472).

(12) a. *Diese Studenten, die alles den Test nicht bestanden haben, ...
these students, REL.NOM WQ the.ACC test not passed have.3PL
Intended: 'These students, which have all passed the test, ...'

b. *Diese/einige Bücher$_1$ hat Maria alles$_{1/n}$ rezensiert.
these/some books have.3SG Maria WQ reviewed
Intended: 'Maria reviewed all of these books.'

The interpretation of *alles* depends on the sentence. In restrictive relative clauses and *wh*-exclamatives, *alles* seems to exhaustively quantify over whatever its associate denotes. In *wh*-questions, *alles* additionally presupposes a plurality of *answers*, and creates the expectation for an exhaustive list of *answers* for the question-answer interaction to be felicitous. The question in (13A) can only be asked with the understanding that multiple answers will be given (and the usual understanding that the hearer knows this). The only acceptable answer to (13A) is (13Bi). It is the only answer that exhausts all answers; (13Bii) is not an exhaustive answer, while (13Biii) is not a plurality of *answers*.

(13) *Context: Lotte ate blueberries, raspberries, and blackberries with her breakfast.*

A: Welche Früchte hat Lotte alles zum Frühstück gegessen?
which fruit.PL have.3SG Lotte WQ to.the breakfast eaten
'What kinds of fruit did Lotte have for breakfast?'

B: i. Heidelbeeren, Himbeeren und Brombeeren.
blueberries raspberries and blackberries

ii. #Himbeeren und Brombeeren.
raspberries and blackberries

iii. #(Verschiedene) Beeren.
various berries

It should be noted, also, that in *wh*-questions *alles* is interpreted with the operator (or the variable) rather than with what is denoted by the operator's restriction. Reis (1992) shows this to be especially clear in the case of complex *wh*-phrases, such as 'which NP' or 'whose NP'. In the case of a possessor question, for instance, familiar A-dependency quantifier float (henceforth A-QF) quantifies over the restrictor NP, and this stays true in a *wh*-question.

This is shown in (14ab), such that (14b) could be understood as an answer to (14a).

(14) a. Wessen Münzen sind all-e auf 'Kopf' gelandet?
whose coins.NOM are all-NOM.PL on head landed
'Whose coins have all landed on heads?'

"What are all the people such that all of their coins landed on heads?"

b. Deine Münzen sind all-e auf 'Kopf' gelandet.
your coins.NOM are all-NOM.PL on head landed
'Your coins have all landed on heads.'

Alles differs in that it seems to quantify over, or be interpreted with, the operators *wessen* – or *wem* in the colloquial dative possessor construction – 'whose' directly, rather than with the *wh*-phrase restrictor NP as is the case with A-QF above.

(15) a. Wessen Münzen sind alles auf 'Kopf' gelandet?
whose coins are WQ on head landed
'What are all the people for which (some) coins have landed on heads?'

b. Wem seine Münzen sind alles auf 'Kopf' gelandet?
who.DAT his.PL coins are WQ on head landed
'What are all the people for which (some) coins have landed on heads?'

The difference between *alles* in (15a)-(15b) and A-QF in (14a) is, respectively, whether the question is a question about all the coin-throwers that had some (and maybe all) heads, or whether it is a question about only-heads coin-throwers. In other words, *alles* is not the same as the A-QF 'all-' plus some requirement to exhaust a salient non-atomic set of answers. Otherwise, it would be expected that in both cases only possessors (i.e. coin-throwers) for which all coins landed heads could be included in an answer. Consider the following model of the (relevant) world, where three people have thrown three coins each: Mina got three heads, Anouk got two heads and one tails, and Sig got one head and two tails.

(16)

PERSON	OUTCOME
Mina	{H, H, H}
Anouk	{H, H, T}
Sig	{H, T, T}

What seem to be the only truthful answers are given in (17): the A-QF question can only be answered by 'Mina's', because she is the only thrower of all heads, while the *alles*-question can only be answered by 'Mina's, Anouk's, and Sig's', because they are all the head-throwers there are, by virtue of each getting heads from at least one coin.

(17)

QUESTION	TRUE ANSWER
'whose coins A-QF heads'	'Mina's'
'whose coins WQ heads'	'Mina's, Anouk's, and Sig's'

Another way to see the difference of what the questions in (14a) and (15a)-(15b) are *about* is to look at possible answers. (14b), containing the quantifier, is a natural way to answer (14a) with a full sentence. Adding the same quantifier to a full-sentence answer to (15a)-(15b), as in (14b) or (18), on the other hand, feels like information is provided in the answer that was not asked for in the question—an infelicitous answer in a way. ('#' thus only pertains to (18) as an answer to (15b), independent of the context in (16).)

(18) Der Mina ihre Münzen sind (#all-e) auf 'Kopf' gelandet.
the.DAT Mina her.PL coins are all-NOM.PL on head landed
'Mina's coins have (all) landed on heads.'

2.3 Homophonous expressions that are not '*alles*'

Alles must be kept separate from two other expression which do or can surface as *alles*. The first is the A-QF inflecting *all-* 'all'. Though in certain context it can have the form *alles*, it must be kept separate from invariant *alles*. Sufficient justification was given in the previous section (but see Reis (1992) for more justification). The second one is *alles* appearing in predicational contexts or in impersonal imperatives (see in particular Giusti 1991).

In addition, the *alles* that is investigated in this paper appears to vary across speakers, such that it may be necessary to distinguish it from a third kind of *alles*. The reader is referred to appendix A for initial directions on the properties that appear to cluster together across speakers.

Throughout the paper, unless specified otherwise, '*alles*' will refer exclusively to the invariant *alles* introduced in the previous section.

2.4 Summary

Overall, it seems we can assign the following list of basic properties to invariant *alles*. The stress/focus property in particular will constitute an important potentially confounding factor when assessing certain judgments.

(i) *Stress/Focus: alles* cannot be stressed or focused.

(ii) *Associates: alles* requires its associate to contain an indefinite/non-anaphoric operator.

(iii) *Interpretation: alles* is interpreted with the operator/variable in its associate.

(iv) *Presupposition: alles* presupposes a plurality of (possible) answers.

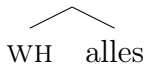
(v) *Answers: alles* creates the expectation of a full list of answers within a discourse-established domain.

These properties, given their systematicity ought to be important for any theory of *alles*, but will not receive more attention in this paper. We turn to the main set of *explananda* for *alles* in the next section, right after discussing the range of first-principles predictions the Same Source vs. Different Source hypotheses generate.

3 Generalizations about *alles*, and the same source different source question

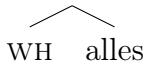
The main goal of this paper is to establish whether adjacent and distal *alles* are to be assigned the *same source*, or a *different source*. The two competing hypotheses are schematically repeated in (19)–(20).

(19) *Same source for adjacent and distal* alles*:*

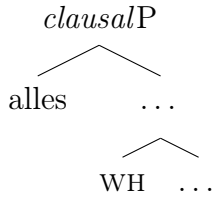


(20) *Different sources for adjacent vs. distal* alles*:*

a.



b.



This section lays out a number of core properties of German *alles*, and what are believed to be some of the empirical generalizations that need to be explained.[3] While doing so, it is argued that adjacent and distal *alles* need to be assigned a single common source. The conclusion is reached not based on a single fact or argument, but rather by the collection of generalizations established in this section, all pointing in the same direction. Most of these generalizations will in fact not, by themselves, refute the Different Source Theory. Bearing the two schemata in (19)–(20) in mind, it is thus helpful to point out the general,

[3]For more extensive empirical characterization of the phenomenon, the reader is referred to Reis (1992), as well as to Zimmermann (2007).

first-principles predictions that the two theories make, and highlight the ways in which they differ.

The SST, having only one source, predicts that there be a one-to-one correspondence between the associate and the quantifier. For each quantifier, there must be a corresponding associate, because the quantifier and the associate are introduced into the structure together—there is no other way of introducing the quantifier. This property will be called the *uniqueness property*. The fact that for each quantifier there must be a corresponding associate also means that they are in a dependency, and given that the dependency can be established in only one way, both adjacent and floated quantifiers must make the same contribution in meaning (cf. Sportiche, 1988). We can call this the *synonymy property*. The dependency is predicted to have other consequences, as well. For one, the dependency must be established, at least in part, by Merge. One would predict selection and agreement to be possible. We can call this the *form property*. In addition, it is predicted that the quantifier can, from its privileged position, be sensitive to transformations that apply to the associate in the course of a derivation. We can call this the *insider property*. The insider property has the broadest implications. Informally speaking, it predicts the quantifier to be sensitive to the associates derivational history. Depending on how the quantifier gets to be floated, it may be sensitive to all the steps leading up to the step that floats the quantifier (e.g. if stranding is achieved via movement), or it may even be sensitive to all steps, including those after (in derivational time) the quantifier is floated (e.g. if stranding is achieved via pronunciation of different parts of the chain). One important such 'sensitivity' concerns surface distribution, with the prediction that the positions that a floated quantifier can occur are constrained by those that the associate occupied in the course of a derivation—a floated quantifier's surface distribution is predicted to be a subset of its associate's distribution *per* given derivation. We can specify this consequence of the insider property as the *stranding property* (cf. Sportiche, 1988). Finally, because on standard minimalist assumptions any analysis that derives the float from a shared constituent necessarily involves movement (internal Merge), the insider property has the consequence that locality restrictions are expected to apply to the dependency between an associate and a floated quantifier, and that these mirror the locality restrictions that are observable for movement. We can call this property of the insider property the *local link property*. There is one last property that the SST entails. Because the quantifier and the associate enter in a relation at, essentially Deep Structure, the relation is 'licensed' right away. It is predicted, then, that while the relation between the quantifier and the associate can be negatively affected (or 'disrupted') by later stages of the derivation (i.e. yield strings that are ungrammatical for yet unknown reasons), the relation can not be established by later stages of the derivation (e.g. at the the interfaces). We can

call this last property the *early link property*.

The predictions for the DST are more difficult to foresee. The only clear prediction that the DST makes is that adjacent quantifiers and floated quantifiers have no systematically shared properties, *other than those* that follow from their shared semantics. According to the DST, in fact, it is their shared semantics that creates the illusion of a shared source (cf. Bobaljik, 2003: 125–128). Another way to say it, is that the only thing that the DST shares with the SST is the *synonymy property*.[4] It is thus possible that one quantifier may be agreeing, and the other not (no form property); the floated quantifier may appear in TP, while the associate could not have been in TP in that derivation (no stranding property); there may be locality restrictions on the adjacent quantifier for where inside the DP it may occur, while no locality restrictions apply to the floated quantifier (no local link property). Of course, it is *possible* that more properties be shared between the SST and the DST. If, e.g., the synonymy property follows from the fact that a quantifier and its associate must enter into an Agree relation for an index (cf. Heck and Himmelreich, 2017), the form property, the local link property, and to some extent the early link property, would be added with the synonymy property to the list of entailed properties due to the mechanics of Agree, making this version of the DST less distinguishable from the SST.

The fact that all these properties follow directly from the SST is a strong conceptual argument for the SST if empirically supported, and it's important to recognize that this remains true even if the SST is empirically supported only at varying degrees across languages: even if only different subsets of the properties named above correlated across different languages, the SST offers a straightforward learning account given that the richness of the facts would all fall out from independently available properties of the grammatical architecture and their interactions with the initial source and the mechanics of stranding. Possibly counter to that, it has been argued that different domains of quantifier floating may require separate explanations. Quantifier float in A-dependencies appears to exhibit some rather systematic gaps. For instance, when comparing A- and $\bar{\text{A}}$-quantifier float within West Ulster English, McCloskey (2000) questions whether the differences are reducible to A- vs. $\bar{\text{A}}$-movement properties. Reis (1992) raises the same issue for German, as mentioned in the previous section, and Fitzpatrick (2006) raises the issue more broadly, observing a semantic distinction between A-floats and $\bar{\text{A}}$-floats . English also systematically disallows A-floating in the base position (see Bošković (2004) for an analysis of the gap), while Merchant (1996) argues

[4] As noted by Howard Lasnik (p.c.), however, stating that the DST entails the synonymy property is too strong as it is not clear how *synonymy* is to be enforced in an explanatory way in a theory where adjacent and floated quantifiers are separate lexical items. The discussion below thus takes it for granted that there is such a solution, e.g. a UG-given condition on quantifiers so that they will always relate to some associate via the same mechanism—a property of all members of this syntactic category.

that A-float in the base position is possible in German. The domain of quantifier float in Ā-dependencies paints a different picture, where SST-analyses have been applied successfully across a variety of languages: Irish English *all* (McCloskey, 2000; Henry, 2012), Japanese numeral quantifiers (Miyagawa, 1989), Korean numeral quantifiers (Ko, 2007; Fitzpatrick, 2006), Russian non-inflecting quantifiers (Fitzpatrick, 2006), English quantity-measure adverbs (Davis, 2018).

The generalizations that this section will establish for *alles* will make it clear that the full bouquet of properties displayed by *alles* are all within the scope of predictions made by the SST, while any specific version of the DST would predict some but not others and eventually run afoul of re-coding the facts. *Alles*-floats thus lend further support to the SST in the domain of Ā-movement.

The remainder of this section is structured to follow the properties listed above, arguing step by step in favor of a SST approach and against a DST one, for *alles*, and until reasons to the contrary universally so.

(i) *Synonymy property:* Adjacent *alles* and distal *alles* make the same meaning contribution (section 3.1)

(ii) *Uniqueness property:* There is a one-to-one correspondence between *alles*s, whether adjacent or distal, and associates (section 3.2).

(iii) *Stranding property:* Distal *alles* occurs in a subset of the positions that its associate may occupy independent of *alles* (section 3.3).

(iv) *Insider property:* Distal *alles* occurs in Ā-trace but not in A-trace positions (section 3.4).

(v) *Local link property:* Distal *alles* must be c-commanded by a clausemate link of a suitable Ā-chain, in overt syntax (section 3.5).

(vi) *Early Link property:* The dependency between distal *alles* and its associate cannot be established at LF; it must be licensed in narrow syntax (section 3.5).

3.1 Synonymy property of adjacent-*alles* and distal-*alles*

The first generalization, given in (21), concerns minimal pairs such as (22a) and (22b).

(21) *Synonymy generalization:*
Adjacent and distal *alles* make the same contribution to sentence meaning.

(22) a. Wem alles musstest du helfen?
who.DAT WQ must.PST.2SG you help
'Who all did you have to help?'

b. Wem musstest du alles helfen?
who.DAT must.PST.2SG you WQ help
'Who all did you have to help?'

The two sentences in (22) differ only in the position of *alles*. In (22a) *alles* is in an associate-adjacent positions; in (22b) *alles* is in an associate distal position. The two sentences are synonymous.[5] This parallelism holds for Q-adjacent sentences and Q-distal sentences more generally, across grammatical role or function.

The synonymy property is predicted by both the SST and the DST.

3.2 Uniqueness property

Note: As Howard Lasnik (personal communication) points out to me, the argument in this section can ultimately not be established without resorting to interpretation. In order to exclude that *alles* can (but need not) associate with multiple *wh*-phrases, the following could be attempted: (i) an utterance plus context are created where *alles* could naturally be construed as exhausting both *wh*-phrases, and where (ii) the question is embedded as to create an assertion that could be challenged, but where (iii) it is not possible for the interlocutor to make a comment that grants the plurality/exhaustivity of one the *wh*-phrases while at the same time **denying** the plurality/exhaustivity of the other *wh*-phrase—importantly, as Alexander Williams (p.c.) points out, at a level of truth like "No/You are wrong". However, this turns out to be not at all trivial, both because of assumptions about exhaustivity and what is presupposed and what is asserted, and because of merely establishing clean judgments on this. As far as I can tell, what is out there is not sufficient to establish as much yet.

There is a uniqueness relation between *alles* and its associate.[6] This is predicted by the *uniqueness property* of the SST, but is not generally predicted by the DST.[7]

(23) *Uniqueness relation between* alles *and associate*
Each occurrence of *alles* A is uniquely mapped to an associate W.

[5] LF-identity, however, would be too strong a notion, as there are LF-intervention effects with *alles* (Beck, 1996), and quantifier-scope interactions (Pafel, 1991). I leave these issues aside in this paper. Note, though, that whatever explanation will be given for the A- vs. Ā-trace distribution generalization in section 3.4 could in principle account for these effects as well under plausible assumptions: (a) that only A-movement can escape LF-interveners; (b) scope reversal is only possible via A-movement in German.

[6] See also Zimmermann (2007) for discussion of this relation between *alles* and its associate.

[7] This particular formulation is due to Alexander Williams (p.c.).

There are three parts to the generalization. First, even though *alles* may be found in various positions of a sentence, there can be at most one *alles* per licit associate, cf. (24).[8] Sentences like (24) are perfectly fine sentences with only one of the three *alles*, but not acceptable, in fact word salad, as soon as a second or a third *alles* is added, regardless of which.

(24) *Wen alles hat der US-President alles an jenem Tag alles dem
who.ACC WQ have.3SG the.NOM US-president WQ at that day WQ the
Aussterben einen Schritt näher gebracht?
extinction a step closer brought
'Who all did the US-president bring one step closer to their extinction on that day?'

Second, for any one *alles* there must be a (licit) associate. This part amounts to a basic licensing requirement, and was already established in section 2.2: *Alles* needs to have a relation with a specific kind of operator, and if that operator is absent, i.e. *alles* has no associate, the sentence is ungrammatical. This is shown in (25), where *alles* fails to associate with the topic *der US-President*.

(25) Der US-President hat uns an jenem Tag (*alles) unserem
the.NOM US-president have.3SG us.ACC at that day WQ our
Aussterben einen Schritt näher gebracht.
extinction a step closer brought
'The US-president brought us (*all) one step closer to our extinction on that day.'

Lastly, each *alles* is *uniquely* mapped to an associate. This means that, as argued by Zimmermann (2007), each *alles* has at most one associate. Consider (26).

(26) (Und) wem_1 hat wer_2 dieses Foto $\text{alles}_{1,*2,*1+2}$ gezeigt?
and who.DAT have.3SG who.NOM this.ACC photo WQ shown
OK:'Who showed this photo to who all?'
#:'Who all showed this photo to who?'

In (26), *alles* may associate with the DAT, but not with the NOM *wh*-phrase. Conversely, *alles* may associate with the NOM but not with the ACC *wh*-phrase in (27).

(27) Wer_1 hat $\text{alles}_{1/*2/*1+2}$ was_2 der Bevölkerung gezeigt?
who.NOM have.3SG WQ what.ACC the.DAT population shown

[8] One exception seems to be in cases of Across The Board movement out of a conjunct:

(i) Wen wirst du [morgen (alles) trösten] und [übermorgen wieder (?alles) verletzen]?
who.ACC will.2SG you tomorrow WQ console and day.after.tomorrow again WQ hurt
'Who all will you console tomorrow and hurt again the day after?'

OK:'Who all showed the population what?'
#:'Who showed the population what all?'

This will come as no surprise under the SST given that in each case *alles* occurs in a position that the wrong associate can itself not occupy either. (This will become clearer in the following sections, in particular section 3.3.) In addition, Zimmermann (2007) shows that the unavailability of relating *alles* to the 'wrong' associate can be made clearer by manipulating the semantic compatibility of the one licit associate so that the sentences will become wholesale unacceptable. For instance, a context can be assumed in which the possible associate (DAT in (26), and NOM in (27)) is known to be atomic – or one can use a singular complex *wh*-phrase for the possible associate–, while the 'wrong' associate is left semantically licit: *alles* becomes completely unavailable in such cases. *Alles* requires its associate to denote an open class for which a plurality of answers can be given. The fact that failing the restriction only for one of the two in principle compliant *wh*-phrases makes *alles* completely unavailable thus means that *alles* cannot relate to the semantically compliant *wh*-phrases here. (28a) is the minimal pair to (26), (28b) to (27).

(28) a. *(Und) welcher Freundin hat wer dieses Foto alles gezeigt?
and which.DAT friend have.3SG who.NOM this.ACC photo WQ shown
Intended: 'For each x, x a friend, tell me what y showed this photo to x.'

b. *Welche Demonstrantin hat alles was der Bevölkerung
which.NOM.SG protester have.3SG WQ what.ACC the.DAT population
gezeigt?
shown
Intended:'For each x, x a woman protester, x showed the population some y, tell me what y x showed the population.'

Conversely, it is possible to have one *alles* for each of the two associates, without creating redundancy (or word salad):

(29) a. (Und) wem_1 $\{\text{alles}_1\}$ hat wer_2 alles_2 dieses Foto $\{\text{alles}_1\}$
and who.DAT WQ have.3SG who.NOM WQ this.ACC photo WQ
gezeigt?
shown
'Who showed this photo to who all?'

b. Wer_1 $\{\text{alles}_1\}$ hat $\{\text{alles}_1\}$ was_2 alles_2 der Bevölkerung
who.NOM WQ have.3SG WQ what.ACC WQ the.DAT population
gezeigt?
shown
'Who all showed the population what all?'

c. ?(Und) wem_1 hat wer_2 $alles_2$ $alles_1$ dieses Foto gezeigt?
and who.DAT have.3SG who.NOM WQ WQ this.ACC photo shown

d. ?Wer_1 hat was_2 $alles_2$ $alles_1$ der Bevölkerung gezeigt?
who.NOM have.3SG what.ACC WQ WQ the.DAT population shown

(28)-(29) together still do not constitute conclusive evidence that one *alles can never* satisfy more than one associate. However, they are strongly suggestive given that it is possible to have multiple *alles*s in sentences with multiple associates even though it is impossible to have more than one *alles* in sentences with only one associate. We know now that one *alles* is sufficient for one associate, and that adding another *alles* causes unacceptability. Whatever the reason causing unacceptability in sentences with 1 *alles* and 1 associate, if it were possible for a single *alles* to associate with multiple associates, then adding another *alles* should cause unacceptability in sentences with multiple associates, too. The reasoning is that, informally speaking, in both cases there is an *alles* that "isn't doing any work". A unique mapping from *alles* to associates, on the other hand, seems to explain the fact that sentences with multiple *alles* are acceptable so long as each *alles* can uniquely relate to one of the associates. We may thus conclude that one *alles* cannot have multiple associates.

3.3 Surface distribution of *alles*, Part 1: the stranding property

Turning to the surface distribution of *alles*, there are two generalizations that should be noted. The first generalization concerns the distribution of distal *alles* as compared to the distribution of its associate. The second generalization concerns the distribution of distal *alles* in A-trace vs. $\bar{\text{A}}$-trace positions and is introduced in section 3.4.

When *alles* occurs right-adjacent to its associate, it clearly forms a constituent with it. For the various positions that distal *alles* may occupy, the following generalization can be made:

(30) *Distribution generalization for distal* alles*, #1:*
Distal *alles* occurs in a subset of the positions that its associate may occupy independent of *alles*.

Alles may be floated in a variety of positions. Reis (1992), building on Pafel (1991), observes that in addition to *alles* occurring (a) in one constituent with, and right-adjacent to the associate, it may also occur in (b) the associate's base position, and (c) any position the

associate can reach via scrambling.[9,10] This section establishes (a)–(c) in detail, and adds a fourth distribution, namely *v*P of intermediate clauses in the path of long-distance *wh*-movement. Overall it establishes that the distribution of distal *alles* is always one of the positions that its associate occupied at some point in the derivation, i.e. distal *alles* has the distribution of the traces of its antecedent. Section 3.4 will refine this statement. The generalization follows from the SST's *stranding property*. A DST would need to either derive these facts from massive roll-up movements, or by allowing a high, and yet systematically restricted, degree of freedom for the addition of distal *alles* to the clause.

3.3.1 Right-adjacent to associate

We start off with the basic case where *alles* is right-adjacent to and in a constituent with the associate, as e.g. in (31). (The *wh*-phrase must be stressed, to at least some relative degree, for these sentences to satisfy the de-stressed property of *alles*.)

(31) Was alles hätte man besser nicht *e* tun sollen?
what.ACC WQ would.3SG one better not do should
'What all should one have better not done?'

Alles can also occur right-adjacent to its associate in other positions, e.g. in the multiple *wh*-question in (32), where the *wh*-phrase cannot move to Spec,C given that in German only one *wh*-phrase moves to Spec,C (as in English).

(32) Wo kann man was alles kaufen?
where can.3SG one what WQ buy
'Where can one buy what all?'

While it is less clear whether *alles* and the associate form a constituent in (32) (for reasons that will be made apparent in the following section), they must form a constituent in (31) given the generalization about German that only a single constituent can occupy the so-called

[9] Scrambling (in German) is adjunction to maximal verbal projection. In more modern terms it is movement to a specifier of a clause-medial projection, *v*P or TP; possibly both are possible with different interpretations associated as e.g. proposed in Heck and Himmelreich (2017). Scrambling has been classified as either A-movement (e.g. Haider and Rosengren, 1998; Frey, 2006), Ā-movement (e.g. Grewendorf and Sabel, 1999; Grewendorf, 2005), or as a third type altogether sharing properties of both (Webelhuth, 1992). Scrambling generally creates new A-binding possibilities.

[10] Reis (1992: 486) notes two exceptions to this generalization: "(a) in the position immediately adjacent to the Wackernagel position; (b) in the position immediately preceding the non-scrambling elements in the rightmost position(s) of the middle field". She dismisses them by observing that they all seem to be connected to the 'clitic property' that *alles* has in common with modal particles. It seems that an additional process is affecting *alles* and modal particles alike. Another, less construction specific way to understand the effects would be if some degree of reshuffling of prosodic units is possible at PF to accommodate de-accented material more broadly.

prefield, i.e. Spec,C of V2 clauses.[11]

3.3.2 Floated in the thematic/base position of the associate

More nuanced diagnostics are required to establish that distal *alles* can occur in the base position of its associate. German is a right-headed language–in the verbal domain–, with rather free word order. Identifying the base position can therefore be challenging, and particularly so for objects. Scrambling, i.e. movement to a clause-medial position targets positions outside the VP. Finding *alles* in positions below the subject, an object, or certain adverbs, will not always be evidence for *alles* occurring in base position. It may well be that *alles* is occurring in a position higher than the thematic position, but still lower than the position reached by other elements via scrambling. To exemplify, the object *den Kuchen* 'the cake' in (33) may occur either to the right (*viz* below) the subject, or to its left (*viz* above it). In (34), however, with the subject out of the way, it is unclear whether the object is in its base position or not, i.e. how to distinguish between (34a) and (34b).

(33) a. [$_{CP}$ dass keiner den Kuchen mag]
that noone.NOM the.ACC cake like.3SG

b. [$_{CP}$ dass [$_{vP}$ [$_{DP}$ den Kuchen]$_1$ [$_{vP}$ keiner e_1 mag]]]
that the.ACC cake noone.NOM like.3SG
'that no-one likes the cake'

(34) Ich weiß nicht,
'I don't know'

a. [$_{CP}$ wer [$_{C'}$ ... [$_{VP}$ [$_{DP}$ den Kuchen] mag]]]
who.NOM the.ACC cake like.3SG

b. [$_{CP}$ wer$_2$ [$_{C'}$ ... [$_{vP}$ [$_{DP}$ den Kuchen]$_1$ [$_{vP}$ e_2 [$_{VP}$ e_1 mag]]]]]
who.NOM the.ACC cake like.3SG
'who likes the cake.'

Reis (1992) provides a compelling argument, based on the verb *aussetzen*, whose two objects cannot revert their relative base order through scrambling. I will add two further arguments, so that there are a total of three diagnostics to diagnose the base position of the associate:

(i) verbs with rigid object order, which constrain the availability of scrambling of one object

(ii) the position of *wh*-indefinites, which, in the general case, mark their thematic position,

[11] For a brief overview and some discussion see for example Müller (to appear: section 1).

(iii) the position of certain focused adverbs, which mark the left edge of the focused VP.

Fixed relative object word order verbs Marga Reis gives the following minimal pair to argue for the availability of *alles* in the base position of its associate (Reis 1992: 483; glosses added). (Throughout the paper, following the convention in the German literature, capitalized words or syllables indicate focus.)

(35) a. *Wen hat er der Prüfung alles HEUTe ausgesetzt?
who.ACC have.3SG he.NOM the.DAT exam WQ today subjected
Intended: 'Who all did he subject to the exam today?'

b. Wen hat er der Prüfung denn HEUTe ausgesetzt?
who.ACC have.3SG he.NOM the.DAT exam DPRT today subjected
'Who did he subject to the exam today?'

She notes that the relative word order of the two objects of the verb *aussetzen* 'subject to' can, for some reason, not be reversed through scrambling. This can be shown with regular DPs, both of which being definite should be able to scramble one over the other, cf. (36ab).

(36) a. Hat er den Peter der Prüfung HEUTe ausgesetzt?
have.3SG he.NOM the.ACC Peter the.DAT exam today subjected
'Did he subject Peter to the exam today?'

b. *Hat er der Prüfung den Peter HEUTe ausgesetzt?
have.3SG he.NOM the.DAT exam the.ACC Peter today subjected

In (35a), then, if *alles* is marking the base position of the accusative object, scrambling the dative over it induces unacceptability for the same reason that scrambling the dative object over the accusative object itself does in (36b). Reis (p.484) strengthens her argument by showing that the effect is induced also when the dative object scrambles over the accusative *wh*-phrase (with or without *alles*). *Wh*-phrases stay in-situ in multiple *wh*-questions in German, but may themselves scramble.[12]

(37) *Wer hat der Prüfung wen (alles) HEUTe ausgesetzt?
who.NOM have.3SG the.DAT exam who.ACC WQ today subjected
Intended: 'Who subjected who (all) to the exam today?'

This paradigms shows that distal *alles* behaves, at a point in the derivation where the associate may not have moved on, as if its associate was there, too. Other adverbials (*heute*) or particles (*denn*) do not induce the same effect. It is overall a strong argument in favor of the SST. On a DST, *alles* cannot be identified with the associate, syntactically, so that it would need to be the relation between *alles* and the associate that is responsible for the

[12] This is a contested, but incontrovertibly true fact. See footnote 20.

contrast. However, it seems that the outcome of the relation, if not its derivation (cf. section 3.1), cannot be responsible for the contrast. Adjacent *alles* has no effect:

(38) Wen alles hat er der Prüfung HEUTe ausgesetzt?
who.ACC WQ have.3SG he.NOM the.DAT exam today subjected
'Who all did he subject to the exam today?'

Reis's paradigm does not yet, however, show that *alles* is in the base position of its associate. *Alles* is to the left of the adverb *heute* and therefore outside the VP and not in the object's base position. This detail is easily fixed, and it is indeed possible for *alles* to occur to the adverb's right in such sentences, cf. (39) (more details on how adverbs can diagnose base positions follows in the section on focused adverbs).

(39) Wem seinen Tests hat er HEUTe den Peter *e* alles
who.DAT his.DAT.PL tests have.3SG he.NOM today the.ACC Peter WQ
ausgesetzt?
subjected
'The tests of who all did he subject Peter to today?'

***Wh*-indefinites** The second diagnostic for base positions is based on *wh*-indefinites. The position of *wh*-indefinites can be used as a diagnostic because *wh*-indefinites in German cannot scramble.[13] As such, they mark their thematic position, in the case of arguments, or simply more generally their base position. Consider a ditransitive verb such as *zeigen* 'show', which has the underlying object structure DAT≫ACC (where '≫' stands for 'c-commands').[14]

[13] See Haider (1993: 200, fn2). *Wh*-indefinites can also not be topicalized, as in (i), or, as a subject, occur to either side of weak object pronouns, as regular subjects can, cf. (ii).

(i) *Wen habe ich gesehen.
WND.ACC have.1SG I.NOM seen
Intended: 'I saw someone.'

(ii) a. dass {Peter/*wer} ihn dir {Peter/wer} vorgestellt hat
that Peter/WND him.ACC you.DAT Peter/WND introduced have.3SG
'that Peter/someone introduced him to you'

However, it would be too strong to say that *wh*-indefinites cannot move at all, as they can be the subject of a(n extraposed) raising complement.

(iii) dass was droht, anzubrennen.
that WND.NOM threaten.3SG to.burn.at
'that something threatens to get burnt.'

[14] For example, a negative quantifier can bind a pronoun from the DAT into the ACC, but not vice-versa:

(i) a. dass 'n Lehrer [keinem Schüler]$_i$ sein$_i$ neues Klassenzimmer gezeigt hat.
that a teacher no.DAT student his new classroom shown have.3SG
'that a teacher showed no student their new classroom.'

The word order facts in (40) reflect the asymmetry of objects with *zeigen*. In (40a), the DAT *wh*-indefinite successfully precedes the ACC indefinite; in (40b), the inverse word order fails; in (40c), the DAT indefinite successfully precedes the ACC *wh*-indefinite; in (40d), the opposite word order fails.[15]

(40) a. dass 'n Lehrer wem 'ne Schlange gezeigt hat.
that a.NOM teacher WND.DAT a.ACC snake shown have.3SG
'that a teacher showed a snake to someone.'

b. *dass 'n Lehrer 'ne Schlange wem gezeigt hat.
that a.NOM teacher a.ACC snake WND.DAT shown have.3SG

c. dass 'n Lehrer 'nem Schüler was gezeigt hat.
that a.NOM teacher a.DAT student WND.ACC shown have.3SG
'that a teacher showed something to a student.'

d. *dass 'n Lehrer was 'nem Schüler gezeigt hat.
that a.NOM teacher WND.ACC a.DAT student shown have.3SG

If all the examples in (40) indeed reflect base positions, i.e. do not involve any scrambling, we can understand why (40ac) are acceptable word orders, but (40bd) are not. The acceptable sentences reflect a licit base structure for *zeigen* (DAT≫ACC), while the unacceptable sentences reflect an ungrammatical base structure for *zeigen* (ACC≫DAT). We can show the same facts with two *wh*-indefinites:

(41) a. dass 'n Lehrer wem was gezeigt hat.
that a.NOM teacher WND.DAT WND.ACC shown have.3SG
'that a teacher showed something to someone.'

b. *dass 'n Lehrer was wem gezeigt hat.
that a.NOM teacher WND.ACC WND.DAT shown have.3SG

Crucially, it is possible to scramble a definite (and sometimes indefinite) full DP to a higher position. (42b), but especially (42a) are exactly what wasn't possible in (41) and (40). If *wh*-indefinites cannot scramble and rather have to stay in their base position the word order facts fall out directly from the underlying order of objects of *zeigen*. Similar patterns hold for verbs of the same class, and the reverse pattern hold for ditransitives of the class with underlying ACC≫DAT.

(42) a. dass 'n Lehrer die Schlange wem gezeigt hat.
that a.NOM teacher the.ACC snake WND.DAT shown have.3SG

b. *dass 'n Lehrer sein-en$_i$ Mitschülern [keinen neuen Schüler]$_i$ gezeigt hat.
that a teacher his-DAT.PL peer no.ACC new student shown have.3SG
'that a teacher showed no new student to their peers.'

[15]It is important to give monotonically falling declarative intonation to avoid effects of contrastive focus.

b. dass die Schlange 'n Lehrer wem gezeigt hat.
that the.ACC snake a.NOM teacher WND.DAT shown have.3SG
'that a teacher showed the snake to someone.'

Finding an expression below a dative *wh*-indefinite with *zeigen* thus constitutes evidence for that expression occupying a position in the thematic nucleus of the VP, plausibly the ACC-argument position itself. This is indeed possible for *alles*, cf. (43A).

(43) A: Weißt du, [$_{CP}$ was der Lehrer wem alles gezeigt haben
know.2SG you what.ACC the.NOM teacher WND.DAT WQ shown have
soll?
MOD.3SG
'Do you know what all the teacher supposedly showed to someone?'

B: Ne, nur, dass er wem 'ne Schlange gezeigt haben soll.
no only that he WNDDAT a.ACC snake shown have MOD.3SG
'No, just that he supposedly showed someone a snake.'

Focused adverbs The next diagnostic involves focused adverbs.[16] Focused adverbs are a good diagnostic for argument base positions because, as a general understanding of German focus-background structure, focused constituents do not scramble.[17] As a consequence, VP-level adverbs mark the left edge of the focused VP when they are focused. For distal *alles*, we predict that it can occur to the right of the focused adverb if it can be in base positions of its associate. This is indeed what we find in (44). To get the right focus structure, one can understand the sentences as answers to the question *Was weißt du über Susi?* 'What do you know of/about Susi?' (ignoring the fact that they are perhaps unusually specific).

(44) Ich weiß zum Beispiel,
I know for instance

a. wem Susi GERne alles ein Geschenk mitbringen würde.
who.DAT Susi gladly WQ a.ACC present bring.with would.3SG
'I know, for instance, who all Susi would bring a present with pleasure.'

b. was Susi GERne einem Kind alles mitbringen würde.
who.ACC Susi gladly WQ a.DAT child bring.with would.3SG
'I know, for instance, what all Susi would bring a child with pleasure.'

As a test case, consider Lenerz' generalization (Lenerz, 1977), which states that direct objects

[16] Focus here is to be understood in a way that excluded contrastive focus, which has different prosodic properties (a "hut"-contour), and allows for more syntactic freedom of movement. The VP-focused intonation rather is monotonically falling.

[17] Reis (1992: fn18) points to Lenerz (1977), von Stechow and Sternefeld (1988: 466), Grewendorf and Sternefeld (1990: 15), Fanselow (1990: 115ff).

may only scramble when definite. This generalization captures the word order facts in (45), where the definite ACC-object may scramble over the DAT-object (in a DAT≫ACC verb) only when it is definite ((45b) vs. (45d)).

(45) a. Susi hat dem Kind das Geschenk mitgebracht.
Susi have.3SG the.DAT child the.ACC present brought.with
'Susi brought the child the present.'

b. Susi hat das Geschenk dem Kind mitgebracht.
Susi have.3SG the.ACC present the.DAT child brought.with

c. Susi hat dem Kind ein Geschenk mitgebracht.
Susi have.3SG the.DAT child a.ACC present brought.with
'Susi brought the child a present.'

d. *Susi hat ein Geschenk dem Kind mitgebracht.
Susi have.3SG a.ACC present the.DAT child brought.with

If we add in a focused adverb, e.g. *gerne* 'with pleasure', we see that it may occur in positions preceding a string that is compatible with the canonical order, but not in positions preceding a string that is not compatible with the canonical order, as indicated by the impossibility of having focused *gerne* in front of the ACC-object in (46b). The prosody will be homogeneously falling from the focused adverb on to the rest of the VP. (Schematics are included below each sentence for expository purposes.)

(46) a. Susi hat (GERne) dem Kind (GERne) das Geschenk (GERne)
Susi have.3SG gladly the.DAT child the.ACC present
mitgebracht.
brought.with
'Susi brought the child the present with pleasure.'

b. Susi hat (*GERne) das Geschenk (GERne) dem Kind (GERne)
Susi have.3SG gladly the.ACC present the.DAT child
mitgebracht.
brought.with

c. [(ADV) DAT (ADV) ACC (ADV) V]

d. [(*ADV) ACC (ADV) DAT (ADV) V]

In addition, we see that when the (lower) ACC-object is indefinite, any string where the focused adverb is preceded by the indefinite ACC-object is illicit (again with the monotonically falling prosody).

(47) a. Susi hat (GERne) dem Kind (GERne) ein Geschenk (*GERne)
Susi have.3SG gladly the.DAT child a.ACC present
mitgebracht.
brought.with

'Susi brought the child a present with pleasure.'

b. *Susi hat (GERne) ein Geschenk (GERne) dem Kind (GERne)
Susi have.3SG gladly a.ACC present the.DAT child
mitgebracht.
brought.with

c. [(ADV) DAT (ADV) ACC-INDEF (*ADV) V]

d. *[(ADV) ACC-INDEF (ADV) DAT (ADV) V]

Generalizing, scrambling minimally targets a position higher than the projection minimally including the focused adverb: (i) non-canonical word orders are derived by scrambling, such that strings containing non-canonical object word orders cannot follow the focused adverb as scrambling would have then targeted a position lower than the focused adverb (as in (46b)); (ii) the (lower) indefinite ACC-object may not scramble and therefore may not precede the focused adverb (as in (47ab)).[18]

The fact that adding *alles* below the ACC-object causes unacceptability when questioning the DAT-object is thus further support for the construal of the diagnostic, as well as for the idea that *alles* occurs in the base position of its associate when it occurs to the right of focused VP-adverbs. The unacceptability of (48a) mirrors the unacceptability of (48b). As was the case with the *aussetzen* 'subject to' paradigm, the relation between distal *alles* and its associate, alone, cannot be responsible for the effect given that changing distal *alles* in (48a) to adjacent *alles* in (48c) makes the sentence fully acceptable again.[19]

(48) a. *Wem hat Susi GERne ein Geschenk alles mitgebracht?
who.DAT have.3SG Susi gladly WQ a.ACC present bring.with
Intended: 'Who all did Susi gladly bring a present?'

b. *Hat Susi GERne ein Geschenk dem Kind mitgebracht?
have.3SG Susi gladly a.ACC present the.DAT child brought.with
Intended: 'Did Susi gladly bring a child a present?'

c. Wem alles hat Susi GERne ein Geschenk mitgebracht?
who.DAT WQ have.3SG Susi gladly a.ACC present bring.with
'Who all did Susi gladly bring a present?'

[18] The *wh*-indefinite test and the focused adverb test can be combined to check one against the other. Indeed, a *wh*-indefinite may not occur to the left of a focused adverb:

(i) Susi hat (GERne) wem (*GERne) ein Geschenk mitgebracht.
Susi have.3SG WND.DAT a.ACC present brought.with
'Susi brought someone a present with pleasure.'

[19] (48a) looks like an intervention effect configuration as discussed in Beck (1996) and Heck and Himmelreich (2017). It should be thus noted that at least some examples of that kind can be explained by the stranding property of *alles*.

3.3.3 Scrambling positions

As discussed in the previous section, focused adverbs mark the left edge of focused VPs. The fact that *alles* can occur to their left thus indicates that *alles* may occur in positions that their associate reaches via scrambling.

(49) Ich weiß nicht,
I don't know

a. wem Susi alles GERne *e* ein Geschenk mitgebracht hat.
who.DAT Susi WQ gladly a.ACC present brought.with have.3SG
'I don't know who all Susi brought a present with pleasure.'

b. was Susi alles GERne einem Kind *e* mitgebracht hat.
who.ACC Susi WQ gladly a.DAT child brought.with have.3SG
'I don't know what all Susi brought a child with pleasure.'

The *wh*-phrase itself may in fact also occupy this position. (50) shows that the "in-situ" *wh*-phrase of multiple *wh*-questions may scramble, given that the *wh*-phrase may occur either to the left or the right of the focused adverb.[20]

(50) a. Wer würde UNgerne wem ein Geschenk mitbringen?
who.NOM would.3SG not.gladly who.DAT a.ACC present bring.with
'Who wouldn't like to bring whom a present?'

b. Wer würde wem UNgerne *e* ein Geschenk mitbringen?
who.NOM would.3SG who.DAT not.gladly a.ACC present bring.with

There are two more facts that show that *alles* may occupy scrambling positions (that its associate itself may also scramble to). The first set is about the relative order of *alles* and the higher object of a ditransitive, the second set about the position of *alles* relative to weak object pronouns. The first set is positive data, where both the associate, and *alles*, may occupy a position that can be reached via scrambling of a comparable category; the second set is negative data, where both the associate, and *alles*, may *not* occupy a position that can*not* be reached via scrambling of a comparable category.

First, consider where *alles* can occur relative to the DAT object, which is underlyingly higher than the associate ACC object of *zeigen* 'show', in (51a).

(51) a. Was hat die Demonstrantin {alles} der Bevölkerung
what.ACC have.3SG the.NOM protester WQ the.DAT population

[20] The claim that *wh*-phrases may scramble is contended in the literature. Other authors' examples and conclusions, however, are in line with the results discussed here (e.g. Reis, 1992; Wiltschko, 1997; Heck and Himmelreich, 2017). Given a general understanding of scrambling as variable word order in the *middle field*, the opposing view, that *wh*-phrases cannot scramble, seems empirically untenable, then. See also Reis (1992: fn19) and references therein.

{alles} gezeigt?
WQ shown
'What (all) did the woman protester show the population?'

b. Wer hat {was (alles)} der Bevölkerung {was (alles)} gezeigt?
who.NOM have.3SG what.ACC WQ the.DAT population what.ACC WQ shown
'Who showed the population what (all)?'

Alles may occur to either the left or the right of the DAT object in (51a). Given that the underlying order of objects for *zeigen* is DAT≫ACC, the *alles* to the left of the DAT object is occupying a scrambling position. In addition, notice that the same position may be filled, with or without *alles* by the associate *wh*-phrase itself, as e.g. in (51b) where a *wh*-subject is preventing the ACC *wh*-phrase from moving up further. Notice also that the scrambled ACC *wh*-phrase is part of a genuine multiple *wh*-question rather than a scrambled echo-questioned DP as witnessed by the possibility of adding the discourse particle *denn* to the sentence, cf. (52). *Denn* is parasitic on interrogative force and resists echo-questions, cf. (53).[21]

(52) Wer hat denn was (alles) der Bevölkerung gezeigt?
who.NOM have.3SG DPRT what.ACC WQ the.DAT population shown
'Who showed the population what (all)?'

(53) a. ??Wer hat denn WAS (alles) der Bevölkerung gezeigt?
who.NOM have.3SG DPRT what.ACC.ECHO WQ the.DAT population shown
'Who showed the population WHAT (all)?'

b. *Sie hat denn WAS der Bevölkerung gezeigt?
she.NOM have.3SG DPRT what.ACC.ECHO the.DAT population shown
'She showed the population WHAT?'

Next, consider where *alles* may occur relative to weak object pronouns, for instance *sich* '3-self' in (54a), and *ihm* '3DAT' in (54b):[22]

(54) a. Wen soll man {?*alles} sich {alles} dabei vorstellen?
who.ACC MOD.3SG one.NOM WQ REFL WQ that.by imagined
'Who all is one supposed to think of based on that?'

[21] For more general information about *denn*, see Bayer et al. (2016) and references therein.

[22] 'Weak' stands for the phonological status. Weak objects in German are de-stressed, and, depending on dialect may be reduced to various degrees. They occupy what is traditionally referred to as the *Wackernagel position*, which is either the leftmost edge of *v*P, or some specifier of TP, such that definite DP subject may occur on either side of weak object pronouns, cf. (55). See e.g. Müller (2001); Anagnostopoulou (2008) and references therein.

b. Was hat {?*alles} ihm {alles} keiner *e* geben wollen?
what.ACC have.3SG WQ him.DAT WQ noone.NOM give want
'What all did no-one want to give him?'

Alles may not occur to the left of the weak pronouns, a position that subjects may occupy, but non-pronominal objects may not, not even via scrambling. This is shown for a subject in (55a), for a definite accusative object in (55b), and a *wh*-phrase in (55c). Once again these facts hold for intonations that exclude the contrastive focus hut-contour raising on ACC and falling on NOM.

(55) a. dass {Maria} ihm {Maria} GERne was abgegeben hätte.
that Maria him.DAT Maria gladly WND.ACC give.away have.COND.3SG
'that Maria would have gladly given him something.'

b. dass {*den Apfel} ihm {den Apfel} keiner *e* abgegeben
that the.ACC apple him.DAT the.ACC apple noone.NOM give.away
hätte.
have.COND.3SG
'that Maria would have given him the apple.'

c. Wann/wo {*was} ihm {was} keiner *e* abgegeben
when/where what.ACC him.DAT what.ACC noone.NOM give.away
hätte, ist unklar.
have.COND.3SG be.3SG unclear
'It is unclear when/where no-one would have given him what.'

Alles thus may also *not* appear in positions that its associate may *not* reach via scrambling. To complete the picture, one would expect based on the asymmetry between subjects and objects that subject-*alles* can to the left of a weak object pronoun. The facts are less clear here. It seems that at least for *sich* 'self' it is marginally possible to get *alles* to the left of a weak pronoun. However, the availability mostly goes away with a complex *wh*-phrase associate.[23]

(56) a. Wer möchte sich *e* alles {NOCH/BITte} selbst anzeigen?
who.NOM want.3SG REFL WQ still/please self lay.charge.against
'Who all wants to lay charges against themselves?'

b. ?Wer möchte alles sich *e* {NOCH/BITte} selbst anzeigen?
who.NOM want.3SG WQ REFL still/please self lay.charge.against

[23] Weak objects other than *sich* (e.g. *mich/dich* 'me/you.SG') seem to fully resist *alles* to their left. For this reason, there is a focused element separating *sich* and *selbst*, to ensure that the monotonically falling prosody of the focused VP would start after *sich*, and therefore *sich* be an unfocused weak pronoun. Thanks to Julian Schlöder for noting that the modal particles *noch* and *bitte* ameliorate judgments (for obscure reasons).

c. Wem seine Mandanten möchten sich *e* alles NOCH selbst
who.DAT his.PL clients want.3PL REFL WQ still self
anzeigen?
lay.charge.against
'Whose clients all want to lay charges against themselves?'

d. ??Wem seine Mandanten möchten alles sich *e* NOCH selbst
who.DAT his.PL clients want.3PL WQ REFL still self
anzeigen?
lay.charge.against

It seems, then, that in fact not all positions that can be reached via scrambling are possible positions for distal *alles*. However, this asymmetry is just apparent because a *wh*-phrase subject is also marginal when to the left of a weak object, cf. (57).[24] This contrast lends support to the idea that something special seems to be going on with subject scrambling, and, more importantly, that the distribution of distal *alles* is correlated with the distribution of its associate itself.

(57) Wann {?/??wer} ihm$_i$ {wer} seine$_i$ Medizin geben sollte,
when who.NOM him.DAT who.NOM his.ACC medicine give
war unklar.
should.PST.3SG be.PST.3SG unclear
'It was unclear who was to give him his medicine when.'

3.3.4 Intermediate clauses of long-distance *wh*-movement

The last kind of position where distal *alles* can occur is Spec,*v* of intermediate clauses in the path of long-distance *wh*-movement of the associate. Under the assumption that long-distance *wh*-movement is bounded (Chomsky, 1973, 1977), and that specifically it must proceed through the specifier of certain projections, including at least CP and *v*P (Chomsky, 1986a, 2000, 2001),[25] the SST predicts that distal *alles* may, in principle, occur in these positions as a consequence of the *stranding property* of the SST. Floating of *alles* in CP is not possible, for any speaker asked; floating of *alles* in *v*P, however, is possible (for speakers who allow extraction from finite CPs in the first place–a regional characteristic of southern varieties of German):[26]

[24] The degree of unacceptability depends once again on whether there is a raising intonation on *wer* as the beginning of a hut-contour falling on *Medizin*; having this prosody corresponds to the less marginal, perhaps even fully acceptable judgment. Definite DP subjects do not require this prosody to be fully acceptable when occurring to the left of weak object pronouns.

[25] Specifically, in Chomsky (1986a), long movement was broken up by intermediate steps of *adjunction* to maximal projections that allowed it (according to a heterogeneous set of criteria).

[26] This is yet another way in which invariant *alles* is to be distinguished from inflecting *all-*. Bobaljik (2003: 121) notes that *all-* cannot be floated in CP, to which we can add that floating in the same position

(58) [$_{CP}$ Wem$_1$ hat [der Andreas]$_2$ [$_{vP}$ (alles) e_2 gedacht, [$_{CP}$ (*alles) dass
who.DAT has the.NOM Andreas all thought all that
die Georgine noch e_1 einen Schnapps einschenken würde]]] ?
the.NOM Georgine yet a.ACC schnapps pour would.3SG
'Who (all) did Andreas think that Georgine would pour another schnapps?'

We can establish that *alles* in the matrix clause in (58) is in some outer specifier of *v*P by comparing the position of *alles* with respect to a subject in the associate's clause of origin with the position of *alles* with respect to a subject in the matrix clause. Consider (59). *Alles* is in its associate's clause of origin, and it can occur on either side of a *wh*-indefinite subject.[27]

(59) (A:) Was weißt du noch von gestern? (B:) Ich weiß noch,...
(A:) What can you remember from yesterday? (B:) I still remember...

a. ?wen gestern wer alles getroffen hat.
who.ACC yesterday WND.NOM WQ met have.3SG
'I still remember who all someone met, yesterday.'

b. wen gestern alles wer getroffen hat.
who.ACC yesterday WQ WND.NOM met have.3SG
'I still remember who all someone met, yesterday.'

Now compare the (absence of a) contrast in (59ab) with the contrast in (60ab).[28] *Alles* can occur to either side of the *wh*-indefinite subject in the clause of origin of its antecedent ((59)), but it can occur only to the right of the *wh*-indefinite subject of the matrix clause

where invariant *alles* can be floated is also impossible.

(i) Welche Würste hat der Peter [$_{vP}$ (*all-e) gesagt [$_{CP}$ (*all-e) dass der Hund gegessen hat]]?
which sausages has the Peter all-PL said all-PL that the dog eaten has
'Which sausages did Peter (*all) say that the dog ate?'

Notice that Merchant (1996) argues that inflecting *all-* can be stranded in any position from which scrambling may occur, both A- and Ā-scrambling. If Merchant's analysis was correct, and both A- and Ā-scrambling were to exist in German, then the difference between inflecting *all-* and invariant *alles* cannot be attributed to postulating that inflecting *all-* can occur only in A-trace positions, and invariant *alles* can occur only in Ā-trace positions (as will be argued is the case for *alles*; cf. section 3.4). If, on the other hand, scrambling was uniformly A-movement in German, it might be possible that the complementary distribution of floated inflecting *all-* and floated invariant *alles* would correlate with A- vs. Ā-traces, respectively.

[27] It is plausible that the slight degradation of (59a) is due to the interfering factor of trying to parse the sentence as an embedded multiple question, which would yield stress on *wer*, and more easily satisfy the de-accented property of *alles*.

[28] The question marks in front of '*' in (60a), and nothing in (60b), are to acknowledge that these sentences are harder to judge. The relative contrast is clear, nonetheless. It may thus turn out that (60b) is worse than indicated, or that the relative difference between the (a) and (b) examples in (59) and (60) is the same if tested in more controlled conditions. In the latter case, in particular, it would mean that *wh*-indefinite subjects are not a valid diagnostic.

((60)).

(60) (A:) Was weißt du noch von gestern? (B:) Ich weiß noch,...

(A:) What can you remember from yesterday? (B:) I still remember...

wen gestern wer gemeint hat, dass Susi getroffen hat.
who.ACC yesterday WND.NOM reckoned have.3SG that Susi met have.3SG
'I still remember who all someone said, yesterday, that Susi met.'

a. ?*wen gestern *wer* **alles** gemeint hat, dass Susi getroffen hat.

b. ?wen gestern **alles** *wer* gemeint hat, dass Susi getroffen hat.

Identifying *v*P both as the clause medial phase (or escape hatch), and as the projection in which the subject is introduced, one must conclude that *alles* in matrix clauses occurs no lower than *v*P, specifically a specifier of *v*P that is higher than where the subject is introduced.[29] Identifying an upper bound is more difficult in terms of naming specific clause medial projections. It is clear, however, that *alles* in intermediate/matrix clauses, just as in an associate's clause of origin, may not occur to the left of a weak, unstressed object pronoun:

(61) (A:) Was weißt du noch von gestern? (B:) Ich weiß noch,...

(A:) What can you remember from yesterday? (B:) I still remember...

wen der Peter ihm gesagt hat, dass Susi getroffen
who.ACC the.NOM Peter him.ACC said have.3SG that Susi met
hat.
have.3SG
'I still remember who all Peter told him, yesterday, that Susi met.'

a. wen der Peter *ihm* **alles** gesagt hat, dass Susi getroffen hat.

b. *wen der Peter **alles** *ihm* gesagt hat, dass Susi getroffen hat.

Let's assume that weak pronouns, given their prosodic status, occur at the left edge of the projection they syntactically occur in. If the Wackernagel position, to which weak pronouns move in German, is TP, it is possible that *alles* in these examples is actually in TP given that the weak object would be in the leftmost TP position. If, on the other hand, the Wackernagel position is in *v*P, then *alles* would be fully confined inside of *v*P, specifically

[29] Note, however, that the results are not compatible with a decomposition of the VP into voiceP-*v*P-VP (see, e.g., Marantz 2001 and references therein), where voiceP introduces the external argument, VP the internal argument, and *v*P is the verbalizing projection, and voiceP *and* *v*P are phases. If both voiceP and *v*P were phases, all else equal, one would expect, contrary to fact, that *alles* can appear on either side of an external argument. And if *v*P as understood here were below ApplP, which introduces the second internal argument, we would expect *alles* to be able to appear below an indirect object in base position. To remedy these proposals one would have to assume that *v*P and ApplP do not project escape hatches (at least not for Ā-movement).

to an outer specifier given that it must precede the subject. The facts are not conclusive so far.[30]

3.4 Surface distribution of *alles*, Part 2: the insider property

The second generalization concerning the surface distribution of distal *alles* concerns its distribution relative to the traces of its antecedent:

(62) *Distribution generalization for distal* alles*, #2:*
Distal *alles* can occur in positions corresponding to the tail of Ā-movement, but not in positions corresponding to the tail of A-movement.

Using the distinction between A- and Ā-traces as a formal tool of description, generalization (62) says that, in any given derivation, *alles* may occupy positions from which its associate has Ā-moved, but not positions from which its associate has A-moved. The second distribution generalization thus restricts the first distribution generalization, such that the new generalization in (63) can be given to cover both sets of facts:[31]

[30] To try to distinguish between TP and *v*P one could add adverbs of likelihood, which are presumably in TP (see e.g. Merchant, 1996: 185–188): It is possible, if not obligatory, to have weak objects follow the adverb of likelihood *vermutlich* 'presumably', cf. (ia). If *alles* could only appear in the rightmost position of the four indicated in (ibc)–i.e. only to the right of *ihm* in (ib)–, then one could be more confident that *alles* cannot be in TP in these cases. The logic is that if the adverb is in TP, and relative word orders have to do with free ordering of adjunction to TP, then one would expect *alles* to occur freely in any of the TP-adjoined positions. While this seems to be going in the right direction for my own judgments, these sentences are very difficult to judge so that I do not report judgments here. I indicate this by prefixing the sentences with '$'.

(i) a. Dann hat der Peter {ihm} vermutlich {ihm} gestern gesagt, [$_{CP}$ dass
then have.3SG the.NOM Peter him.DAT presumably him.DAT yesterday said that
er ihm 'ne LaSAgne vorbeibringt].
he.NOM him.DAT a.ACC Lasagne by.bring.3SG
'(That must mean that) Peter presumably told him, yesterday, that he is bringing him some lasagna.'

b.$ Was hat der Peter vermutlich (alles) ihm (alles) gestern gesagt, [$_{CP}$
what.ACC have.3SG the.NOM Peter presumably WQ him.DAT WQ yesterday said
dass er ihm vorbeibringt]?
that he.NOM him.DAT by.bring.3SG
'What (all) did Peter presumably tell him, yesterday, that he is bringing him?'

c.$ Was hat der Peter (alles) ihm (alles) vermutlich gestern gesagt, [$_{CP}$
what.ACC have.3SG the.NOM Peter WQ him.DAT WQ presumably yesterday said
dass er ihm vorbeibringt]?
that he.NOM him.DAT by.bring.3SG
'What (all) did Peter presumably tell him, yesterday, that he is bringing him?'

[31] Reis (1992) already states, essentially, the generalization in (63). However, the data provided in the paper were equally consistent with a generalization that included also A-traces into the distribution statement.

(63) *Distribution generalization for distal* alles, COMPLETE*:*
The distribution of distal *alles* corresponds to the distribution of $\bar{\text{A}}$-traces of its associate.

The possibility of a split as described by (63) follows from the *insider property* of the SST. However stranding is achieved, the procedure will involve the constituent that *alles* is a part of, or a sub-constituent of which *alles* is a part of and in a dependency with–the associate. From this privileged position, while not necessary, it is possible for *alles* to be sensitive to how the stranding procedure is effected. This sensitivity would need to be re-encoded into the DST on any version as far as I can see.

There are two kinds of facts that motivate (62) (and therefore (63)): facts based on raising, and facts based on scrambling. The raising facts cut across A- and $\bar{\text{A}}$-chains more clearly. Scrambling in German, on the other hand, is a domain that has received many different treatments in the literature (where it is characterized either as A, or as $\bar{\text{A}}$, or as something in between). The facts presented here, however, are not easily reconcilable with a (pure) treatment of German scrambling as an $\bar{\text{A}}$-transformation.

The data will all follow the same pattern. While we have seen so far that *alles* can be found in both A- and $\bar{\text{A}}$-*positions*, this section will show that *alles* can occur in an A-position only if the associate has $\bar{\text{A}}$-moved from there. Sentences that are only compatible with A-movement having applied from that position will prevent *alles* from occurring there. (64)-(65) visualizes the pattern: Each line in (64) is a close-in of four positions of a sentence, two A-positions and two $\bar{\text{A}}$-positions. The higher $\bar{\text{A}}$-position is occupied by the associate *wh*-phrase in each case, while *alles* occurs in a different positions in each example, twice in an A-position, and once in an $\bar{\text{A}}$-position.

(64) a. $[_{\bar{\text{A}}2}$ wh_1 $[_{\bar{\text{A}}1}$ $[_{\text{A}2}$ $[_{\text{A}1}$ alles_1]]]]
b. $[_{\bar{\text{A}}2}$ wh_1 $[_{\bar{\text{A}}1}$ $[_{\text{A}2}$ alles_1 $[_{\text{A}1}$]]]]
c. $[_{\bar{\text{A}}2}$ wh_1 $[_{\bar{\text{A}}1}$ alles_1 $[_{\text{A}2}$ $[_{\text{A}1}$]]]]

(65) now adds information about the derivation, in particular whether the *wh*-associate has A- or $\bar{\text{A}}$-moved from any of the lower positions. (65a,a′) are derivations involving only $\bar{\text{A}}$-movement: wh_1 $\bar{\text{A}}$-moves from A1 to $\bar{\text{A}}$2, through $\bar{\text{A}}$1. (65b,b′) are derivations that additionally include a step of A-movement: wh_1 first A-moves to A2, and only then $\bar{\text{A}}$-moves to $\bar{\text{A}}$2 through $\bar{\text{A}}$1. *Alles* is licit in its position in (65a,a′) and (65b′) because it occurs in a position from which $\bar{\text{A}}$-movement has applied. *Alles* is *illicit* in its position in (65b) because

Also, Fitzpatrick (2006) had independently come to the same conclusion for Japanese and Korean numeral quantifiers, and Russian non-agreeing quantifiers, also using Weak Cross Over and anaphor binding as diagnostics, lending cross-linguistic support to the naturalness of the pattern with *alles*.

the movement step that applied to wh_1 from that position was an instance of *A-movement*.

(65)

a.	$[_{\bar{A}2}$ wh$_1$	$[_{\bar{A}1}$ $\bar{A}$-t'_1	$[_{A2}$	$[_{A1}$ $\bar{A}$-t_1 alles	$]]]$
a′.	$[_{\bar{A}2}$ wh$_1$	$[_{\bar{A}1}$ $\bar{A}$-t'_1 alles$_1$	$[_{A2}$	$[_{A1}$ $\bar{A}$-t_1	$]]]$
b.*	$[_{\bar{A}2}$ wh$_1$	$[_{\bar{A}1}$ $\bar{A}$-t'_1	$[_{A2}$ $\bar{A}$-t_1	$[_{A1}$ **A-**t_1 alles$_1$	$]]]$
b′.	$[_{\bar{A}2}$ wh$_1$	$[_{\bar{A}1}$ $\bar{A}$-t'_1	$[_{A2}$ $\bar{A}$-t_1 alles$_1$	$[_{A1}$ A-t_1	$]]]$

The following sections are evidence for this state of affairs. In each case, independent properties of the sentence will force an intermediate step of A-movement – i.e. without the step of A-movement the sentence would be ungrammatical. While the sentence would be grammatical on its own, i.e. A-movement can successfully lead to a convergent derivation, adding *alles* in the trace-position of A-movement will cause the sentence to deteriorate significantly.[32]

[32] An important caveat arises in this context: why are sentences which seem to violate superiority, like (i), not significantly deteriorated, and in fact fully acceptable with the right intonation (e.g. with a pause after *wer*).

(i) Was$_2$ hat wer$_1$ alles$_2$ gekauft?
what.ACC have.3SG who.NOM WQ bought
'Who bought what all?'

The availability of scrambling is often implicated with the absence of superiority effects in German (e.g. **?**, but see also the discussion in Fanselow and Féry 2008, where the empirical facts are called into question and analyzed as artifacts): there are no superiority effects between clausemate *wh*-phrases, but there are superiority effects between non-clausemate *wh*-phrases; scrambling is clause-bound in German. If scrambling is a necessary step in a derivation of a superiority (non-) violating sentence as (i), then *alles* in (i) seems to be in the trace-position of scrambling, counter to the generalization that this section defends. There are two options to account for the acceptability of (i): either (a) the subject scrambled over *alles* and the scrambling-trace of the object, or (b) the scrambling step nullifying superiority is $\bar{A}$-movement, thus being able to strand *alles*. The two derivations are represented in (iia,b), respectively.

(ii) a. $[_{CP}$ obj$_2$ $[_{TP}$ sbj$_1$ $[_{vP}$ $t_{2,wh}$ alles $[_{vP}$ $t_{1,\Sigma}$ $[_{VP}$ $t_{2,\Sigma}$ $]]]]]$
b. $[_{CP}$ obj$_2$ $[_{vP}$ sbj$_1$ $[_{VP}$ $t_{wh,2}$ alles $]]]$

Option (b) seems a bad idea, generally, as it is not clear how the asymmetries shown to hold in this section would arise if the object could $\bar{A}$-scramble. As for option (a), there is some suggestive evidence that it might be on the right track. For one, (a) would predict that subjects that like to be in a higher, more topical position would resist *alles* from occurring to their left. This is shown to be true with a D-linked *wh*-phrase subject: it is worse than a simple *wh*-phrase when *alles* occurs to its left, cf. (iii). In addition, we might expect certain adverbs that occur in positions higher in the clause, which are less accessible for objects than for subjects (if not *in*accessible for objects) (cf. Merchant, 1996), to be able to appear between the subject *wh*-phrase and *alles*. Again, this seems to be correct, cf. (iv). The same relative contrasts hold between *wo* 'where' and a subject (e.g. *Wo hat wer alles einen Apfel gekauft?* 'Who bought an apple where?'), overall shifted somewhat into the less-acceptable spectrum. Note that for (a), as based on (i–iv), to be viable, it must also be the case that this scrambling step of the subject does *not* make the subject "superior" for the purposes of Superiority.

(iii) a. Wem$_2$ seine Kinder hat [welche der Lehrerinnen, über die wir
who.DAT his.ACC children have.3SG which.NOM the.GEN teachers about REL we

3.4.1 Raising

Note: Stefan Keine (personal communication) made me aware of the generalization in German that extraposed clauses are always CPs. This would make the relevant test cases in this section control rather than raising . For control, the explanation of the asymmetry *could* depend on A-movement – assuming the movement theory of control (**??**) –, but it *need not*. If PRO is assumed instead, then the unavailability of *alles* in the relevant sentences may simply follow from the fact that PRO is not a *wh*-phrase. Stay tuned for a different argument based on scope facts in favor of the same asymmetry in raising contexts.

Alles may be found in the matrix clause of raising verbs, irrespective of whether the associate is a subject or an object (the object one seems slightly less acceptable):

(66) Einfach nur unglaublich, ...
'Just unbelievable ...'

a. [$_{CP}$ was$_1$ alles$_1$ droht], [$_{TP}$ e_1 dem Max das Leben zu
what.NOM WQ threaten.3SG the.DAT Max the.ACC life to
versauen] !
spoil
'what all threatens to ruin Max's life!'

b. ?[$_{CP}$ was$_2$ [dieses Ergebnis]$_1$ alles$_2$ droht], [$_{TP}$ e_1 dem Max e_2
what.ACC this.NOM result WQ threaten.3SG the.DAT Max
zu versauen] !
to spoil
'what all this result threatens to ruin for Max!'

---

gesprochen haben,]$_1$ alles$_2$ verwöhnt?
taked have WQ spoiled
'Which of the teachers that we talked about spoiled the children of who all?'

a′. ??Wem$_2$ seine Kinder hat alles$_2$ [welche der Lehrerinnen, über die wir gesprochen haben,]$_1$ verwöhnt?

b. Wem$_2$ seine Kinder hat wer$_1$ alles$_2$ verwöhnt?
who.DAT his.ACC children have.3SG who.NOM WQ spoiled
'Who spoiled the children of who all?'

b′. Wem$_2$ seine Kinder hat alles$_2$ wer$_1$ verwöhnt?

(iv) a. Was$_2$ hat wer$_1$ wahrscheinlich alles$_2$ gekauft?
what.ACC have.3SG who.NOM probably WQ bought
'Who bought what all?'

a′. ??Was$_2$ hat wahrscheinlich wer$_1$ alles$_2$ gekauft?

b. ??Wer$_1$ hat was$_2$ alles$_2$ wahrscheinlich gekauft?
who.NOM have.3SG what.ACC WQ probably bought
'Who bought what all?'

b′. Wer$_1$ hat wahrscheinlich was$_2$ alles$_2$ gekauft?

Inside the raising complement, however, *alles* is acceptable only when the associate is the object, not if it is the subject. Compare (67b) with (67a), respectively.[33]

(67) Einfach nur unglaublich, ...
'Just unbelievable ...'

a. ?*[$_{CP}$ was$_1$ droht], [$_{TP}$ [dem Max]$_2$ e_1 alles$_1$ e_2 das Leben zu versauen] !
what.NOM threaten.3SG the.DAT Max WQ the.ACC life to spoil
'what all threatens to ruin Max's life!'

b. [$_{CP}$ was$_2$ [dieses Ergebnis]$_1$ droht], [$_{TP}$ e_1 dem Max e_2 alles$_2$ zu versauen] !
what.ACC this.NOM result threaten.3SG the.DAT Max WQ to spoil
'what all this result threatens to ruin for Max!'

The only difference between (68ab) is whether the *wh*-phrase had to raise into the matrix clause before *wh*-moving to matrix Spec,C (the subject), or whether it did not have to raise into the matrix clause prior to *wh*-movement (the object). Given that the subject had to

[33] These examples contain a dative object to ensure that there is a viable prosodic host to the left of *alles*, required by the fact that *alles* is always de-stressed. Therefore, given the conclusions in section 3.3, in (67a), the dative object must have scrambled over the subject position to appear to the left of *alles*, while no such additional movement is required in (67b). This fact alone, however, is unlikely to have caused the unacceptability of the sentence. For one, the dative object of *versauen* may scramble over an overt subject (ia). Giving the same over-the-subject-scrambling prosody as in (i) to (67a) does not improve the status of the sentence: to me, the sentence sounds like a good sentence, but the sentence meaning remains fuzzy. In addition, distal *alles* inside a control complement shows no asymmetry between being to the right of an object in base position, or being to the right of an object that has scrambled over it (iiab). (It does not matter here which is which; a slight intonational break is preferred before *alles* in (iia)–'(#)'.)

(i) dass dem Max die Maria das Leben versauen wollte.
that the.DAT Max the.NOM Maria the.ACC life spoil want.PST.3SG
'that Maria wanted to ruin Max's life.'

(ii) a. Wem$_1$ hat der Peter versucht, [$_{CP}$ Micheal Jackson (#) alles$_1$ vorzustellen]?
who.DAT have.3SG the.NOM Peter tried MJ.ACC WQ to.introduce
'Who all did Peter try to introduce Micheal Jackson to?'

b. Wen$_1$ hat der Peter versucht, [$_{CP}$ Micheal Jackson alles$_1$ vorzustellen]?
who.ACC have.3SG the.NOM Peter tried MJ.DAT WQ to.introduce
'Who all did Peter try to introduce to Micheal Jackson?'

raise out of the complement first (e.g. to get Case),[34] and raising is A-movement,[35] there must be an A-trace in the raising complement. Conversely, no *wh*-movement prior to raising may have occurred given Improper Movement. The object, on the other hand, can *wh*-move out of the raising complement, directly, and thus leave an $\bar{\text{A}}$-trace there. There seems to be no other difference. Therefore, the availability of *alles* inside the raising complement must be tied to the A- vs. $\bar{\text{A}}$-status of the trace left in the position where *alles* occurs.

As for *alles* in the matrix clause in (66), *alles* can occur there regardless of whether its associate is a subject or an object for two separate reasons: the subject *must* A-move into the matrix prior to *wh*-movement; the object *can scramble* into the matrix clause prior to *wh*-movement, as shown by the possibility to move the dative argument into the matrix clause:

(68) ?Einfach nur unglaublich, was$_1$ dieses Ergebnis [dem Max]$_2$
simply only unbelievable what.ACC this.NOM result the.DAT Max
DROHT, e_2 e_1 zu versauen!
threaten.3SG to spoil
'It's just unbelievable what this result threatens to ruin for Max!'

Finally, notice that the inflecting A-QF *all-* does not have the same asymmetric distribution. *Alle* can be found in matrix and embedded positions, both with subject and object associates; (69) shows the two possible subject floats, and (70) the ones for the object. The results with the object are slightly worse. Nonetheless, the asymmetry does not go in the same direction as with *alles* (inside the infinitival, inflecting *all-* is worse with the object, but *alles* is (much) worse with the subject), and the slight contrast does not impact interpretability in the same way as is the case with the low subject-*alles*, where the *alles*-specific interpretation becomes very difficult to impossible. What matters primarily is the fact that the strong degradation experienced with low subject-*alles* is not due to a general impossibility of floating a subject-quantifier inside a raising complement.

(69) a. Weißt du, [$_{\text{CP}}$ [$_{\text{DP}}$ wem seine Welpen]$_1$ {alle$_1$} drohen],
know.2SG you who.DAT his.NOM.PL puppies all.3PL threaten.3PL

[34] It does not matter what the 'trigger' is here. Raising must happen, as e.g., leaving the subject inside the complement, with an expletive subject in the matrix, yields ungrammatical results:

(i) a. *Es droht, dieses Ergebnis dem Max das Leben zu versauen.
it threaten.3SG this.NOM result the.DAT Max the.ACC life to spoil
Intended: 'This result threatens to ruin Max's life.'
b. *dass (es) t_i droht, [$_{\text{TP}}$ dieses Ergebnis dem Max das Leben zu versauen]$_i$.
that it threaten.3SG this.NOM result the.DAT Max the.ACC life to spoil
Intended: 'This result threatens to ruin Max's life.'

[35] For example, the derived subject controls agreement of the matrix clause.

[$_{TP}$ {alle$_1$} dem Max den Garten zu verwüsten]?
all.3PL the.DAT Max the.ACC garden to vandalize
'Do you know whose puppies (all) threaten to vandalize Max's garden?'

b. Ich weiß nicht, [$_{CP}$ [$_{DP}$ wessen Apps]$_1$ {alle} drohen], [$_{TP}$
I know.1SG not whose Apps.NOM all.3PL threaten.3PL
{alle} ein riesen Erfolg zu werden] (– deine sicher nicht).
all.3PL a.NOM giant success to become yours surely not
'I don't know whose Apps (all) threaten to become a huge success – yours clearly not.'

(70) Weißt du, [$_{CP}$ [$_{DP}$ wem seine Blumen]$_1$ dieser Welpe {?alle$_1$}
know.2SG you who.DAT his.NOM.PL flowers this.NOM puppy all
droht], [$_{TP}$ {?alle$_1$} zu verwüsten]?
threaten.3SG all.3PL to vandalize
'Do you know whose flowers this puppy threatens to (all) vandalize?'

3.4.2 Scrambling

The logic of the following three sections is as follows: when a derivation containing a step of scrambling over a certain constituent X is necessary for convergence, *alles* can be floated above the constituent X, but not below it. In addition, in two out of the three paradigms, scrambling is arguably A-movement given the kind of violation it obviates. The three paradigms are:

1. scrambling over a subject containing a bound pronoun (Weak Cross Over)
2. scrambling over a subject containing a reflexive (Condition A)
3. scrambling by an in-situ *wh*-phrase over an adverb

It should be noted that the judgments of this section are more subtle. Not surprisingly, the degree of contrast, and the range of acceptability in which the contrast arises (i.e. in the more-acceptable or less-acceptable range) varies across speakers. Nonetheless, the relevant contrasts arise reliably across speakers for Weak Cross Over and multiple *wh*-questions. Judgments for anaphor binding are only my own. Across all paradigms, stacking non-exhaustive *so* over *alles* makes judgments even clearer for myself.

Weak Cross Over Consider the following contrast in binding possibilities:

(71) *Intended:* 'Who are all the individuals x, such that x's teacher hit x?'

a. Wen$_i$ hat **alles** sein$_i$ Lehrer geschlagen?
who has WQ his teacher hit

b. ??Wen$_i$ hat sein$_i$ Lehrer **alles** geschlagen?
who has his teacher WQ hit

The intended reading is one where the possessive pronoun inside the subject is interpreted as bound by the *wh*-phrase. The availability of the bound pronoun interpretation depends on the position of *alles*. In (71a), *alles* marks a scrambling position, above the subject. In (71b), *alles* marks either the base position of its associate, or a scrambling position below the subject. In other words, *alles* above the subject is acceptable, but *alles* below the subject induces a Weak Cross Over violation. Weak Cross Over (WCO; Postal 1971; Wasow 1979) is essentially the phenomenon where a bound pronoun is not A-bound, and the sentence is degraded with the bound-pronoun interpretation. The specific type of WCO that is relevant in this section is where a possessive pronoun cannot, or can only with great difficulty, be understood as bound by an operator contained in an $\bar{\text{A}}$-dependency that spans over the bound pronoun. As a consequence, an English sentence as (72a) is generally judged as quite degraded. The configuration in (72b), in comparison, where both the head and the tail of the $\bar{\text{A}}$-dependency c-command the possessive pronoun, is perfectly acceptable (adapted from Safir, 2017: 1).

(72) a. ??Who$_1$ did [his$_1$ mother] praise t_1?
b. Who$_1$ t_1 praised [his$_1$ mother]?

Scrambling in German (and other languages) can obviate this effect (cf. Webelhuth, 1992). German (73a), corresponding to the degraded English (72a), is perfectly acceptable. The idea is that *wh*-movement of *wen* (with or without *alles*) in (73a) is preceded by a step of scrambling, as indicated in the corresponding (73b).

(73) a. Wen$_i$ (alles) hat sein$_i$ Lehrer geschlagen?
who WQ has his teacher hit
'Who (all) did his teacher hit?'
b. Wen$_i$ hat [$_{v\text{P}}$ t' [$_{v\text{P}}$ sein$_i$ Lehrer t geschlagen]]?

In addition, WCO cannot be obviated when the bound possessive pronoun is in a clause that cannot be reached by scrambling. In (74), the bound pronoun *seine* is contained in the matrix subject, while the *wh*-phrase originates in the embedded clause. Because scrambling is clause-bound in German, a long scrambling step preceding *wh*-movement is impossible. The *wh*-movement dependency will necessarily span over the pronoun that the operator binds, and induce a WCO effect.

(74) ??[$_{CP}$ Wen$_i$ dachte seine$_i$ Mutter, [$_{CP}$ dass der Lehrer t geschlagen hat]]?
who thought his mother that the teacher hit have.3SG
Intended: 'For which person x did x's mother think that the teacher hit x?'

Against this background, it seems that (71a) has the derivation in (75); Σ stands for scrambling:

(75) [WH [t_{wh} alles [SBJ [t_Σ]]]]

I take this to mean that the following derivation is available for (71a), making it possible to obviate WCO: *Wen* scrambles above the subject. From there, it A-binds the possessor, obviating WCO. Finally, *wen* *wh*-moves on to Spec,C. *Alles* thus occurs in a position corresponding to the *wh*-trace of *wen*.

If scrambling over the subject is what allows obviation of WCO, then, were (71b) acceptable, it would need to have the same derivation as (71a) in (75), differing only in the position of *alles*, as in (76).

(76) [WH [t_{wh} [SBJ [t_Σ alles]]]]

However, given that (71b) is unacceptable, (76) is not a possible derivation. In other words, *alles* cannot occur in the position corresponding to a scrambling trace. If this conclusion is correct, and *alles* must instead occur in the position of an $\bar{\text{A}}$-trace of its associate, we can begin to understand why (71b) is not acceptable: the conclusion leaves us, in essence, with three alternative derivation where the $\bar{\text{A}}$-trace requirement is satisfied while stranding *alles* below the subject—(77a), (77b), and (77b′).

(77) a. [WH [SBJ [t_{wh} alles]]] ⇒WCO
b. [WH [t_{wh} [t_Σ [SBJ [t_{wh} alles]]]]] ⇒IMPROPER MOVEMENT
b′. [WH [t_{wh} [SBJ [t_Σ [t_{wh} alles]]]]] ⇒IMPROPER MOVEMENT

All three derivations fail. (77a) has the right kind of trace in the position of *alles*, but it is missing a step of scrambling that would obviate WCO. The sentence can therefore not have the intended interpretation. (77b/b′) have the right kind of trace in the position where *alles* is pronounced, and there is a position reached by scrambling that c-commands the bound pronoun inside the subject, the position where the higher t_{wh} is. These derivations can therefore obviate WCO. However, (77b/b′) must be blocked: Both derivations interleave *wh*-movement and scrambling. If scrambling is A-movement, the derivations would constitute a case of Improper Movement, and would therefore prevent *alles* from appearing below the subject in a sentence like (71b).[36].

[36] It is in principle possible that the constituent *wen* first scrambles above the subject, followed by

If this conclusion is correct, then (this kind of) scrambling in fact *must* be A-movement, (contra, e.g. Grewendorf and Sabel, 1999), or else *alles* could be floated below the subject while also not incurring into a WCO violation. The distribution of *alles* must be restricted to positions where $\bar{A}$-traces are left. These conclusions are in line with the observation from English that it is indeed A-movement that obviates WCO, while $\bar{A}$-movement cannot. Raising is a typical example for this ((78a) adapted from Safir 2017): In (78a) *everyone* A-moves to the matrix subject position thus making the bound pronoun interpretation of *his* possible; in (78b), though there is *wh*-movement in the matrix clause, *wh*-movement of *who* is preceded by the same step of A-movement to the matrix subject position as in (78a), therefore making the bound pronoun interpretation of *his* available; (78b) thus contrasts with (78c) (=(72a)) where there is no A-movement step to a position where *his* could have been A-bound.

(78) a. Everyone$_i$ seems [to his$_i$ mother] [t to be a genius].
b. Who$_i$ seems [to his$_i$ mother] [t to be a genius].
c. ??Who$_1$ did [his$_1$ mother] praise t_1?

Anaphor binding **Note:** This section should be largely ignored but I leave it anyway: It seems that generally picture-noun reflexives are logophors and require animate antecedents. In fact, the test sentences are clearly unacceptable when judged out of the blue. It still seems to me that when the contexts are taken seriously, the contrasts arise. But judgments should therefore be taken with a (modestly large) grain of salt. Scrambling creates new A-binding possibilities. This is true also for binding of anaphora inside a subject. To avoid potential interfering factors (prominently logophoricity), the antecedent is made inanimate. The context is therefore somewhat "funny", though still plausible and coherent. Compare the facts in (80), given the context in (79). (The sentence in (80b) has a contrastive reading, as made clear by the context.)

(79) Context: *The furniture in the living room is painted completely white. During renovation work, pictures of the furniture were taken, but someone also made the pictures dirty by accident. Some pictures made some of the furniture dirty as they came into contact. The piece of furniture such that the object of the picture matched the piece of furniture that the picture made dirty is the table.* (As a question: For what x, x a piece of furniture, did a picture of x make x dirty?)

scrambling of the subject again over *wen*. WCO could then be obviated in the step preceding scrambling of the subject. Given the contrast found by native speakers, I take it that this derivation is, at least, less probable in some sense, such that the contrast arises at least as a matter of likelihood of parses of the string.

(80) a. *dass [ein Foto von sich$_i$] heute [den Tisch]$_i$ dreckig gemacht hat.
that a.NOM picture of itself today the.ACC table dirty made have.3SG

b. dass [den Tisch]$_i$ [ein Foto von sich$_i$] heute *e* dreckig gemacht hat.
that the.ACC table a.NOM picture of itself today dirty made have.3SG
Literal: 'that a picture of itself$_i$ made the table$_i$ dirty, today.'

Assuming that this headache-causing scenario can be kept clear in one's mind, the following contrast becomes of importance.

(81) a. [Welche Möbelstücke]$_i$ hat (alles) [ein Foto von sich$_i$] *e* dreckig gemacht?
which.ACC furniture.pieces have.3SG WQ a.NOM picture of itself dirty made
Literal: 'What piece of furniture$_i$ did a picture of itself$_i$ make dirty, today?'

b. [Welche Möbelstücke]$_i$ hat [ein Foto von sich$_i$] (??alles) dreckig gemacht.
which.ACC furniture.pieces have.3SG a.NOM picture of itself WQ dirty made

(81a) shows that *alles* may be floated above the subject, while (81b) shows that *alles* cannot be floated below the subject in this sentence. However, previous sections have established that an ACC-object related *alles* can, in fact, be floated in positions below the subject, both in a position reached via scrambling, and in the base position. The issue of (81b), then, is that the anaphor inside the subject must be A-bound (as, e.g., according to Condition A of the Binding Theory; Chomsky 1981). (81a), ignoring *alles*, shows that the anaphor may be bound as long as the string is consistent with an intermediate scrambling step given that scrambling creates new A-binding possibilities. *Alles* in (81a) thus marks the position which its associate would have targeted with an intermediate scrambling step in order to successfully A-bind the anaphor. From there, the associate *wh*-moves to Spec,C, leaving an $\bar{\text{A}}$-trace in the position occupied by *alles*. Conversely, in (81b), *alles* occupies a position from which *wh*-movement of its associate would leave the anaphor unbound, causing the sentence to be ungrammatical. The only alternative derivation is one where the associate scrambles to the same position targeted in (81a). This would either leave *alles* in a position occupied by a scrambling-trace, which we must then conclude is not licit. The derivation where *alles* occupies an $\bar{\text{A}}$-trace position, but there is nonetheless scrambling to a position from which the *wh*-phrase can A-bind the anaphor would, again, necessarily interleave scrambling and *wh*-movement in a way that violates Improper Movement. Scrambling must again be assumed to be an instance of A-movement in these derivations to explain the facts, lending in its turn

support to the generalization that the distribution of distal *alles* is restricted to the Ā-trace positions of its associate of a given derivation.

Scrambled *wh*-in-situ In multiple *wh*-questions, only one *wh*-phrase may move to Spec,C, leaving the other *wh*-phrase *in-situ*. An *in-situ wh*-phrase may, however, scramble, under certain circumstances. For instance, both (82a) and (82b) are acceptable (also as a pair-list reading).

(82) a. Wo hast du der Maria was gezeigt?
where have.2SG you.NOM the.DAT Maria what.ACC shown
'Where did you show what to Maria?'

b. Wo hast du was$_1$ der Maria e_1 gezeigt?
where have.2SG you.NOM what.ACC the.DAT Maria shown

However, in spite of the fact that the *in-situ wh*-phrase may appear in either position, *alles* may not occur in the pre-scrambling position in such paradigms. Compare (83a) with (83b).[37]

(83) a. Wo hast du was$_1$ alles$_1$ der Maria e_1 gezeigt?
where have.2SG you.NOM what.ACC WQ the.DAT Maria shown
'Where did you show what all to Maria?'

b. ??Wo hast du was$_1$ der Maria alles$_1$ gezeigt?
where have.2SG you.NOM what.ACC the.DAT Maria WQ shown

Similar contrasts hold for relative word order with adverbs: *alles* can occur right-adjacent to the scrambled ACC *wh*-phrase, cf. (84b), but it cannot not occur in the *wh*-phrase base position if the *wh*-phrase scrambled out of that position.

(84) a. Welcher Manager hat wen$_1$ heute (*alles$_1$) gefeuert?
which.NOM manager have.3SG who.ACC today WQ fired
'Which manager fired who (all) today?'

b. Welcher Manager hat wen$_1$ (alles$_1$) heute e_1 gefeuert?
which.NOM manager have.3SG who.ACC WQ today fired

Given that section 3.3 established that *alles* may in fact occur in base positions as well as scrambling positions, these contrasts show that the distribution is restricted by the kind of trace that is left by the associate in the position where *alles* occurs. When the trace is

[37] As a baseline, (i) is acceptable, such that the contrast is about distal *alles* rather than adjacent *alles*.

(i) Wo hast du der Maria was alles gezeigt?
where have.2SG you.NOM the.DAT Maria what.ACC WQ shown
'Where did you show what to Maria?'

an Ā-trace left by *wh*-movement, be it from a base position, or from a position previously reached by scrambling, *alles* may occur in that trace's position. When the trace is a trace left by scrambling, however, *alles* may not occur in that trace's position. This paradigm does not further distinguish scrambling as an A- or an Ā-dependency. Together with the other results of this section, it is nonetheless suggestive of the A- vs. Ā-distinction in the distribution of distal *alles*.

3.5 Licensing of *alles*: the local link property

The following generalization emerges concerning the structural description of where (distal) *alles* is licensed. (85) is a necessary but not sufficient condition. The *local link property* is directly entailed by the SST given that "stranding" necessarily makes reference to movement. The DST, on the other hand, does not directly entail this property.

(85) *Licensing generalization for distal* alles*:*
Distal *alles* must be c-commanded by a clausemate link of a suitable Ā-chain, in overt syntax.

There are essentially three parts to the licensing statement in (85): a direction of licensing, a local domain within which licensing must take place, and the kind of object that *alles* is in a licensing relation with. For the direction of licensing, *alles* must be c-commanded by a suitable object, rather than c-command it. The local domain is the clause, specifically CP. (This is an upper bound to the size of the local domain.) The kind of object that *alles* requires as an associate is a *suitable Ā-chain*. *Suitable* refers to the kinds of operators that *alles* can relate to: citing again Reis (1992), it is the class of operators that quantify over an indefinite variable, denoting an open set; see again section 2.2. *Ā-chain* is intended as an object that contains (a) an operator, and (b) all the traces that are created by Ā-movement and bound by the operator. A link of the chain is intended as a neutral term that refers equally to the *head* (the structurally highest instance), the *tail* (the structurally lowest instance), or an intermediate link of a chain. *Clausemate* means within the same CP. Finally, the licensing generalization is a generalization only if covert movement is ignored altogether for its purposes: distal *alles* must be a clausemate with a chain-link created in overt syntax (before spell-out to the interfaces in a Y-modular organization of Grammar) to rule out sentences where the relation between *alles* and its associate is licensed purely by where the quantifier and the *wh*-phrase take scope at LF.

Subsections 3.5.1-3.5.2 show that CP is the largest possible structure within which the licensing relation must hold, and that it is the operator that licenses *alles* by c-commanding

it.

3.5.1 Locality between *alles* and associate

Distal *alles* cannot relate to a *wh*-phrase that originates in a different clause. The relation cannot be established if, as in (86a), the question is completely contained in the matrix clause but *alles* is in an embedded clause, nor if, as in (86b), *alles* is in the matrix clause, but the question is fully contained in the embedded clause:

(86) a. *[$_{CP1}$ Wem$_1$ hat Peter e_1 erzählt, [$_{CP2}$ dass Maria alles$_1$ Susi
who.DAT have.3SG Peter told that Maria WQ Susi.DAT
geholfen hat]]?
helped have.3SG
Intended: 'Who all did Peter tell that Maria helped Susi?'

b. *[$_{CP1}$ Peter hat alles$_1$ gewusst, [$_{CP2}$ wen$_1$ Maria e_1 liebt]].
Peter have.3SG WQ known who.ACC Maria love.3SG
Intended: 'Peter knows who all Maria loves.'

However, note that the clausemateness need not be *overt*. In fact, as can be seen in (87), *alles* may occur in the embedded clause, with the associate in the matrix clause, as long as the base position of the associate is a clausemate of *alles*.

(87) [$_{CP1}$ Wem$_1$ hat Peter gemeint, [$_{CP2}$ dass Maria alles$_1$ e_1 geholfen hat]]?
who.DAT have.3SG Peter reckoned that Maria WQ helped have.3SG
'Who all did Peter say that Maria helped?'

In addition, *alles* is acceptable, if difficult, in intermediate clauses along the path of movement more generally:

(88) ?[$_{CP1}$ Wem$_1$ hat Peter gemeint, [$_{CP2}$ dass Rudi alles$_1$ dachte, [$_{CP3}$
who.DAT have.3SG Peter reckoned that Rudi WQ think.PST.3SG
dass Maria e_1 geholfen hat]]]?
that Maria helped have.3SG
'Who all did Peter say that Rudi thought that Maria helped?'

(88) therefore lends further support to the fact that distal *alles* must be an associate with a *chain-link* of its associate (and indirectly again to a successive-cyclic analysis of long *wh*-movement proceeding through *v*P in German such that intermediate *alles* is licensed).

3.5.2 Direction

In all the acceptable examples so far, *alles* was c-commanded in overt syntax by a chain-link of its associate. If this condition is not met, the string becomes unacceptable. Consider the

contrast in (89).

(89) a. [Wem alles] hat der Lehrer was alles gezeigt?
who.DAT WQ have.3SG the.NOM teacher what.ACC WQ shown
'Who all did the teacher show what all?'

b. *[Wem alles] hat der Lehrer alles was gezeigt?
who.DAT WQ have.3SG the.NOM teacher WQ what.ACC shown

In both examples, the fronted DAT *wh*-phrase ensures that the ACC *wh*-phrase *was* stays *in-situ*. (The presence of *alles* on both *wh*-phrases makes it easier to understand the complex question in the absence of a clear context.) In (89a), *alles* is to the right of *was*. *Was alles* presumably form a constituent, given also the most natural prosody, just as *wem alles* do in Spec,C. Though it is not clear at this point whether the associate c-commands *alles* when they form a constituent, placing *alles* in a position to the left of *wen* in (89b) causes unacceptability. Given that in the general case precedence maps to c-command in the German *middle field*, *alles* requires to be c-commanded by its unique associate.

The c-command requirement is also visible with associates that are embedded inside a DP. Neither of the two buried *wh*-phrases in (90) may act as an associate for *alles*, making the sentences unacceptable if *alles* is included.[38]

(90) a. [Den Vater [von welchen Kindern]$_1$] hat die Maria
the.ACC father of which.DAT.PL children have.3SG the.NOM Maria
(*alles$_1$) angezeigt?
WQ pressed.charge.against
'The father of which children did Maria press charges against?'

b. [Den Anwalt [von wem sein-en Kindern]$_1$] hat die Maria
the.ACC lawyer of who.DAT his-PL children have.3SG the.NOM Maria
(*alles$_1$) angezeigt?
WQ pressed.charges.against
'The lawyer of whose children did Maria press charges against?'

It is worth noticing, finally, that covert movement does not, at first approximation, contribute to the licensing of *alles*.[39] (86a) from the previous section, for instance, already lent support this observation. Even though both the *wh*-phrase and *alles* clearly scope over

[38] The large pied-piped DP in (90) is likely not fronted via *wh*-movement. Rather, its seems to be a topicalized DP that contains an echo- or quizmaster-like question. For instance, the particle *denn* cannot be added. Nonetheless, as mentioned previously, *alles* is compatible with echo- and quizmaster-like questions, too, such that this is not an interfering factor.

[39] Covert movement does not contribute in the positive sense, i.e. it cannot feed licensing of *alles*; it seems not to be sufficient for the licensing of *alles*. However, certain LF-configurations seem to negatively impact the availability of *alles*, such that *alles* is illicit in certain configurations where additional covert movement is required. See Beck (1996).

the whole proposition, this kind of scope, if analyzed as covert movement, cannot create final or intermediate LFs that would license *alles*—the sentence would be acceptable, otherwise. Similarly, though the associate does not c-command *alles* in (89b) and (90), both expressions must take scope over the whole proposition, such that in whatever order one does it, there will be a final or an intermediate LF in which the associate c-commands *alles*. Yet, this scope-taking is not able to license *alles*. This argument relies on some assumption – for instance that scope-taking maps onto LFs with c-command relations, and that the *wh*-expression in echo questions scopes over the whole proposition just as in regular *wh*-questions. If the assumptions are granted, then we must conclude that there is a condition on the licensing of *alles* that must be stated in narrow syntax before spell-out to the interfaces, be it first Merge or the course of the derivation.

4 Taking stock: SST vs. DST

Section 3 established that, except for the *form property*, all other properties that are entailed by the SST are displayed by *alles* in one way or another. The results are summarized once more:

(a) *Synonymy property:*
The same source for adjacent and floated quantifiers entails the same relation between them and their associate, and therefore the same meaning contribution.

⇒ Adjacent *alles* and distal *alles* make the same meaning contribution (section 3.1)

(b) *Uniqueness property:*
The same source for adjacent and floated quantifiers entails that (these) quantifiers cannot be introduced in other ways into the sentence and therefore a one-to-one correspondence between quantifiers and associates given.

⇒ There is a one-to-one correspondence between *alles*, whether adjacent or distal, and associates (section 3.2).

(c) *Stranding property:*
The same source for adjacent and floated quantifiers entails that the stranding procedure involves movement. If the quantifier may not be moved further by itself, the distribution of the associate bounds the distribution of the floated quantifiers.

⇒ Distal *alles* occurs in a (proper) subset of the positions that its associate may occupy independent of *alles* (section 3.3).

(d) *Insider property:*

However stranding is achieved, the procedure will affect the constituent that *alles* is a part of, or a sub-constituent of that constituent–the associate. The same source for adjacent and floated quantifiers entails that the quantifier can be sensitive to how the stranding procedure is effected.

⇒ Distal *alles* occurs in $\bar{\text{A}}$-trace positions of its associate but not in A-trace positions of its associate (section 3.4).

(e) *Local link property:*
The same source for adjacent and floated quantifiers entails that the stranding procedure involves movement, and that therefore locality restrictions that apply to a floated quantifier relative to its associate mirror locality restrictions on movement.

⇒ Distal *alles* must be c-commanded by a clausemate link of a suitable $\bar{\text{A}}$-chain, in overt syntax (section 3.5).

(f) *Early Link property:*
The same source for adjacent and floated quantifiers suggests that the relation between a floated quantifier and its associate is established in narrow syntax.

⇒ Distal *alles* cannot be licensed at LF; it must be licensed at least in narrow syntax (section 3.5).

The form property, too, presumably, can be included to the list given that *alles* is restricted to certain kinds of operators—a form of selection:

(g) *Form property:*
The same source for adjacent and distal quantifiers entails that the quantifier may select for a form of the associate, or that the associate may select for some form of the quantifier, and that the selection is the same for both adjacent and floating quantifiers.

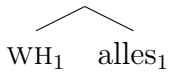
⇒ *Alles* can only associate with indefinite operators, in the sense of Reis (1992) (section 2.2).

The abundance of overlap of predictions of the SST and facts about *alles* is clear. It is equally clear then that, even if a version of the DST could capture the majority of facts by tailoring the analysis, unless the facts also followed from inherent properties of the analysis, the SST should be favored on conceptual grounds. This would be especially true if similar facts were true of other languages. In particular the fact that the A- vs. $\bar{\text{A}}$-trace split was found also for West Ulster English (McCloskey, 2000), Korean, Japanese, and Russian (Fitzpatrick, 2006), and the stranding property seems to hold in the same way in more broadly, in West Ulster English (Henry, 2012) and other varieties of English (see McCloskey, 2017 for an overview) support this conclusion.

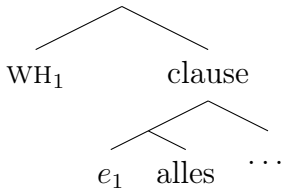
On the basis of these observations, I take it as settled that some version of the SST is to be chosen to explain the syntax of distal *alles*.[40]

What is not clear yet, however, is how the shared constituency in the common source gets to be actually (or seemingly) "undone" in cases of distal *alles*. In other words, how do we get from some structure like (91a), to some structure like (91b). The next section is devoted to narrowing down the answer.

(91) a.



b.



[40] For this reason I do not get into the details of why, specifically, which DST-analysis falls short, or on which data points. As a general case, the availability of *alles* in base positions is a big challenge for any DST analysis. The facts would have to be derived by massive remnant movement, or structure-preserving movement internal to the minimal projection containing the gap and the floated quantifier (as, e.g., proposed in Koopman, 2010; Heck and Himmelreich, 2017). Alternatively, the facts need to be accounted by assigning to the "adverb" *alles* the freedom of occurring in various projections, as is in fact common for adverbs more generally; this route is non-explanatory given that the distribution of *alles* varies with kind of associate (subject vs. object, etc.), and range of movement possibilities of the kind of associate.

Those facts aside, the uniqueness facts may partially follow under the plausible assumption that repetition of the same meaning leads to unacceptability due to redundancy (as in *#The smart student is smart.*). Specifically, this would account for why there cannot be multiple *alles* for one associate. However, the fact that one *alles* cannot relate to multiple associates is more challenging. The relation cannot be established from *alles* via Binding of the associate(s): Were it a Binding relation into the c-command domain of *alles*, two issues would arise. First, one would expect that *alles* may relate to multiple associates just as one QP can bind multiple pronouns, and in different clauses (e.g. *No-one$_i$ believed himself$_i$ more capable than his$_i$ mother believed him$_i$ to be.*, or *No-one$_i$ told [the friend$_j$ that he$_i$ doesn't like] that he$_i$ likes him$_j$*). Second, one would expect that *alles* may relate to associate in-situ objects, or associates in embedded clauses from the matrix clause. The locality is wrong. To avoid these issues, a DST analysis would have to adopt that the relation is an Agree relation, and that the associate must agree with *alles*. (If *alles* agreed with associates instead, the locality is fixed, but the question of why *alles* cannot multiply agree with, e.g., multiple associate clausemate objects remains unsettled: the only promising route on an Agree account is to resort to Minimality, but Minimality will undergenerate as it will predict that *alles* can only relate to the structurally highest potential associate–an incorrect prediction.) This endpoint is (part of) the analysis of Heck and Himmelreich (2017). The base position puzzle remains, and so does the A- vs. Ā-trace split. Certain LF-intervention effects, on the other hand, that Heck and Himmelreich derive, and that are the goal of their paper, will remain untouched in this paper. In that regard, it is worth pointing out that the LF-intervention effects seem to extend beyond just the indefinite DPs that Heck and Himmelreich address, and in fact include also negative DPs, sentential negation, and negative adverbs (e.g. *nie* 'never'). Each of these elements would have to be able to enter into an Agree relation (as the probing head assigning an index) with the adverb *alles*—a likely undesirable conclusion.

5 How (not) to float?

The remainder of the paper explores how floated *alles* is derived from a shared constituent with its associate. We know that the associate moves on when a float is created, but that in itself does not entail that the float is 'stranded' by sub-extraction as proposed by Sportiche (1988), and in a way already by Dougherty (1970); Kayne (1975) with Q-Post followed by movement of NP. The first decision in the *how to float* decision tree thus concerns (a) via movement, vs. (b) not via movement.

Within the copy theory of movement (Chomsky, 1993), a non-movement approach might consist in differentially deleting copies of a chain, as for example with *Distributed Deletion* (Fanselow and Ćavar, 2002), or following other chain-pronunciation computations, e.g. the one defended in Nunes (2004). This line of inquiry will not be pursued further in this paper. Rather, the following sections will be aimed at narrowing down a stranding analysis that is based on movement directly.

I will limit myself here to pointing out that an approach based on Distributed Deletion (DD) will, at least at first gaze, run into a serious issue when searching for an explanation of the generalization that *alles* is only floated in $\bar{\text{A}}$-trace positions. In a DD approach, there is no *stranding*, i.e. the whole constituent containing *alles* is moved at each movement step in the derivation. It is unclear then how DD could apply in such a way as to yield the A- vs. $\bar{\text{A}}$-trace difference in the distribution of distal *alles* given that DD presupposes the copy theory of movement. Postulating a difference between A- and $\bar{\text{A}}$-copies would pretty much lead back to trace theory, a contradiction. The only avenue that seems to be pursuable on such approaches is to limit "reconstruction", of meaning and of pronunciation to copies left by $\bar{\text{A}}$-movement. In other words, A-movement must be assumed not to leave any copies at all, and all A-copy related semantics must be computed and recorded as the derivation proceeds (as e.g. famously argued for Condition A by Belletti and Rizzi (1988).

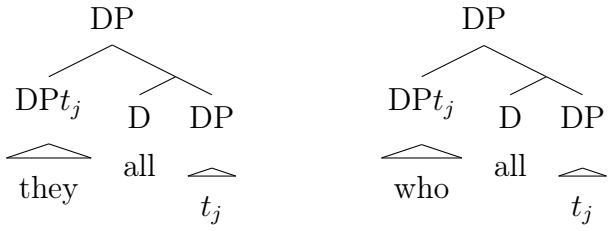
5.1 McCloskey's analysis

McCloskey (2000)'s analysis of West Ulster English (WUE) *wh*-quantifier float follows the SST-line of work (specifically Giusti 1990; Shlonsky 1991; Sportiche 1996; Merchant 1996). He further follows Postal (1974) and Koopman (1999) in assuming that pronoun-quantifier units of sentences as in (92ab) have the structure in (93a). Assuming that (simplex) *wh*-phrases are pronouns, McCloskey postulates the parallel structure in (93b) for WUE *wh-all* (McCloskey, 2000: 59).

(92) a. *What all* did you get t for Christmas? (McCloskey, 2000: 58)

b. *What* did you get *all* for Christmas?

(93) a. b.



In this analysis, the movement to the specifier creates a sub-constituent, the intermediate projection of the quantifier. This sub-constituent can be stranded by further movement of the *wh*-phrase associate. This kind of analysis captures the stranding property, in particular the occurrences of distal *alles* in base positions (just as it does for WUE). It can also capture the selectional relation between the quantifier and the operator, the uniqueness property, the locality facts and the early link property. In addition, as McCloskey discusses in section 7, if the first movement step to the escape hatch of the DP involves $\bar{\text{A}}$-movement, which seems plausible given that it continues to move via *wh*-movement from there on, the *wh*-phrase will not be able to A-move, and therefore not be able to strand *all* via A-movement.[41] McCloskey concludes that this must indeed be the case as there can be no available derivation where *all* is stranded in the VP-internal subject position via A-movement. He reasons that if such a derivation were available, then the contrast between floating *all* in an $\bar{\text{A}}$-context in (94a), and floating *all* in an A-context in (94b) could not exist within the same grammar. Yet, this is true in WUE as reported (labels and brackets added):[42]

(94) a. Who was throwing stones [$_{v\text{P}}$ all around Butchers' Gate?
b. *They were throwing stones [$_{v\text{P}}$ all around Butchers' Gate. (McCloskey, 2000: 77)

The reason is that at the point in the derivation where it must be chosen whether the DP containing *who all*/*they all* moves to TP (a) as a whole, or (b) only the pronoun, the two derivations in (94ab) are indistinguishable. McCloskey thus takes at face-value the fact that

[41] To be precise, McCloskey assumes that given that the first step of movement is connected to the *wh*-feature, the quantifier itself must bear this feature to trigger the movement to its specifier. Spec,Q will then be an $\bar{\text{A}}$-position because the Q has a *wh*-feature which it projects. It is not clear whether such an analysis would entail an A-over-A violation for sub-extraction, then, given that both the associate and the larger structure will bear the same feature when stranding in non-base positions. This issue seems to be a general issue for stranding approaches based on movement; it will be ignored here.

[42] McCloskey (2000) argues that there is overt object shift in West Ulster English: the object moves to a VP-external object position (Spec of Agr$_O$ or the head that introduces the external argument), followed by movement of the verb to yet a higher position. Thus *all* in (94) is in VP/*v*P-internal subject base position even though on the surface it follows the object *stones*. See McCloskey (2000: sections 6.3–6.4). McCloskey refers to 'the higher V' instead of the label '*v*P' that I choose for illustration in (94).

stranding via A-movement cannot be possible given the unacceptability of (94b): whatever makes that derivation unavailable would make the derivation of (94a) also unavailable, because if *all* is not stranded right away via A-movement to Spec,T, any subsequent *wh*-movement will come too late to strand *all* in the post-verbal subject position. Because, however, this is possible in (94a) McCloskey concludes that the stranding derivation in (94a) can (and must) proceed directly to CP in order to prevent an instance of Improper Movement, and that this is forced by *who* being in a DP-internal Ā-position.[43] In short, the same state of affairs seems to be in place with WUE *all* as with German *alles*: the associate may strand the quantifier via Ā-movement, but not via A-movement.[44] By adopting McCloskey's analysis we can then account for the stranding facts of *alles* in full.

5.2 The complexity puzzle

There remains, however, one last puzzle about *alles*, this time a puzzle concerning adjacent *alles*. It exhibits a complexity restriction, in the sense that it can occur with simplex *wh*-phrases but not with complex *wh*-phrases:[45]

(95) a. *[Wessen Freunde alles] möchten mir *e* beim Umzug helfen?
whose friends WQ want.3PL me.DAT by.the move help
Intended: 'The friends of who all want to help me move?'

b. *[Welche Teilnehmer alles] sollten wir *e* aufmuntern?
which.ACC.PL participants WQ should.3PL we.NOM cheer.up
Intended: 'What all (kinds of) participants do we need to cheer up?'

c. *[Wem seinen Studenten alles] soll ich die Lösung *e*
who.DAT his.DAT.PL students WQ shall.3SG I.NOM the.ACC solution
schicken?
send
Intended: 'The students of who all shall I send the solutions?'

[43] McCloskey (2000: 79f) therefore assume, as necessary, that the satisfaction of the EPP is not fully obligatory.

[44] There is a caveat, however, that figures earlier in McCloskey's paper. McCloskey argues that objects move to Agr_O for case, and that because of this movement *all* can be found in either of the two object positions in sentences as the one in (i): in the lower base position as in (ia) or in the derived object position as in (ib).

(i) a. What did you put all in the drawer.
b. ?What did you put in the drawer all. (McCloskey, 2000: 71)

If movement to a Case position is A-movement, as standardly assumed, then it seems that *all* can indeed be stranded in positions corresponding to A-traces of its associate. McCloskey discusses that these sentences are marginal, and even outright rejected by some speakers. This state of affairs is reminiscent of the speaker variation found with *alles* throughout section 3.4. It is an open question, then, what conclusion should be drawn from these contrasts, or whether WUE and German do in fact differ on this aspect.

[45] See also Reis (1992: fn23) for observations about this asymmetry.

Another way of saying it is that *alles* must be stranded when associating with complex *wh*-phrases.[46] The minimal pairs to (95) with floating *alles* are fully acceptable, similar examples of which have figured previously in the paper:

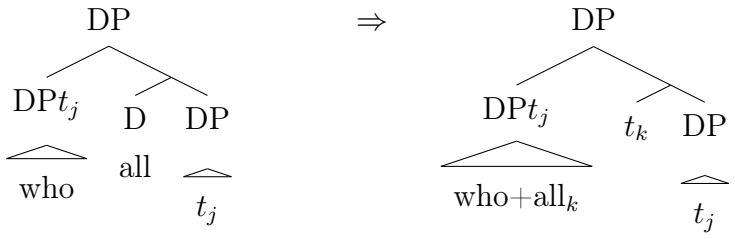
(96) a. [Wessen Freunde] möchten mir *e* alles beim Umzug helfen?
whose friends want.3PL me.DAT WQ by.the move help
'The friends of who all want to help me move?'

b. [Welche Teilnehmer] sollten wir *e* alles aufmuntern?
which.ACC.PL participants should.3PL we.NOM WQ cheer.up
'What all (kinds of) participants do we need to cheer up?'

c. [Wem seinen Studenten] soll ich die Lösung *e* alles schicken?
who.DAT his.DAT.PL students shall.3SG I.NOM the.ACC solution WQ
send
'The students of who all shall I send the solutions?'

What could yield this contrast on McCloskey's analysis? The quantifier is a D^0 in that analysis. A straightforward direction would be to posit that the left branches of the complex *wh*-phrases, which will contain a D^0 themselves, are somehow in competition with the quantifier. This approach would be too strong, however. The quantifier could not merge with a complex *wh*-phrase in the first place, thus preventing it from ever stranding it at any later stage of the derivation. Assigning some clitic properties to *alles* (cf. Reis, 1992) will also be too strong. If for instance this property is cashed in by saying that *alles* needs to m-merge with an adjacent *wh*-operator (similar in spirit to clitics in Matushansky (2006), where the PF cycle and narrow syntax get interleaved), as in (97ab), stranding would become an issue due to the impossibility of excorporation after m-merger.

(97) a. b.



The choice of m-merging must be made at some point in the derivation. If it is simply a choice to m-merge or not, then it is not clear why complex *wh*-phrases *have to* strand *alles*. If it is fully obligatory, then again complex *wh*-phrases would not be able to associate with

[46] It is of course an option to resort to multiple lexical entries, such that adjacent *alles* and distal *alles* are separate lexical items. This line of analysis seems untenable, or at least highly unlikely given the evidence in favor of assigning the same structural description to their insertion.

alles at all. It seems that the only other option is that m-merger is *obligatory if possible*, in some sense. But even if 'if possible' can be defined successfully, this option forces the undesired outcome that simplex *wh*-phrases will never be able to strand *alles* given that it will always be possible to m-merge with them.

The choice of pied-piping or not is also not conditioned by whether the associate moves (overtly) or not. The complexity split applies to complex *wh*-phrases in-situ just as it does to those in Spec,C:

(98) Possible answer: *Snowden showed someone XYZ, Ellsberg showed someone UVW, and 'John Doe' showed someone RST.*

a. [Wer alles] hat wem [was alles] gezeigt?!
who.NOM WQ have.3SG WND.DAT who.ACC WQ shown
'Who all showed what all to somebody?'

b. [Wer alles] hat wem [welche geheimen Dokumente (*alles)] gezeigt?!
who.NOM WQ have.3SG WND.DAT which.ACC secret documents WQ shown
'Who all showed what secret documents all to somebody?'

The contrast is also unlikely the product of a prosodic restriction given that it cannot be repaired by ellipsis. Common (and in fact often preferred) NP-ellipsis in sluicing of complex *wh*-phrases has two positive outcomes: it makes *alles* and the operator adjacent at the surface, and it dramatically reduces the absolute size of the prosodic unit. Nonetheless, the effect does not disappear, cf. (99b). *Alles* is not *per se* incompatible with *welch-* and an elided NP as can be seen from (100), where *welche alles* are string adjacent but *alles* is stranded.

(99) Die Presse hat Athleten aus aller Welt gepriesen...
the press have.3SG athletes from all.GEN world praised
'The press praised athletes from all over the world...'

a. Ich weiß aber nicht, wen alles (genau).
I know but not who.ACC WQ exactly
'...but I don't know who all.'

b. *Ich weiß aber nicht, welche (Athleten) alles (genau).
I know but not which.ACC.PL athletes WQ exactly
'...but I don't know what (athletes) all.'

(100) [Welche (Süssigkeiten)] alles PEter mag, bleibt ein Mysterium.
which.ACC.PL sweets WQ Peter.NOM like.3SG remains a mystery
'What all the kinds of sweets are that Peter likes, remains a mystery.'

Counting the number of syllables, so to exclude *we.ssen alles* and *wel.che alles* ($\sigma\sigma$ alles)– for instance because *alles* wants to be parsed into a foot while keeping its own syllables unstressed–, also seems hopeless. For one, the separable complex *wh*-phrase in (101) is at least marginally compatible with adjacent *alles*, and so are certain other (floating) expressions, as for example *genau* 'exactly' in (101b). In addition, this line of analysis would have to apply more broadly, without making reference to the syntactic label of its environment. A monosyllabic NP inside a complex *wh*-phrase cannot fix the problem, and neither does the word left-adjacent to *distal alles* have to be monosyllabic:[47]

(101) a. ?[Was für Leute alles] hat er damals bei dir treffen wollen?
what for people WQ have.3PL he.NOM then by you meet want
'What all sorts of people did he want to meet at your place back then?'

b. ... Ich weiß aber nicht, wen (genau) alles.
I know but not who.ACC exactly WQ
'... but I don't know what (athletes) all.'

(102) [Wessen *p*s/*p*-Wärte {*alles}] wurden {alles} gehackt?
whose *p*.PL/*p*-values WQ PASS.PST.3PL WQ hacked
'The *p*(-value)s of who all were hacked?'

Was..für NP is a separable expression, distinguishing its structure from the other complex *wh*-phrases' structures. In addition, it is possible to add *alles* to the simplex *wh*-phrase contained in a complex *wh*-phrase directly:

(103) a. ?[Was alles für Leute] hat er damals bei dir treffen wollen?
what WQ for people have.3PL he.NOM then by you meet want
'What (all) sorts of people did he want to meet at your place back then?' It seems appealing, then, to assume that the restriction is of structural nature.

b. ?[Wem alles seinen Studenten] soll ich die Lösung *e*
who.DAT WQ his.DAT.PL students shall.3SG I.NOM the.ACC solution
schicken?
send
'The students of who all shall I send the solutions?'

It seems appealing then to turn again to structural analyses of the restriction, while keeping in mind the difficulties faced above, to which one final complication must be added: The restriction appears to really be about the shape of the operator, and not about the size of the associate. Pied-piped PPs may be associates of *alles* just as DPs, cf. (104). PPs are structurally more complex than just a simplex *wh*-phrase, such that it might be predicted

[47] An IPA transcription of *wurden* in (102) is ['vuɐ̯.dn̩], to be compared to *wessen* ['vɛ.sn̩].

that any PP would require *alles* to be stranded. However, the simplex-complex distinction continues to be about the *wh*-phrase even inside the PP; compare the boldfaced position in (104c) with the italicized ones in (104ab). The mini-typology is summarized in (105).

(104) a. [Wo {*alles*}] kann man hier *e* {alles} gut essen?
where WQ can.3SG one here WQ good eat
'Where all can one eat well around here?'

b. [In was {*alles*}] muss man den *e* {alles}
in what.ACC WQ must.3SG one.NOM that.ACC.M.SG WQ
eintunken?
dip.in
'In what all must one dip that in?'

c. [In welchen Restaurants {***alles**}] kann man hier *e* {alles} gut
in which.DAT.PL restaurants WQ can.3SG one here WQ good
essen?
eat
'In what all (kinds of) restaurants can one eat well around here?'

(105) *Complexity typology:*

CONTEXT:	DP	PP
[SIMPLEX WH *alles*]	OK	OK
[COMPLEX WH *alles*]	*	*

Let us take stock of what the complexity split might be about by recapitulating what it is *not* about. Overall, the restriction could be a *narrow syntax* effect, or it could be an *interface* requirement, either of *LF* or of *PF*. We have excluded lexical competition, selection, cliticization, surface-adjacency, and prosodic structure from the options of narrow syntax and PF. Three more options will be explored in the next three sections. The first is, for lack of better understanding, a narrow syntactic effect. *Alles* is assumed to require sisterhood with a *wh*-pronoun, where '*wh*-pronoun' is understood as either a simplex *wh*-phrase or a *wh*-trace. The second option attempts to derive the first attempt as an LF condition – specifically a requirement that *alles* be sister to a variable. The third option is an attempt to derive the first attempt from a PF/Spell-Out condition, namely a condition on linearization/labeling, or a surface filter similar to restrictions on multiple-*wh* clusters in Serbo-Croatian and Romanian or the Double-*ing* Filter.

5.3 *Wh*-pronoun toy theory

Based on the discussion of the complexity split so far, we can come to a new generalization for the overall distribution of *alles*. Adjacent *alles* must be adjacent to a simplex-*wh*, call it

a *wh*-pronoun. Distal *alles* must be adjacent to an Ā-trace of its associate. Right-adjacency to *wh*-pronouns and Ā-traces then captures the total distribution of *alles*. We may thus, as a theory of first approximation, postulate that Ā-traces are essentially *wh*-pronouns.[48,49] To implement this idea, let us next assume that *alles* selects for *wh*-pronouns, but that this requirement can be satisfied at any stage of the derivation. It follows that (i) *alles* may be the sister of *wh*-pronouns both bare and inside of PPs, cf. (106), and that (ii) *alles* can be stranded by PPs, with simplex or complex *wh*-phrases inside, and complex *wh*-phrase DPs, because both of them will leave a "*wh*-pronoun" behind such that the requirement of *alles* can be satisfied at some point in the derivation, cf. (107). Note the scare-quotes here as *wh*-traces are being treated not at actual *wh*-pronouns, but as categories that are treated by the relevant part of syntactic computation as the same as *wh*-pronouns. Otherwise, it would need to be assumed that *wh*-traces of PPs are *wh*-pronouns, too—while this assumption may seem innocuous for some PPs, e.g. locations, for which pronouns do exist, the assumption should seem highly suspicious for other PPs, e.g. comitatives, for which there are no equivalent pronouns in German. Notice also that it must be possible for *alles* to merge with either an associate inside a PP, or the PP directly: Stranding of *alles* with PPs would not be possible without P-stranding, otherwise, which is disallowed in German.[50]

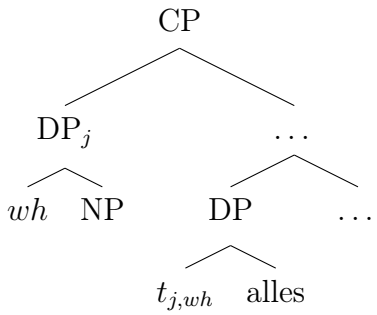
(106) a. b.

[48] This idea, its potential consequences for reconstruction, and its potential implications for the copy theory of movement are due to Jeff Lidz (p.c.).

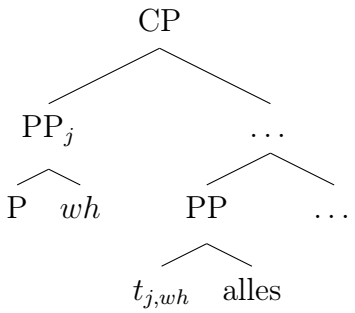
[49] The term 'pronoun' in '*wh*-pronoun' is not to be understood in the sense of, e.g., Binding Theory as discussed in Chomsky (1981). In binding theoretic terms they are rather to be understood as R-expressions, as often assumed for the lowest Ā-trace of a chain.

[50] Most versions of a DST analysis will also need to for (part of) the dependency to hold between *alles* and the PP directly. See footnote 40.

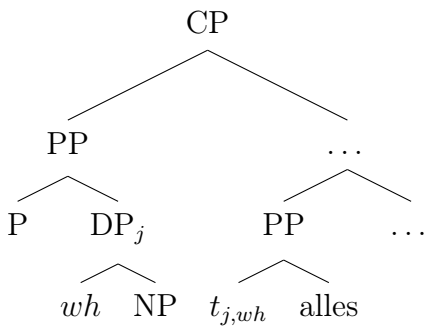
(107) a.



b.



c.



This toy theory derives the complexity split, all stranding cases, and the ban on stranding via A-movement (under the plausible assumption that A-movement does not leave *wh*-traces behind, even when it applies to a *wh*-phrase). The restriction is essentially a quirky lexical property of *alles* at this point, so it is essential that this theory make some testable predictions and that these hold true.

One prediction this toy theory makes, given that the restriction is lexical, is that it is idiosyncratic of the element itself, and that other similar elements do not have this restriction. Reis (1992) names three other elements as part of the class of "quantifying particles" as she calls them: *genau* 'exactly', *zum Beispiel* 'for example', and *so alles*, the combination of the non-exhaustive particle *so*, which has a plurality presupposition like the one of *alles*, and *alles*.[51] All three can occur in Spec,C with a simplex *wh*, cf. (108); *genau* and *zum Beispiel*

[51] Non-exhaustive *so*, as discussed in Reis (1992), while displaying many of the characteristics of *alles*, and perhaps all, is not a clean comparison given that it can also occur with seemingly the same meaning contribution in declarative sentences, i.e. in the absence of a specific operator. Thanks to Elena Herburger (p.c.) for bringing this difference to my attention:

(i) a. Ich forsche so an schwebenden Quantoren.
I research SO at floating quantifiers
Approximately: 'I kinda do a bunch of research on floating quantifiers.'

b. Ich hab' mir so gedacht, dass das ganz gut sein könnte.
I have me SO thought that that whole good be could
Approximately: 'I kinda thought about it a bunch and concluded that that could be pretty good.'

can occur with a complex associate, but *so alles* cannot, cf. (109).

(108) a. [Wem genau] wolltest du das geben?
who.DAT exactly want.PST.2SG you.NOM that.ACC give
'Who exactly did you want to give that to?'

b. [Wem zum Beispiel] wolltest du das nicht geben?
who.DAT for example want.PST.2SG you.NOM that.ACC not give
'Who, for example, did you not want to give that to?'

c. ?[Wem so alles] hast du das gegeben?
who.DAT SO WQ have.2SG you.NOM that.ACC given
'Who are some of the people that you gave that to?'

(109) a. ?[Wem seinen Kindern genau] wolltest du das
who.DAT his.DAT.PL children exactly want.PST.2SG you.NOM that.ACC
geben?
give
'Whose children exactly did you want to give that to?'

b. ?[Wem seinen Kindern zum Beispiel] wolltest du das
who.DAT his.DAT.PL children for example want.PST.2SG you.NOM that.ACC
nicht geben?
not give
'Whose children, for example, did you not want to give that to?'

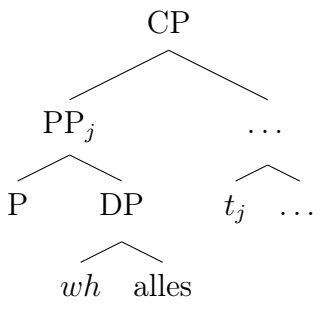
c. ?*[Wem seinen Kindern so alles] hast du das gegeben?
who.DAT his.DAT.PL children SO WQ have.2SG you.NOM that.ACC given
'Who are some of the children's parents that you gave that to?'

The variation seems to support the *wh*-pronoun theory: *genau* and *zum Beispiel* have the same syntax except that they do not further require to be sisters to a *wh*-pronoun at some point in the derivation.

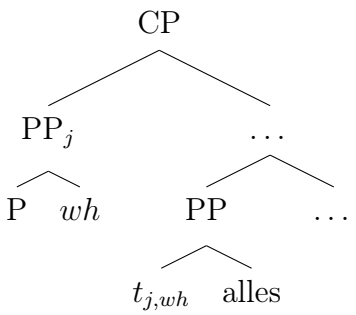
Another consequence is that the starting structure for PPs with simplex *wh*s that strand *alles*, is different from the starting structure for PPs that don't. When *alles* is not stranded, it must be the sister of the simplex *wh*, but when it is stranded it must have started out as the sister of the PP so that the PP may move out; compare (110ab). It is unclear, however, what the predictions are that are tied to this distinction.

(110) *Two different starting configuration for the same association with PPs:*

a.



b.



The third consequence is that under certain assumptions *alles* would allow for reconstruction of its associate only when the associate is a simplex *wh*. If one assumes that the requirement must stay satisfied once it is satisfied, simplex *wh*s may reconstruct because *wh*-pronouns can themselves satisfy the *alles*. Complex *wh*s, however, and PPs that stranded *alles*, cannot reconstruct to the trace position that is sister to *alles*. Reconstructing there would cause the requirement to go unsatisfied again. Testing this prediction could involve testing Condition A and Condition C effects with picture-NPs. The judgments are not trivial and require more controlled investigation. If they turned out to be correct, however, they would still face some important conceptual challenges. For reconstruction effects to affect or be affected by the form requirement of *alles*, the requirement must be an LF-requirement: it cannot be a first-merge requirement, as that had proven to be too strong, and if it is a narrow syntax (or even S-Structure) requirement, then what is forcing it to stay "active" once it is satisfied? If it is an LF-requirement then it becomes clear how it may be bled by reconstruction. However, what is the relevant vocabulary to restate the form restriction as an LF-restriction, as its current statement essentially conflates operators and variables?[52]

If these issues can be overcome, and the predictions are borne out, it seems that this theory would constitute some push-back against the Copy Theory of Movement (Chomsky, 1993) and the Inclusiveness Condition (Chomsky, 1995) given that the analysis rests on the idea that the syntactic process of *wh*-movement may create elements that are identical to simplex *wh*.

I turn to two alternatives in the last section, where two different proposals are made for how the starting point of the *wh*-pronoun toy theory may be derived. The starting point is that *alles* is

(a) the sister of a (*wh*-)pronoun, *or*

[52] Furthermore, if the form requirement is an LF-restriction we risk over-generation in cases of LF-movement, for instance multiple *wh*-questions where the in-situ *wh* is a complex *wh*, or a PP. In those cases, it would have to be the case that only the operator itself moves, so that such movement may not satisfy the LF-restriction. This distinction will become relevant again in section 5.4.

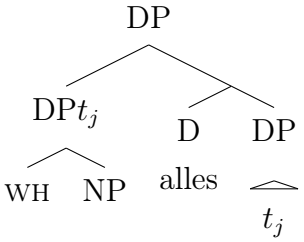
(b) the sister of a (*wh*-)trace.

The *wh*-pronoun toy theory unified the disjunctive statement by postulating the identity of simplex *wh*-phrases and *wh*-traces as both being *wh*-pronouns. The first alternative to this unification tries to understand the restriction as an LF condition, and unifies the disjunction by treating *wh*-pronouns and *wh*-traces as variables alike. The second alternative tries to understand the restriction as a PF/Spell-Out condition on linearization, and resolves the disjunction by pointing to the existence of two separate mechanisms to avoid certain un-linearizable/un-labelable structures, namely Movement and Projection.

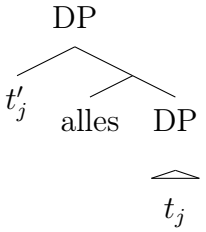
5.4 LF: Operators sisters of variables

The complexity split allowed us to restate the distributional statement of *alles* as occurring either right-adjacent to a simplex *wh*, or right-adjacent to a *wh*-trace. We can also conclude that, at least in the case of right-adjacency to the trace, right-adjacency must be sisterhood, and that movement to the specifier position of that projection must be blocked. In fact, if the configuration in (111a) could license *alles* it would be unclear why the complex *wh*-phrase couldn't stay in the specifier when it doesn't have to move on overtly, e.g. when it stays in-situ in multiple *wh*-questions. By the same token, we must also exclude the notion that *alles* could be licensed by a *wh*-trace in its specifier, as in (111b). This means that McCloskey's analysis for WUE, as previewed in section 5.2, cannot be extended to *alles*.

(111) a.

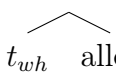


b.

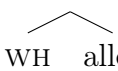


We are left with the configuration in (112a) as the sole licensing configuration for *alles* with complex *wh*-phrases; the argument extends to PPs with stranded *alles*. We may thus assume that a parallel configuration must hold for *alles* licensed by simplex *wh*-phrases: (112b).

(112) a. DP/PP



b. DP



We would thus have a single structural configuration, but there is still a disjunction between the two elements that are sisters of *alles*. Semantically, the *wh*-expression in (112a) is an

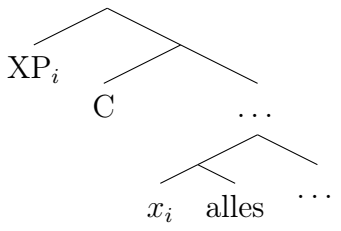
operator, and the trace in (112b) is a variable when corresponding to the base position. We would solve the disjunction if both simplex *wh*-phrases and *wh*-traces were an operator, or if both were a variable. As Howard Lasnik (p.c.) points out, the second route would further require that intermediate traces, too, act semantically as variables. Both options seem undesirable, such that the LF branch does not seem like a good avenue to derive the complexity split for now, but I will nonetheless explore some of the consequences in this section.

The first option, that both are operators, seems unlikely given that operators require to have a variable that they can bind, i.e. there is no vacuous quantification (in natural language).[53] The other option, that both the simplex *wh*-expression and the *wh*-trace are variables, seems approachable at first. If both the *wh*-expression itself and its traces are variables, we can understand *alles* as an operator. Then, the complexity restriction may be understood as a condition on *alles*, namely that it requires to be the sister of the variable it binds/quantifies over, as stated in (113). The *wh*-operator that is usually assumed to correspond directly to the *wh*-phrase must be somewhere else in the structure in this approach, e.g. in C. From there, the *wh*-operator can c-command both the complement domain and the specifier given Bare Phrase Structure (cf. Chomsky, 1995), and bind the variable(s) occurring in the *wh*-chain.

(113) *LF condition on* alles*:*
Alles must be sister to a variable.

The LF condition stated in (113) has the benefit that whenever the associate strands *alles*, *alles* becomes sister to the trace of that associate; in terms of the copy theory of movement, *alles* becomes the sister of a copy that is interpreted as a variable. All cases where *alles* is stranded by an associate follow for the same configuration illustrated in (114) (x is a shorthand for a variable, or a copy interpreted as a variable, and XP can be a simplex *wh*, a complex *wh*, or a PP with either a simplex or a complex *wh*).

(114)



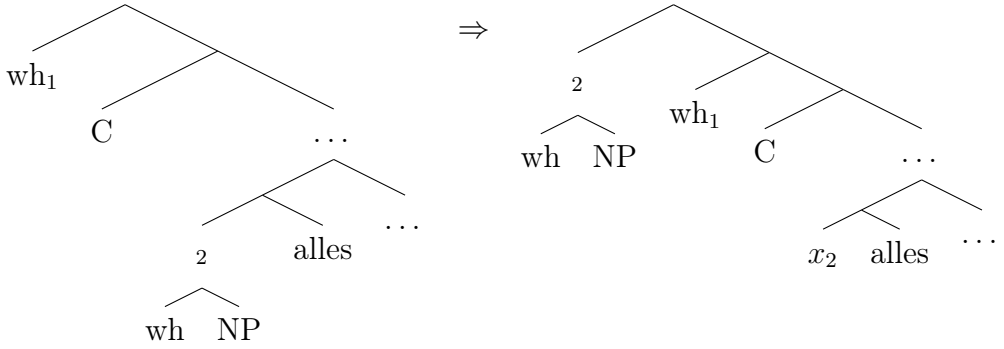
[53] As given, e.g., by the principle of Full Interpretation, and blocking examples such as (i) (Chomsky, 1986b: 99).

(i) *Who did John like Bill?

For the cases where *alles* is not stranded, there are two cases to consider: first, those where *alles* occurs in an in-situ *wh*-phrase of multiple questions, and second, those where *alles* is pied-piped by the *wh*-phrase, to Spec,C or anywhere else.

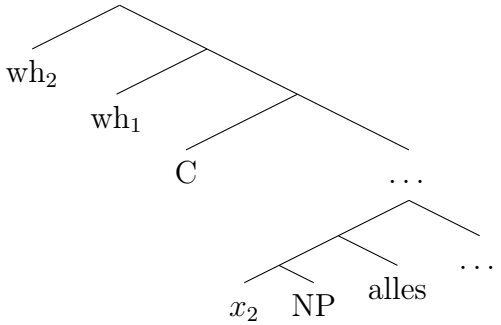
Under the assumption that in-situ *wh*-phrases of multiple questions move covertly to take scope over the proposition at LF, we would expect this kind of movement to be able to feed the condition in (113) given that (113) is an LF condition. For simplex *wh*-phrases this assumption is inconsequential. For complex *wh*-phrases (in a PP or not), this assumption is potentially problematic. If the whole XP sister to *alles* may move covertly, configurations like (115a) would be expected to converge, yielding e.g. (115b), when the respective sentences are in fact unacceptable.

(115) a. b.



This issue is resolved if covert movement does not pied-pipe, i.e. if only the variable moves, such that (115a) yields (116) instead of (115b). *Alles* is then no longer sister to a variable, and its LF condition is not met; the same is true of complex *wh*-phrases inside of PPs. Simplex *wh*-phrases inside a PP, on the other hand, are always trivially satisfied as long as *alles* is merged sister to the *wh* rather than sister to the PP. This is the desired outcome given that respective sentences are acceptable.

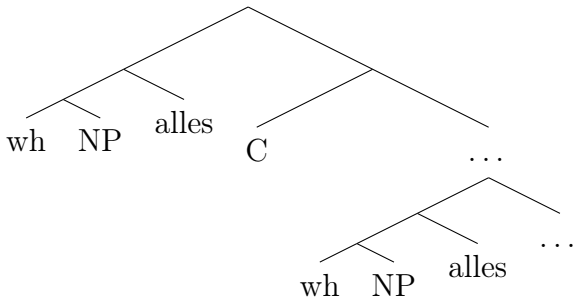
(116) ⇒



Let's turn now to pied-piped *alles*, i.e. *alles* that moves along with complex *wh*-phrase associates, for instance the German equivalent of (117).

(117) [What European ministers all] support the White House's policies on Venezuela?

(118)

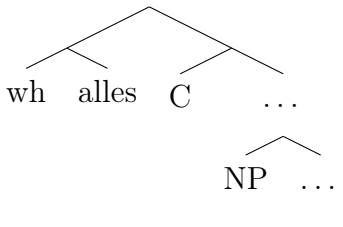


The issue here is the inverse one of covert movement going up the tree, namely "reconstruction". Whatever "reconstruction" is, let's understand it as the marking of 'which part of a chain is interpreted in which link of the chain' as, for instance, in Chomsky (1995), or as discussed in Uriagereka (2012: chapter 1.3).[54] For instance, the syntactic object in (119a) can undergo reconstruction as in (119b) such that the restrictor NP is interpreted in the lower copy, and the operator (on the traditional interpretation) in the higher copy:

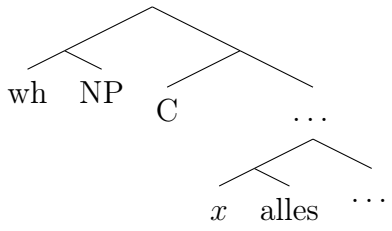
(119) a. $[_{\text{CP}}$ [which European ministers] $[_{\text{TP}}$ [which European ministers] $[_{\text{VP}} \ldots]]]$
b. $[_{\text{CP}}$ [which] $[_{\text{TP}}$ [European ministers] $[_{\text{VP}}$ support $\ldots]]]$

With this in mind, let's turn again to complex *wh*-associates. Given that we have been assuming that both *wh*-expressions *and* traces/lower copies are variables, and that *alles* could be licensed in the stranded position where it is sister to a variable, it is necessary to block reconstruction from feeding the LF condition in (113). Were it possible for reconstruction to feed (113), nothing could prevent derivations like those in (120ab) from being licit continuations to (118).

(120) a.



b.



In (120a), *alles* is licensed because reconstruction of the restrictor NP made *alles* sister of the variable–*wh*. In (120b), *alles* is reconstructed directly to a lower position. From here, given that *alles* was assumed to be satisfiable in such a configuration, it is unclear what would prevent *alles* from being licensed this way. Either way, *alles* pied-piped by complex *wh*-associates is expected to yield acceptable sentences if reconstruction can feed the LF

[54] The main difference here is that, given that *wh*-expressions are variables, there is, a priori, more freedom of what can be reconstructed where (as opposed to being limited to reconstructing the restrictor but not the operator, e.g.).

condition on *alles*. It must be concluded that reconstruction can *not*, in whatever way possible, feed the condition in (113).

Overall, this route is capable of deriving the complexity split while also doing away with its disjunctive generalization. The cost, however, is a rather construction specific condition that cannot be fed by processes of the same domain of computation. Of course it is possible that there is principled reason for why reconstruction cannot feed the condition, but the motivation of such a reason is this account's burden to find. In addition, if reconstruction is not able to feed the condition, we would equally expect it not to be able to bleed it. The cost would be that the predictions that were discussed in the previous section for the *wh*-pronoun toy theory would not apply to the supposed derivation of that toy theory. Finally, there are some conceptual limitations of this approach. Specific to this account is the limitation that intermediate copies/traces must be able to be variables. As Howard Lasnik (p.c.) points out, we should be suspicious of this assumption from a stand-point of logic. Instead of having a function $f(x)$, we would end up with a function $f(x, x, x, x, \ldots)$ with as many xs as members of the chain that end up as variables. At first stroke it appears to me that depending on the semantics one chooses, the multiplicity of variables can either be innocuously redundant, as perhaps in systems based on conjunction, or it might lead to the wrong semantics entirely, as perhaps in systems based on function application; clearly more must be said here. Another issue specific to this account, as pointed out by Juan Uriagereka (p.c.), is that it would lead to higher order quantification. Both *alles* and the *wh*-operator (in C, e.g.) would be quantifying over the same variable. While this is not impossible, it is added complexity. The issue could be resolved if *alles* were instead a predicate, whose semantics is a set of instructions for how answers should be picked from the domain over which the *wh*-expression quantifies, e.g. adding the restriction that the domain be partitioned in a certain way as to give a list of answers. The next section will indeed assume that *alles* functions as a predicate. Finally, there are some broader issues with an LF approach to the complexity restriction. As Jeff Lidz (p.c.) points out, at first stroke it seems suspicious to have the meaning component drive the choice of stranding *alles* or not, and just in case the *wh*-phrase is complex, in particular given that floated and adjacent *alles* make the same meaning contribution. Perhaps part of the reasoning of this section can instead be applied to an analysis of the complexity restriction that is not based on stranding as sub-extraction, but instead on stranding as "reconstruction" at PF, or more generally the procedure responsible for the pronunciation and non-pronunciation of (parts of) copies. See the preamble of section 5 for some discussion.

5.5 PF conditions and filters

5.5.1 PF Linearization condition/Labeling condition

In this section we turn to the PF side or Spell-Out more broadly, for an alternative interface motivation for the complexity split and the disjoint distribution statement that it left us with. The motivation will be based on the need for linearization, or interpretability more broadly, and the idea that simplex *wh*-pronouns and *wh*-traces are both special in that they trivialize the mapping from syntactic structure to the interfaces.

Consider again the disjunctive distribution found for *alles*: it is either the sister of a *wh*-trace, or the sister of a simplex *wh*-pronoun (=(112)):

(121) a. DP/PP b. DP

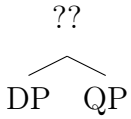
This section proposes that the disjunctive character of (121) is not to be resolved directly, but that it is rather best understood as a disjunction between two mechanisms that can resolve a structure that does not meet interface requirements. In particular, there are two ways to understand it, both resorting to the following two mechanisms: projection, and movement. The first is based on linearization, a PF condition, essentially stating that symmetric XP-YP structures cannot be linearized. The structure *can* be linearized, however, if one of the two elements can move in a way akin to Moro's *Dynamic Antisymmetry* (Moro, 2000). In addition, we may add that projecting, i.e. functioning as a head, is a second way to avoid the tension. The second is based on *labeling* in Ott's *local instability* (Ott, 2011, 2012). The architecture of grammar is such that Merge applies freely and does not by itself lead to labeling/projection. Rather, it is an independent procedure that establishes labels of syntactic objects. As an interface condition, syntactic structures must be labeled, with essentially two ways to establish a label: again, movement, and "projection" (this time in scare-quotes as there is no 'projection mechanism' by itself or as a result of Merge in this system, but rather it is an outcome of the labeling procedure).

The two approaches lend themselves equally to an analysis of *alles*. I will remain agnostic here. Getting more concrete, Ott (2012) proposes in chapter 4 that German quantifier float (of the inflecting A-variety) is best analyzed as one of the topic split constructions that motivated his system. He proposes the following symmetric base structure, as for any other of these constructions.

(122) {DP, QP} (Ott, 2012: 154)

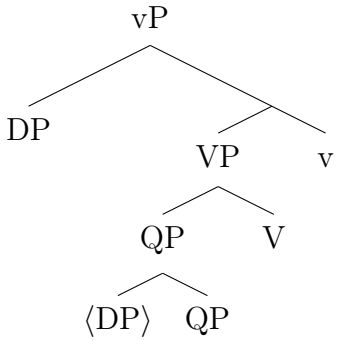
The QP is thus an XP, just as its associate.[55] Given its XP status, the predicate that is thus formed must be broken up. Essentially, (122) would lead to a point in the derivation (at the interfaces) where the question marks as illustrated in (123) would cause a crash.

(123)



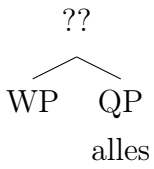
Movement of one of the constituents fixes the tension. Ott argues that the DP moves out, cf. (124) (Ott, 2012: 155).

(124)



Returning to *alles*, the same movement resolution can be adopted to derive the fact that complex *wh*-associates require *alles* to be stranded. Specifically, the base structure in (125) can be assigned to associate-*alles* pairs, where WP stands for any licit associate:

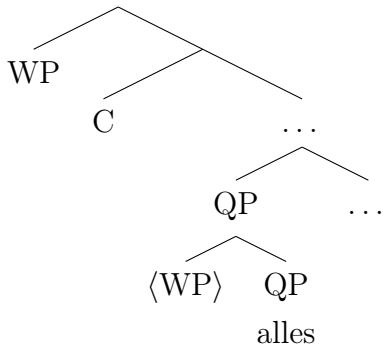
(125) *Symmetric source for* alles*:*



(125) requires movement to resolve the *??*s, be it due to labeling or due to linearization. This fact derives stranding *alles* with any associate. In fact, so far, stranding is even obligatory. In addition, stranding will be possible from intermediate positions, too, to which the symmetric WP-QP structure moved to, so long as the interface condition does not have forced cyclic points of interpretation (à la Fox and Pesetsky 2005, for instance). An example is given in (126):

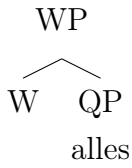
(126) *Stranding of* alles *out of symmetry:*

[55] Ott motivates the XP status. Many of these argument appear not to be extendable to *alles* at first sight.



The way in which pied-piping of *alles* is allowed with simplex *wh*-expression, and only with those, is that simplex *wh*-expressions have a special bare phrase structural status: they are both minimal and maximal, that is to say that they can function both as XPs and as X^0s in X′-theoretic terms. A compelling case for their split personality was made in Donati (2006) based on free relatives. A simplex *wh* may thus undo the tension in (125) in another way, namely by projecting:

(127) *Symmetric source for* alles*:*



Note that the dual status of a simplex *wh* has a specific effect, namely that it has to be one or the other once something in the derivation happens for which one status over the other is needed (see Donati, 2006: 33). For stranding this means that once *wh* projects, subsequent sub-extraction will be impossible for one of two reasons. If *wh* is now a head, and *wh*-phrases are nominals/DPs, the issue is that heads of nominal/DP projections cannot move in German. (128) illustrates this for a determiner.

(128) *Das habe ich [e_i Auto] gestern gekauft.
the.ACC have.1SG I.NOM car yesterday bought
Intended: 'I bought the car yesterday.'

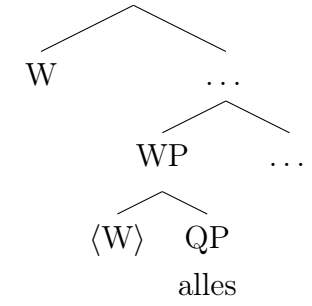
Second, even if *wh* could somehow still function as an XP after it projects, it will become a left branch, e.g. in Ross's original sense of 'the leftmost NP of a larger NP' (here understood as DP) (Ross, 1967: chapter 4.3.2.1). German, as Ross discusses for English, does not allow extraction of the left branch (so-called *Left Branch Condition*). This is illustrated for a *wh*-phrase in (129)– *wem* can only be interpreted as a benefactive (and oddly so given the meaning of the sentence), but not in its possessive interpretation equivalent to English *whose* or more standard German *wessen*.

(129) *Wem hast du gestern [e_i seine Kinder] geschimpft?
who.DAT have.2SG you.NOM yesterday his.PL children scolded
Intended: 'Whose children did you scolded yesterday?'

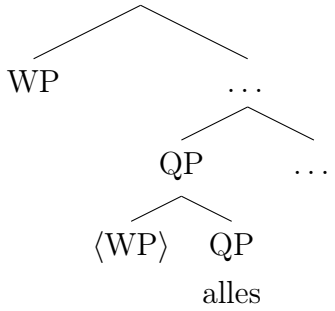
Nonetheless, the dual status also allows a simplex *wh* to be an XP, such that the *wh* will be able to, and actually be forced to move out and strand *alles*. Thus, (130a) will be impossible, but (130b) will be possible. W indicates the minimal status of *wh*, WP the maximal status.

(130) *No stranding after projection:*

a. *Left-Branch



b.



The full typology of derivations is given in (131).

(131) *Typology of derivations:*

		DP	PP
simplex	stranded	WP-QP: movement	PP-QP: movement
	pied-piped	W-QP: projection	[P[W-QP]]: projection
complex	stranded	[W-NP]-QP: movement	PP-QP: movement
	pied-piped	*: impossible	*: impossible

Overall, the complexity split is successfully accounted for under this approach, be it on a labeling or on a linearization basis. What is required is that XP-YP structures can be merged together, and that these cause the interfaces to crash unless the symmetry is resolved. *Alles* must furthermore be assumed to be an internally complex syntactic object. The analysis leaves hopes that these structures are compatible for extension to other floating elements that can co-occur with *alles* (see section 5.2), or to invariant *alles* in predicational structures, such as (132) (see Giusti, 1991).

(132) Wir sind alles Squashspieler.
we.NOM be.1PL ALLES squash.players
'We are all squash players.'

This approach, however, comes at the cost of not offering an explanation for why *alles* occurs in positions from which the associate has $\bar{\text{A}}$-moved but not in ones from which the associate has A-moved. Symmetry-breaking movement does not preclude A-movement from getting the job done—in fact, Ott proposes this analysis for A-QF. This question must remain unanswered here.

5.5.2 PF Filter

The final type of approach that can be taken to explain the complexity restriction is to resort to a PF filter. As Norbert Hornstein (p.c.) notes, low-level filters are indeed necessary for at least two other phenomena: multiple-*wh* clusters in Serbo-Croatian and Romanian, and the Doubl-*ing* Filter in English. I briefly review the two phenomena and their accounts in what follows. I then discuss how the complexity restriction compares to the two phenomena. The conclusion based on their discussion will be that while a filter similar in spirit to the ones proposed for these phenomena *might be possible* to be stated for the complexity restriction on *alles*, it will remain a construction specific quirk with a serious exception. That is, while broader design features or phenomena seem to be involved in the filters applying to multiple*wh* clusters and Double *-ing* configurations, I can think of no potentially broader property of Grammar or Language that could be invoked in the case of the complexity restriction on *alles*. In addition, in contrast to the other two filters, a filter for the complexity restriction resists a local statement, and instead seems to necessarily resort to a global rule that scans larger portions of the tree. The *was für NP* construction remains an exception.

Multiple-*wh* clusters in Serbo-Croatian and Romanian Nunes (2004: 35f) discusses the following minimal pairs from work by Željko Bošković.

(133) *Serbo-Croatian* (from Bošković 2000, 2002):

a. Ko šta kupuje?
who what buys

b. *Ko kupuje šta?
who what buys
'Who buys what?'

c. *Šta šta uslovljava?
what what conditions

d. Šta uslovljava šta?
what conditions what
'What conditions what?'

(134) *Romanian* (from Bošković 2002):

a. *Ce ce precede?
what what precedes

b. Ce precede ce?
what precedes what
'What precedes what?'

(133ab) are meant to be representative of the fact that in multiple *wh*-questions in Serbo-Croatian all *wh*-phrases move to the front. *Šta* can therefore not occur in-situ in (133b) as it would in English or German. This is a feature of the linguistic area that includes Romanian. Crucially, in Serbo-Croatian (133cd), and Romanian (134ab) multiple *wh*-fronting is impossible. The observation is that the non-multiple *wh*-fronting examples involve two *wh*-phrases that sound the same: *šta–šta* and *ce–ce*. Bošković argues that in the examples where the second *wh*-phrase occurs in-situ, this is only an appearance, and that instead the *wh*-phrase moves as expected by multiple *wh*-fronting, but is pronounced in its lower position. In terms of the copy theory of movement, he argues that the lower copy of the chain is pronounced. Suggestive evidence in favor of this hypothesis comes from the pair in (135), which contrasts with (133cd). In (135), the two instances of 'what' are separated by an adverb, and because of that, a higher copy is able to be pronounced. (Whether it is the highest in Spec,C or not does not matter here.)

(135) *Serbo-Croatian* (from Bošković 2000, 2002):

a. Šta neprestano šta uslovljava?
what constantly what conditions

b. *Šta neprestano uslovljava šta?
what constantly conditions what
'What constantly conditions what?'

Bošković adds more compelling evidence for Romanian, where he shows that the "exceptionally in-situ" *wh*-phrase is able to license parasitic gaps, which are generally argued not be licensed by genuine in-situ *wh*-phrases.

(136) *Romanian* (from Bošković 2002):

Ce precede ce fara sa influenteze?
what precedes what without SUBJ.PRT influence.3SG
'What precedes what without influencing?'

The conclusion that Nunes draws is that there is good evidence for the existence of a default pronunciation of chains, i.e. pronunciation of the highest copy, based on the fact that this

default can be disrupted such that other copies, e.g. the lowest, are pronounced instead. For our purposes, the conclusion is that there are filters that affect the choice of when or where a copy can be pronounced.

Doubl-*ing* Filter in English The so-called Doubl-*ing* Filter is a restriction on certain co-occurrences of two *-ing*-inflected verbs.[56]

(137) a. It continues to rain.
b. It continues raining.
c. It is continuing to rain.
d. *It is continuing raining. (Pullum and Zwicky, 1999: 251)

Without going into the details of all the co-occurrences that are unacceptable for many speakers of English, and which instead are acceptable, Pullum and Zwicky (1999) state the filter as a prohibition on the adjacency between a gerund-participle inflected verb and a VP-complement headed by another gerund-participle inflected verb.[57] They give the following illustration of the filter (Pullum and Zwicky, 1999: 261):

(138) The Doubl-*ing* Filter (third revision):
The following type of local tree is not permitted:

With this formulation Pullum and Zwicky (1999) argues to improve on previous formulations of the filter both empirically and conceptually. Empirically, (138) explains contrasts of the following types in (139)–(140), where intervening material bleeds the filter when it is part of the matrix VP constituent but not when it is part of the complement VP constituent (p. 262).

(139) a. [$_{\text{VP}}$ Keeping right on [$_{\text{VP}}$ drinking]] would be most unwise.
b. *[$_{\text{VP}}$ Keeping [$_{\text{VP}}$ secretly drinking]] would be most unwise.

(140) a. I hope you won't be [$_{\text{VP}}$ continuing throughout next week [$_{\text{VP}}$ going over the same material]].

[56] As far as I can tell from the discussion in Pullum and Zwicky (1999), the filter originally goes back to Ross (1972).

[57] *Gerund participle* is a technical term here, comprising only a certain class of *-ing* contexts, essentially all *-ing* forms of *inflecting* verbs, including progressives, quasi-progressives, VP complements, periphrastic futures, and more. See Pullum and Zwicky (1999: 258).

b. *I hope you won't be [$_{VP}$ continuing [$_{VP}$ suddenly jumping out and scaring people]].

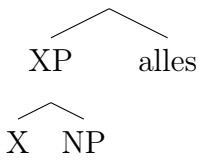
Conceptually, they argue the filter is local, and therefore avoids issues of being defined globally or trans-derivationally, as they argue previous accounts needed to. For our purposes, this filter is relevant because the structural configuration in which the two *-ing*-inflected verbs are matters. For instance, examples such as (141) are simply not affected, because the structural configuration is different (a "progressive in apposition"; Pullum and Zwicky, 1999: 259):

(141) I was *sitting thinking* about my troubles.

Complexity restriction on *alles* Section 5.2 explored various ways in which one could state a *requirement* on *alles*, clearly unsuccessfully given the disjunctive character of the restriction when understood as a requirement: *alles* needs to be right-adjacent either to a simplex *wh* or to a *wh*-trace. The two phenomena and accounts above instead tackle the issue by stating a filter, a *prohibition*. A prohibition would be another way of doing away with the disjunction repeated above, just in case a coherent, and formulable, environment can be stated in which *alles* cannot appear. In that logic, stranded *alles* would be licit almost as an elsewhere case. In the case of multiple-*wh*, the environment was adjacency between phonetically identical lexical items, essentially a haplology. In the case of double *-ing* configurations, the prohibited context was complementation of an *-ing* VP by an *-ing* V, essentially a dissimilation of sorts. We can now ask: what would a filter for the complexity restriction on *alles* look like?

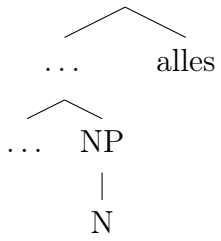
We know that the filter must be able to refer to syntactic structure, just as the Doubl-*ing* Filter above, because ellipsis cannot repair the complexity restriction (see again section 5.2; this conclusion must be coupled with the conclusion that ellipsis, if it applies in narrow syntax, deletes phonetic material but not syntactic structure). For examples involving **[which NP alles]* or **[whose NP alles]*, a filter that makes reference to structure could look like (142).

(142) *Complexity Filter on* alles*:*
Alles cannot be sister to an XP whose complement is an NP; the following tree is illicit.



While (142) restricts the filter to applying to material that is in the same constituent as *alles*, it is still unsatisfactory given that he restriction is not local: it does not resort to either adjacency or subjacency. In complex associates formed by PPs (**[P wh NP alles]*), the NP will be even deeper inside the complement of *alles* than in (142). The only statement that appears to work is then (143), which requires globality (even if in a rater limited sense, perhaps):

(143) *Complexity Filter on* alles (revised version):
Alles cannot be sister to an XP that contains an NP/N; the following tree is illicit



This filter might work under the assumption that simplex *wh*-phrases are D^0s. Alternatively, they may be DPs; what is required is that they be internally simplex in a way that they do not branch down to an NP/N. Assuming Trace Theory, the same consideration must apply to traces, so that they cannot be NP/Ns. Assuming the Copy Theory instead, the filter must apply to a level of representation where "copy reduction" has already applied. If the filter in (143) instead applied to a level of representation where copies are still fully represented, then (143) would apply equally to stranded *alles* as to adjacent *alles* so that associates with complex *wh*-phrases would be illicit in general. The issue that *alles* can occur adjacent to the *was für NP* construction still remains.

(143) is rather unappealing from a conceptual point of view, both because of its lack of locality, and because the filter cannot be understood as a known surface process, e.g. as a haplology or a dissimilation as was the case for the two other filters discussed above. The question then is: can we resort to adjacency to some specific category instead? It seems that all examples have in common the property that *alles* is adjacent to a noun, assuming that the complement of *welch-* 'which/what' and the complement of *wessen* 'whose' are NPs containing a noun; for PPs and dative possessive constructions the same can be assumed to hold since there is, one step further down in the structure, an NP that will be adjacent to *alles*. Conversely, simplex *wh*-phrases must be assumed to belong to a different category, e.g. DPs, and this non-terminal difference must be what matters for the filter. Still, a fundamental issue remains. The adjacency does not care about the internal structure of the constituent the adjacent NP is a part of. In fact, as discussed above, it *must* not care about

it. The predictions become too broad, then, such that we end up undergenerating: we would predict that any configuration where *alles* is right-adjacent to some NP is illicit. However, complex associates composed of the *was...für NP* construction allow *alles* to the right of the NP, and *alles* can be stranded in positions in the clause where it ends up being right-adjacent to some NP, too. It appears that we run yet again into a paradoxical situation. We must then return to the dissatisfying solution in (143). I conclude by reiterating that it does not suffice for a filter to be statable; ideally, a surface filter would fall under one of the classes of processes that we know exist in natural language. I coarsely depicted the multiple-*wh* filter as a haplology, and the Doubl-*ing* Filter as a dissimilation. It is far from obvious to me at this point what category process a surface filter formulation (or any formulation) of the complexity restriction should be understood as.

6 Conclusion

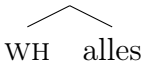
This paper investigated the syntax of German *wh*-quantifier float, "invariant *alles*", asking whether instances of adjacent *alles* and instances of floated *alles* ((144ab)) can be given, should be given, or require a uniform analysis.

(144) a. [Wen alles]$_1$ hat der Peter t_1 zur Party mitgebracht?
who.ACC all have.3SG the.NOM Peter to.the party brought.with
'Who all did Peter bring along to the party?'

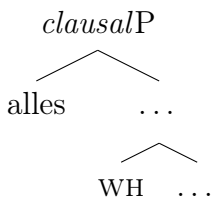
b. Wen$_1$ hat der Peter t_1 alles zur Party mitgebracht?
who.ACC have.3SG the.NOM Peter all to.the party brought.with
'Who all did Peter bring along to the party?'

Specifically, the question was asked whether the two quantifiers share the same starting structural configuration, such that they have the *Same Source* (145a), or a *Different Source* (145a)+(145b).

(145) a.



b.



Building on previously established analyses of quantifier float more broadly, a number of first-principles properties and predictions of theories that assume a same source (SSTs) were established, and compared to first-principles properties and predictions of theories that as-

sume a different source (DSTs). The conclusion reached in the first part of the paper is that the empirical generalization about the distribution of *alles* are fully in line with the SST, while much challenging the DST. The list of properties and findings is repeated here:

(a) *Synonymy property:*
The same source for adjacent and floated quantifiers entails the same relation between them and their associate, and therefore the same meaning contribution.

⇒ Adjacent *alles* and distal *alles* make the same meaning contribution (section 3.1)

(b) *Uniqueness property:*
The same source for adjacent and floated quantifiers entails that (these) quantifiers cannot be introduced in other ways into the sentence and therefore a one-to-one correspondence between quantifiers and associates given.

⇒ There is a one-to-one correspondence between *alles*, whether adjacent or distal, and associates (section 3.2).

(c) *Stranding property:*
The same source for adjacent and floated quantifiers entails that (if the quantifier itself may not be moved further) the distribution of the associates bounds the distribution of the floated quantifiers.

⇒ Distal *alles* occurs in a (proper) subset of the positions that its associate may occupy independent of *alles* (section 3.3).

(d) *Insider property:*
However stranding is achieved, the procedure will involve the constituent that *alles* is a part of, or a sub-constituent of which *alles* is a part of and in a dependency with–the associate. The same source for adjacent and floated quantifiers entails that the quantifier can be sensitive to how the stranding procedure is effected.

⇒ Distal *alles* occurs in $\bar{\text{A}}$-trace but not in A-trace positions (section 3.4).

(e) *Local link property:*
The same source for adjacent and floated quantifiers entails that floats must be created by movement, and that therefore the same locality restrictions that apply to movement apply to a floated quantifier relative to its associate.

⇒ Distal *alles* must be c-commanded by a clausemate link of a suitable $\bar{\text{A}}$-chain, in overt syntax (section 3.5).

(f) *Early Link property:*
The same source for adjacent and floated quantifiers suggests that the relation between

a floated quantifier and its associate is established in narrow syntax.

⇒ Distal *alles* cannot be licensed at LF; it must be licensed at least in narrow syntax (section 3.5).

(g) *Form property:*
The same source for adjacent and distal quantifiers entails that the quantifier may select for a form of the associate, or that the associate may select for some form of the quantifier, and that the selection is the same for both adjacent and floating quantifiers.

⇒ *Alles* can only associate with indefinite operators, in the sense of Reis (1992) (section 2.2).

The second part of the paper aimed to restrict the space of analyses of how floated *alles* is 'stranded'. McCloskey (2000)'s analysis is explored in detail, finally rejecting it on the basis of a complexity split that *alles* exhibits: *alles* must be the sister of either a *wh*-trace, or of a *simplex wh*-phrase:

(146) a. DP/PP b. DP

Three approaches for how to unify the disjunction were explored, one narrow syntactic, one an LF condition, and one a PF/Spell-Out condition. The narrow syntactic one unified the disjunction by assuming that *wh*-traces are *wh*-pronouns, just like simplex *wh*-phrases. The LF condition unified the disjunction by assuming that both *wh*-traces and *wh*-expressions are variables, and that *alles* must be sister to a variable at LF. Finally, the PF/Spell-Out condition unified the disjunction by assuming that *alles* and the associate form a symmetric XP-YP (predicational) structure, that this kind of structure is illegible for the interfaces, and that there are two ways to make the structure legible—one corresponding to *wh*-traces, and one corresponding to simplex *wh*-phrases.

Overall, the complexity split will play an important role in determining the correct analysis of *alles* in the future, and perhaps of *wh*-quantifier float more broadly. Other work has not discussed it, as far as could be seen. If the complexity split is cross-linguistically present, and cross-linguistically not repairable by ellipsis (i.e. not a surface restriction), there must be something "deep" about the nature of *wh*-quantifier float that brings this split about.

The other significant empirical finding is that *alles* can only be stranded in positions corresponding to $\bar{\text{A}}$-traces of the the associate, in a given derivation. This finding replicates the finding by McCloskey (2000) for West Ulster English, and Fitzpatrick (2006) for Korean, Japanese, and Russian, increasing the suspicion that there are universal constraints at work,

and hopefully lending itself as a model for further investigation, and as a viable domain for investigation into the nature of "reconstruction".

I want to conclude by highlighting the implications that the Same-Source finding for floated *alles* has for empirical domains outside of quantifier float. Most prominently, it carried weight into the investigation of long-distance $\bar{\text{A}}$-dependencies. As argued in this paper, the distribution of *alles* motivates that long-distance *wh*-movement in German proceeds successive-cyclically, not only via CP, as has been motivated before (cf. Bayer et al., 2016), but also via *v*P. The distribution of *alles* thus adds to the idea that there is a cyclic-node in the verbal domain as well, as thought since Chomsky (1986a). In addition, given the same-source conclusion, and the conclusion about the $\bar{\text{A}}$-trace distribution of floated *alles*, this may become a fruitful domain of inquiry into the nature of differing $\bar{\text{A}}$-movement dependencies, and into the nature of chains composed by differing movement dependencies: can *alles* be floated in the $\bar{\text{A}}$-trace positions corresponding to Parasitic Gaps, or to the gap in comparatives? Even more so, what is the distribution of *alles* in other long-distance *wh*-questions in German, e.g. in *wh*-scope marking (*aka* partial *wh*-movement), and can it help settle the divide between the Indirect Dependency and the Direct Dependency approaches there? One way or another, the study of *alles* is certain to impose significant boundary conditions on the understanding of other empirical domains, as well.

References

Anagnostopoulou, Elena (2008). Notes on the Person Case Constraint in Germanic (with Special Reference to German). In: *Agreement Restrictions*, edited by Roberta D'Alessandro, Susann Fischer, and Gunnar Hrafn Hrafnbjarganson, Berlin/New York: Mouton de Gruyter, pp. 15–47.

Bayer, Josef, Jana Häussler, and Markus Bader (2016). A New Diagnostic for Cyclic Wh-Movement: Discourse Particles in German Questions. *Linguistic Inquiry*, 47, no. 4: pp. 591–629.

Beck, Sigrid (1996). *Wh*-Constructions and Transparent Logical Form. Ph.D. thesis, Universität Tübingen.

Belletti, Adriana and Luigi Rizzi (1988). Psych-Verbs and Θ-Theory. *Natural Language and Linguistic Theory*, 6: pp. 291?–352.

Bobaljik, Jonathan (1995). Morphosyntax: The Syntax of Verbal Inflection. Ph.D. thesis, MIT, Cambridge, MA.

Bobaljik, Jonathan (2003). Floating Quantifiers: Handle with Care. In: *The Second GLOT International State-of-the-Article Book*, edited by Lisa Cheng and Rint Sybesma, Berlin: Mouton, pp. 107–148.

Bošković, Željko (2000). What Is Special About Multiple *Wh*-Fronting? In: *Proceedings of NELS 30*, edited by Mako Hirotani, Andries Coetzee, Nancy Hall, and Ji-Yung Kim, Amherst: University of Massachusetts, GLSA, pp. 83–107.

Bošković, Željko (2002). On Multiple *Wh*-Fronting. *Linguistic Inquiry*, 33: pp. 351–383.

Bošković, Željko (2004). Be Careful Where You Float Your Quantifiers. *Natural Language and Linguistic Theory*, 22, no. 4: pp. 681–742.

Chomsky, Noam (1973). Conditions on Transformations. In: *Festschrift for Morris Halle*, edited by Stephen Anderson and Paul Kiparsky, New York: Holt, Rinehart and Winston, pp. 232–286.

Chomsky, Noam (1977). On Wh-Movement. In: *Formal Syntax*, edited by P. Culicover, Tom Wasow, and A. Akmajian, Academic Press, pp. 71–132.

Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Chomsky, Noam (1986a). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam (1986b). *Knowledge of Language*. New York: Prager Publishers.

Chomsky, Noam (1993). A Minimalist Program for Linguistic Theory. In: *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by K. Hale and S. J. Keyser, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam (2000). Minimalist Inquiries: The Framework. In: *Step by Step. Essays on Minimalist Syntax in Honour of Howard Lasnik*, edited by Roger Martin, David Michaels, and Juan Uriagereka, Cambridge, MA: MIT Press, pp. 89–155.

Chomsky, Noam (2001). Derivation by Phase. In: *Ken Hale: A Life in Language*, edited by Michael Kenstowicz, Cambridge, MA: MIT Press, pp. 1–52.

Davis, Colin (2018). Crossing and Stranding at Edges, https://ling.auf.net/lingbuzz/004138.

Donati, Caterina (2006). On Wh-Head-Movement. In: *Wh-Movement: Moving on*, edited by L. Cheng and N. Corver, Cambridge, Mass.: MIT Press, pp. 21–46.

Dougherty, Ray (1970). A Grammar of Coordinate Conjoined Structures. *Language*, 46: pp. 850–898.

Dowty, David and Belinda Brodie (1984). The Semantics of 'Floated' Quantifiers in a Transformational Grammar. In: *Proceedings of the 3rd West Coast Conference on Formal Linguistics (WCCFL 3)*, Stanford: CSLI, pp. 75–90.

Fanselow, Gisbert (1990). Scrambling as NP-Movemen. In: *Scrambling and Barriers*, edited by Günther Grewendorf and Wolfgang Sternefeld, Amsterdam/Philadelphia: John Benjamins, pp. 113–140.

Fanselow, Gisbert and Damir Ćavar (2002). Distributed Deletion. In: *Theoretical Approaches to Universals*, edited by Artemis Alexiadou, Amsterdam: John Benjamins, pp. 65–107.

Fanselow, Gisbert and Caroline Féry (2008). Missing Superiority Effects: Long Movement in German (and other languages). In: *Elements of Slavic and Germanic grammars: A comparative view*, edited by Jacek Witkoś and Gisbert Fanselow, *Polish Studies in English Language and Linguistics*, vol. 23, Frankfurt: Lang, pp. 67–87.

Fitzpatrick, Justin Micheal (2006). The Syntactic and Semantic Roots of Floating Quantification. Ph.D. thesis, MIT, Cambridge, MA.

Fox, Danny and David Pesetsky (2005). Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics*, 31: pp. 1–46.

Frey, Werner (2006). Contrast and Movement to the German Prefield. In: *The Architecture of Focus*, edited by V. Molnár and S. Winkler, Berlin: Mouton de Gruyter, pp. 235–264.

Giusti, Giuliana (1990). Floating Quantifiers, Scrambling, and Configurationality. *Linguistic Inquiry*, 21: pp. 633–641.

Giusti, Giuliana (1991). The Syntax of Floating *Alles* in German. *Issues in Germanic Syntax*: pp. 327–350.

Grewendorf, Günther (2005). The Discourse Configurationality of Scrambling. In: *The Free Word Order Phenomenon*, edited by Joachim Sabel and Momoro Saito, Berlin: Mouton de Gruyter, pp. 75–135.

Grewendorf, Günther and Joachim Sabel (1999). Scrambling in German and Japanese. *Natural Language and Linguistic Theory*, 17: pp. 1–65.

Grewendorf, Günther and Wolfgang Sternefeld (1990). Scrambling Theories. In: *Scrambling and Barriers*, edited by Günther Grewendorf and Wolfgang Sternefeld, Amsterdam/Philadelphia: John Benjamins, pp. 3–37.

Haider, Hubert (1993). *Deutsche Syntax – Generativ: Vorstudien zur Theorie einer Projektiven Grammatik*. Tübingen: Gunter Narr Verlag.

Haider, Hubert and Ingrid Rosengren (1998). *Scrambling*, *Sprache und Pragmatik*, vol. 49. Germanistisches Institut der Universität Lund.

Heck, Fabian and Anke Himmelreich (2017). Opaque Intervention. *Linguistic Inquiry*, 48, no. 1: pp. 47–97.

Henry, Alison (2012). Phase Edges, Quantifier Float and the Nature of (Micro-) Variation. *Iberia*, 4, no. 1: pp. 23–39.

Junker, Marie-Odile (1995). *Syntaxe et Sémantique des Quantifieurs Flottants* tous *et* chacun*: Distributivité en Sémantique Conceptuelle*. Genève: Librarie Droz.

Kayne, Richard S. (1975). *French Syntax: The Transformational Cycle*. Cambridge MA: MIT Press.

Ko, Heejeong (2007). Asymmetries in Scrambling and Cyclic Linearization. *Linguistic Inquiry*, 38, no. 1: pp. 49–83.

Koopman, Hilda (1999). The Internal and External Distribution of Pronominal DPs. In: *Beyond principles and parameters: Essays in memory of Osvaldo Jaeggli*, edited by Kyle Johnson and Ian Roberts, Dordrecht: Kluwer, pp. 91–132.

Koopman, Hilda (2010). On Dutch *allemaal* and West Ulster English *all*. In: *Structure Preserved*, edited by Jan Wouter Zwart and Jan de Vries, Amsterdam: John Benjamins, pp. 267–275.

Lenerz, Jürgen (1977). *Zur Abfolge Nominaler Satzglieder im Deutschen*. Studien zur deutschen Grammatik, Tübingen: Narr.

Marantz, Alec (2001). Words, talk given at WCCFL, University of Southern California.

Matushansky, Ora (2006). Head Movement in Linguistic Theory. *Linguistic Inquiry*, 37, no. 1: pp. 69–109.

McCloskey, James (2000). Quantifier Float and Wh-Movement in Irish English. *Linguistic Inquiry*, 31, no. 1: pp. 57–84.

McCloskey, James (2017). Micro-Parameters in a Tiny Space – Stranding at the Edge, handout for 'S-Circle', UCSC, November 3, 2017.

Merchant, Jason (1996). Object Scrambling and Quantifier Float in German. In: *Proceedings of NELS 27*, UMass Amherst: GSLA, pp. 179–194.

Miyagawa, Shigeru (1989). *Structure and Case Marking in Japanese*. San Diego: Academic Press.

Moro, Andrea (2000). *Dynamic Antisymmetry*, *Linguistic Inquiry Monographs*, vol. 38. Cambridge, MA: MIT Press.

Müller, Gereon (2001). Harmonic Alignment and the Hierarchy of Pronouns in German. In: *Pronouns – Grammar and Representation*, edited by Horst Simon and Heike Wiese, Amsterdam: Benjamins, pp. 203–232.

Müller, Gereon (to appear). Structure Removal in Complex Prefields. *Natural Language and Linguistic Theory*.

Nunes, Jairo (2004). *Linearization of Chains and Sideward Movement*. MIT Press.

Ott, Dennis (2011). Local Instability: The Syntax of Split Topics. Ph.D. thesis, Harvard University, Cambridge, MA.

Ott, Dennis (2012). *Local Instability: Split Topicalization and Quantifier Float in German*, *Linguistische Arbeiten*, vol. 544. Berlin: de Gruyter.

Pafel, Jürgen (1991). Zum Relativen Skopus von w- und Q-Phrasen (w/Q-Interaktion). In: *Fragesätze und Fragen*, edited by Marga Reis and Inger Rosengren, Tübingen: Niemeyer, pp. 145–173.

Postal, Paul (1971). *Crossover Phenomena*. New York: Holt, Rinehart and Winston.

Postal, Paul (1974). *On Raising: One Rule of English Grammar and Its Theoretical Implications*. Cambridge, MA: MIT Press.

Pullum, Geoffrey and Arnold Zwicky (1999). Gerund Participles and Head-Complement Inflection Conditions. In: *The Clause in English: In honour of Rodney Huddleston*, edited by Peter Collins and David Lee, Amsterdam/Philadelphi: John Benjamins, pp. 251–271.

Reis, Marga (1992). The Category of Invariant *alles* in *wh*-Clauses: On Syntactic Quantifiers vs. Quantifying Particles in German. In: *Who Climbs the Grammar Tree?*, edited by Rosemary Tracy, Tübingen: Max Niemeyer Verlag, pp. 465–492.

Ross, John Robert (1967). Constraints on Variables in Syntax. Ph.D. thesis, MIT, Cambridge, Mass.

Ross, John Robert (1972). Double-*ing*. *Linguistic Inquiry*, 3: pp. 61–86.

Safir, Ken (2017). Weak Crossover. In: *The Wiley Blackwell Companion to Syntax, Second Edition*, edited by Martin Everaert and Henk van Riemsdijk, John Wiley & Sons, pp. 1–40.

Shlonsky, Ur (1991). Quantifiers as Functional Heads: A Study of Quantifier Float in Hebrew. *Lingua*, 84: pp. 159–180.

Sportiche, Dominique (1988). A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry*, 19: pp. 425–449.

Sportiche, Dominique (1996). Clitic constructions. In: *Phrase Structure and the Lexicon*, Springer, pp. 213–276.

von Stechow, Arnim and Wolfgang Sternefeld (1988). *Bausteine Syntaktischen Wissens*. Opladen.

Uriagereka, Juan (2012). *Spell-Out and the Minimalist Program*. Oxford: Oxford University Press.

Wasow, Tom (1979). *Anaphora in Generative Grammar*. Ghent: E. Story-Scientia, based on 1972 MIT PhD dissertation.

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

Wiltschko, Martina (1997). D-Linking, Scrambling und Superiority in German. In: *GAGL: Groninger Arbeiten zur germanistischen Linguistik*, edited by Werner Abraham, vol. 41, Groningen, Netherlands: Germanistisch Instituut, Rijksuniversiteit, pp. 107–142.

Zimmermann, Malte (2007). Quantifying Question Particles in German: Syntactic Effects on Interpretation. In: *Proceedings of Sinn und Bedeutung 11*, edited by Estela Puig-Waldmüller, Barcelona: Universitat Pompeu Fabra, pp. 627–641.

A Two different *alles*-populations?

The differences between *alles* and the inflecting quantifier highlighted in section 2.2 were a re-elaboration on findings in Reis (1992). I would like to add another way in which *alles* and A-QF must be kept separate, as it constitutes a potential confound when investigating invariant *alles*. While it is straightforward given the diagnostics mentioned so far, to distinguish invariant *alles* from A-QF *all-* with complex *wh*-phrase associates, it seems that instances of *alles* with simplex *wh*-pronoun associates are ambiguous between invariant *alles*, and an A-QF showing singular neuter agreement.[58] For certain speakers, *alles* is not available with complex *wh*-phrases—certainly *welch- NP* 'which NP', *wessen NP/wem sein-/ihr- NP* 'whose NP', but potentially still with *was für NP* 'what NP'. Rather, the only quantifier available with these expressions is the correct form of A-QF *all-*. If indeed these speakers did not have an independent lexical item corresponding to (invariant) *alles*, the lack of *alles* with complex *wh*-phrases would be expected. Their form of *alles* available with simplex *wh*-phrases would be an agreeing A-QF, under the assumption that simplex *wh*-phrases bear 3SG.N for ϕ-features, given that *-/es/* is the 3SG.N suffix, as in *das all-es* 'that all-3SG.N'. There are a number of properties that seem to cluster together with this distinction: (i) *überall* 'everywhere' is preferred to *alles* with *wo* 'where'; (ii) *alles* is much dispreferred in matrix clauses of *was..w*-questions (also known as *wh*-scope marking, or partial *wh*-movement); (iii) stranding with pied-piped PPs is very marginal; (iv) the requirement for an exhaustive list answer is not as clearly present. These properties in addition seem to be further supported by the fact that for me, and at least one other speaker tested that accepts *alles* with complex *wh*-phrases (i) *alles* is as available as *überall* with 'where', (ii) *alles* is accepted in matrix clauses of long questions, including *was..w*-questions (though Marga Reis rejects *alles* in *was..w*-questions Reis 1992: ex65), (i′/ii′) *alles* but not *überall* is accepted in matrix clauses of long questions, [59] (iii) *alles* can be stranded with pied-piped PPs, (iv) the answer is expected to be exhaustive. Two more facts can be added. For one, one speaker who did not accept *alles* with complex *wh*-phrases accepts it with combination of *alles* and the particle *so* (so-called 'non-exhaustive *so*'; see Reis 1992), e.g. with a pied-piped PP. On the other hand, adding *so* disambiguates between A-QF-*alles* and invariant *alles* for myself, too. Without further exploring how this effect may come about, notice the following interaction with the adverb *gleichzeitig* 'at the same time', which, according to Junker (1995), presupposes a plurality of events and therefore interacts with distributivity.

[58] Thanks to Julian Schlöder for helpful discussion leading to this conclusion.

[59] To which can be added that Reis (1992: section 2.4) had already noted that 'non-exhaustive *so*' is incompatible with the inflecting *all-*, and indeed 'non-exhaustive *so*' is also incompatible with *überall* further corroborating the idea that *überall* belongs the class of inflecting *all-* and not to invariant *alles*.

(1) a. Wer (alles) ist (alles) gleichzeitig ins Ziel gekommen?
who.NOM be.3SG same.time in.the goal arrived
'Who (all) reached the finish line at the same time?'

b. Wem seine Kinder sind (all-e /??alles /??so alles)
who.NOM his.PL children be.3PL A-QF-NOM.PL / WQ / so WQ
gleichzeitig ins Ziel gekommen?
same.time in.the goal arrived
'Whose children did (all) reach the finish line at the same time?'

c. Wer (??so alles) ist (??so alles) gleichzeitig ins Ziel gekommen?
who.NOM so WQ be.3SG so WQ same.time to.the goal arrived
'Who (all) reached the goal at the same time?'

How systematic these correlations are across the population is left as an open matter here. It is clear, however, that there are two separate populations, one not having, at a phenomenological level, the *alles* investigated in this paper.