ULI SAUERLAND

THE INTERPRETATION OF TRACES*

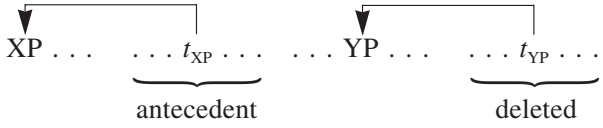
This paper argues that parts of the lexical content of an A-bar moved phrase must be interpreted in the base position of movement. The argument is based on a study of deletion of a phrase that contains the base position of movement. I show that deletion licensing is sensitive to the content of the moved phrase. In this way, I corroborate and extend conclusions based on Condition C reconstruction by N. Chomsky and D. Fox. My result provides semantic evidence for the existence of traces and gives semantic content to the A/A-bar distinction.

In high-school physics class, we did the following experiment: Take two identical, empty glass containers and put them on a scale: they should weigh the same amount. Then use a pump to evacuate the air from one of the containers, and put them back on the scale. Now the scale shows that the evacuated container is lighter than the other one. The experiment demonstrates that despite initial appearance the empty container is not really empty: its invisible content is air which weighs more than nothing.

This paper applies the same logic to show that traces have more content than their initial appearance shows. In this version of the experiment, verb phrase deletion takes the place of the scale: it tests whether two constituents have the same interpretation. The place of the glass containers is taken by two constituents that seem identical, but both contain a trace. The configuration is sketched in (1):

* I started working on this topic at the beginning of 1998 for parts of my dissertation (Sauerland 1998). Noam Chomsky, Danny Fox, Irene Heim, and David Pesetsky gave me a lot of useful advice at this stage. Since then, my analysis has evolved significantly. In 2000, I completed a first paper on the topic (Sauerland 2000b). Thanks to the excellent comments of Danny Fox, Irene Heim, Kyle Johnson, Chris Kennedy, Winfried Lechner, Wolfgang Sternefeld, Arnim von Stechow, and one anonymous reviewer for *Natural Language Semantics*, I managed to rewrite the paper into its present, hopefully much better form. Christopher Kennedy, Marcin Morzycki, and Shoichi Takahashi checked the rewritten version carefully, which helped me to reduce the number of errors significantly. I am grateful for all the helpful comments I received on this topic during the last few years from the people just mentioned, as well as from many others who are not mentioned here for reasons of space. Needless to say, only I should be blamed for the errors that still remain. The German Research Council (DFG) is currently funding me as an Emmy-Noether-Fellow (Grant SA 925/1-1), which I gratefully acknowledge.

Natural Language Semantics **12**: 63–127, 2004.

(1) 

The deletion test will show that the two constituents mean something different, except when XP and YP, the antecedents of the traces in (1), have the same lexical content. I argue that this result shows that parts of the moved phrases are obligatorily interpreted in the trace positions t_{XP} and t_{YP}. VP-deletion reveals this contribution the antecedent makes to the interpretation of traces.

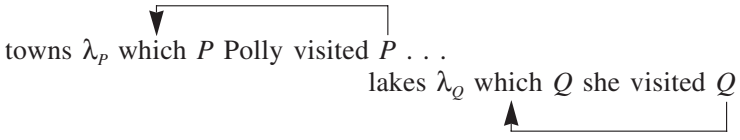
A concrete instance of the contrast I will be analyzing is (2). The examples involve antecedent-contained deletion (ACD) of a kind Kennedy (1994) first discussed – I introduce the term *double-headed ACD* in section 2. The only difference between the two examples is that where, in (2a), the word *lake* occurs, (2b) contains the word *one*, which is anaphoric to *town*. Interestingly, this lexical difference correlates with the indicated difference in grammaticality. (I indicate elided material by Δ, and the intended interpretation of the elided material in parentheses.)

(2) a.* Polly visited every town that is near the lake Erik did Δ. (Δ = visit *t*)
b. Polly visited every town that is near the one Erik did Δ. (Δ = visit *t*)

At first, the examples in (2) may seem needlessly complicated. The examples in (3) also instantiate the schema in (1).

(3) a. I know which towns Polly visited, but not which lakes she did Δ.
b. The cities Polly visited are near the lakes Erik did Δ.

But, as I will argue in detail in section 4 below, the effect of the trace content on deletion is absent in examples like (3) for independent reasons. Specifically, examples like (3) allow a sloppy construal of whatever the content of the trace may be, as sketched in (4) for (3a).

(4) 

Hardt (1999) and Schwarz (1999a, b) argue that such sloppy construals are available in examples with VP-deletion that are like (3) in all relevant

respects, and we therefore expect this kind of sloppiness to be possible in (3) too. However, as I show in section 4.3 below, the sloppy construal of the trace content is blocked in (2). Therefore, I focus on (2) in this introduction.

My paper develops and defends an analysis of contrasts like (2). One assumption I make is that the examples in (2) involve two instances of movement.[1] Specifically, the two movements are *wh*-movement in the relative clause and quantifier raising. (5) sketches the structures created by these two movements on standard assumptions: both create a structure where the VP contains just the verb *visit* and a trace.

(5) a. the lake/one Op Erik visited *t*
deleted
wh-movement

b. [every town that is near the lake/one Erik did] Polly visited *t*
antecedent
quantifier raising

In (5), the deleted VP is identical to its antecedent and therefore deletion should always be licensed. But this is not the case, as (2) shows.

The idea of my account of (2) is the following: Instead of (5), I assume the representations in (6), where the traces contain lexical material from the movement phrase or the head of the relative clause respectively. On the basis of (6), the contrast in (2) is easy to explain: VP-deletion in (2a) is blocked because the two VPs in (6) – *visited lake* and *visited town* – have different interpretations. The corresponding representation for (2b), on the other hand, would contain two VPs that have the same meaning: *visited town* and *visited one*.

[1] For concreteness, I will interpret *movement* literally as a transformational relationship between two representations. For present concerns, however, the term 'movement' could equally well be used as a meaningless mnemonic for a particular kind of relationship that can be established between the argument positions of two predicates in a phrase structure. From this point of view, my paper should be taken to reveal properties of this abstract 'movement' relationship, rather than liberal movement.

(6) a. the lake Op Erik visited *lake* or the one Op Erik visited *one*

deleted deleted

wh-movement *wh*-movement

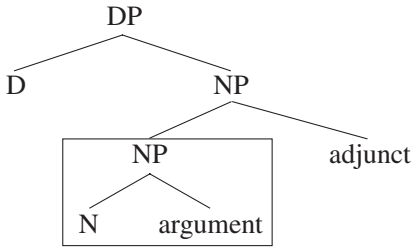
b. [every town that is near the lake/one Erik did] Polly visited *town*

antecedent

quantifier raising

In (6), some but not all lexical material of the moved phrases or the relative clause heads is repeated in the trace position. Specifically, I assume the following: for any DP, the *core NP* is the phrase consisting of the head noun N of the NP that's the complement of the determiner heading the DP and all arguments to N. (7) illustrates this definition: the core NP is marked by the box.

(7)



I argue that the core NP is repeated in the trace position in all cases of DP-movement. For relative clauses, I adopt the matching analysis (Carlson 1977; Sauerland 1998, 2000c, 2003; Bhatt 2002; Hulsey and Sauerland 2003).[2] According to the matching analysis, the relative clause operator is a determiner that must take as its complement an NP that is deleted under semantic identity with the head of the relative clause (see section 1.2). This deleted NP is also the core NP of the relative clause operator DP, and therefore my proposal implies that it must be repeated in the relative clause internal trace position. Therefore, it follows from the matching analysis that an NP semantically identical to the relative clause head must be present in the relative clause internal trace position.

Crucially, I argue that the core NP in the trace position makes a semantic contribution to the VP that contains the trace. As to what this semantic

[2] To be more precise, I assume that relative clauses are ambiguous between two analyses, raising and matching. In examples with ACD like (2), however, the raising analysis is blocked for independent reasons, as discussed by Hulsey and Sauerland (2003).

contribution is, I give a detailed proposal in the appendix. In the main part of this paper, however, I adopt the proposal of Fox (1999b) with a slightly modified notation. I assume that traces are interpreted as follows: The core NP is the argument of the indexed definite determiner defined in (8a). The indexed *the*$_n$ is interpreted like a pronoun as a value of the variable assignment, but with a presupposition that this value be in the extension of the NP-complement of *the*$_n$. Furthermore, this definite determiner must be coindexed with a variable binder that was merged with the sister of the phrase moved from the trace position. Example (8b) shows the logical form representation created by A-bar movement of *every city*.

(8) a. $[\![\text{the}_n]\!]^g(P)$ is defined if $P(g(n)) = 1$
where defined $[\![\text{the}_n]\!]^g(P) = g(n)$

b. $[\![\text{every city}]\!]\ \lambda_n \ldots \text{the}_n \text{ city}$

In (8b), the sister of the quantifier *every city* is a predicate that is defined only for cities. However, this is sufficient to determine whether every city has this property. Hence, the interpretation of (8b) is predicted to be identical to that of a traditional representation where the trace is interpreted as a plain variable. More generally for all conservative quantifiers, representations like (8b) and representations with plain variables as traces are equivalent, as Fox (1999b) points out. If all natural language quantifiers are conservative, the new semantics of traces therefore predicts the correct truth conditions for movement constructions.

The paper consists of seven sections. Section 1 introduces background assumptions concerning deletion licensing and the analysis of relative clauses. Section 2 presents three analyses of ACD structures like (2), focusing on those two that predict some restriction on this type of ACD: the index identity account of Heim (1997a) and the copy identity account of Sauerland (1998). Section 3 presents a number of arguments in favor of the copy identity account. Section 4 provides an analysis of cases of deletion like (3), where a trace is deleted, and explains why these differ from the type of ACD considered in sections 2 and 3. Section 5 establishes that the copy identity account leads to the same conclusions about the lexical context of traces as evidence from Condition C does. Section 6 is the conclusion. Section 7, the appendix, lays out two proposals for the semantics of chains, in particular in the context of functional interpretations.

1. Some Background Assumptions

1.1. *Deletion Licensing*

How deletion of VPs is licensed plays a major role in my analysis. I assume the licensing condition of Rooth (1992a) with the changes argued for by Fox (1999a). For the convenience of the reader, I summarize the deletion licensing condition here with some adjustment of notations. I assume that there is a syntactic feature Δ that renders a phrase unpronounced.[3] Δ must be licensed by occurring in the argument of a ~-operator in the way stated in (9a). The operator ~ in turn presupposes the presence of a discourse antecedent as stated in (9b).[4]

(9) a. ΔX is licensed iff there is a node & such that:
(i) Y is identical to or dominates X, and
(ii) Y bears the feature ~, and
(iii) for no terminal node Z dominated by X, there's a focus feature that dominates Z and is dominated by Y.

b. (To be revised in (21))
$[\![\sim_n X]\!]^g$ is defined iff $domain(g(n)) = domain([\![X]\!]_f)$ and $g(n)(h) \in [\![X]\!]_f^h$ for all assignment functions $h \in domain(g(n))$.
If defined, $[\![\sim_n X]\!]^g = [\![X]\!]^{g^{\backslash n}}$ for any g and n.

The conditions in (9b) make sure that $g(n)$ is exactly identical to the meaning of one element of the focus set of X. The value that g assigns to n, the index on ~, must be of the same semantic type of X: a partial function from partial assignment functions to entities of the semantic type as X. If used without any constraint, assignment functions that have functions from assignment functions to truth values as their values give rise to a version of Russell's paradox: the assignment function g such that $g(1)$ is true only of all assignment functions h such that $h(1)(h)$ is false. Therefore, I propose that the set of possible assignment functions is constrained by (10).

---

[3] Merchant (2001) calls the feature that renders a phrase unpronounced the E-feature. I chose to distinguish myself notationally, because the E-feature plays the roles of both the Δ and the ~-feature of my account. Specifically, Merchant works with a simplified version of Rooth's proposal where the deleted phrase itself has a semantic licensing requirement and unbound variables are closed off existentially. Merchant's account of ellipsis would be insufficient for my purposes. As far as I can see, Merchant's proposal has problems with unbound pronouns and loses Rooth's account of the parallel dependencies requirement, which I discuss below following example (12). Since this requirement plays a big role in my paper, I shall essentially adopt the full version of Rooth's system.

[4] I use the notation $[\![-]\!]_f^g$ for the focus value of a phrase under assignment g as defined by Rooth (1992b).

(10) $\forall g \forall n \in domain(g)$: $n \notin domain(g(n)) \wedge domain(g(n)) \subset domain(g)$

The restriction (10) entails that the assignment function g' that is the value of another assignment function g for some index n must be defined for a properly smaller domain than g itself. Therefore, any assignment function that is defined for itself does not satisfy (10), and Russell's paradox is avoided. The restriction (10) has a consequence for the last part of the definition (9b). There the value of $\sim_n$ X under assignment g could not be defined as the value of X under assignment g because then it would follow from the condition that $domain(g(n)) = domain([\![X]\!]_f)$ that $g \in domain\ (g(n))$. Therefore, the value of $\sim_n$ X must be defined as the value of X under the restricted assignment $g\backslash n$, which I use as an abbreviation for g' that has the domain $domain(g)\backslash\{n\}$ and which for all indices n in its domain yields the same value as g itself. Because of this restriction, one occurrence of $\sim_n$ that is in the scope of another occurrence of a coindexed $\sim_n$ can never be licensed. This restriction will play a role in section 5.5 below.

In (9), I assume that n represents a silent pronominal element of the type of constituent meanings and that usually there is an overt antecedent for the silent pronoun.[5] This requires that there be also variables of this type and an operator that ensures that the value of such a variable be identical to the interpretation of some constituent overtly occurring in the discourse. Therefore, we define a new operator, σ_n, as follows:

(11) $[\![\sigma_n\ X]\!]^g$ is defined if and only if $g(n)(h) = [\![X]\!]^h$ for any assignment function h for which $[\![X]\!]^h$ is defined.
If defined, $[\![\sigma_n X]\!]^g = [\![X]\!]^{g\backslash n}$.

The definition of the σ-operator must also avoid requiring that $g(n)$ have g in its domain. Hence, the complement of σ_n is evaluated relative to the assignment $g\backslash n$. This also blocks σ_n from occurring in the scope of a coindexed $\sim_n$ and vice versa.

Consider example (12) for an illustration of these concepts.

(12) John's coach approved of his contract and Bill did Δ too.

The strict interpretation of (12) is licensed as in (13a). (In (13) and in the following, I use numerals as the indices of variables of the same type as constituent meanings.) If $g(x)$ was John, (13a) would require that both John's

[5] What I account for here is the fact that constituent meanings that occurred once in a discourse are subsequently available for the licensing of deletion. I do not account for the fact that this availability declines with time, nor for cases where constituent meanings become available for deletion licensing without having occurred somewhere else in the discourse.

coach and Bill approve of John's contract. (13b) shows how one sloppy interpretation is licensed. This interpretation requires that John's coach's contract be approved of by John's coach and Bill's contract be approved of by Bill. (13c) is a representation where deletion cannot be licensed because there is no suitable antecedent for deletion licensing.

(13) a. John's coach σ_1 (approved of *x*'s contract) and Bill $\sim_i$ (did Δ (approved of *x*'s contract))
b. John's coach σ_1 (λ_x. *x* approved of *x*'s contract) and Bill $\sim_1$ (λ_y. *y* did Δ (approve of *y*'s contract))
c.* John λ_x. *x*'s coach approved of *x*'s contract and Bill λ_y. *y* did Δ approve of *y*'s contract)

The failure of deletion licensing in (13c) accounts for the absence of an interpretation of (12) that is paraphrased as 'John's coach approved of John's contract and Bill approved of Bill's contract'. Note that such an interpretation is available if the subject of the second conjunct of (12) is changed to *Bill's coach*, which is expected since then *John* and *Bill* are in parallel positions. Example (12) shows how Rooth's analysis predicts the "parallel dependencies" requirement on sloppy readings of Fiengo and May (1994). Fiengo and May argue that, generally, sloppy interpretations are only possible if the binders of the two pronouns are in parallel positions as in (13b). The parallel dependencies requirement will become important in section 4.3.

Rooth's account of the parallel dependencies requirement has another part. Consider representation (14), which has the same interpretation as (13c). In (14), however, deletion could be licensed as indicated because the variables in the two VPs in (14) are coindexed.

(14) John λ_x. *x*'s coach σ_1 (approved of *x*'s contract) and Bill λ_x. *x* $\sim_1$ (did Δ (approve of *x*'s contract))

Rooth assumes with Sag (1976) and others that the coindexation shown in (14) is blocked. Heim (1997a) states this prohibition as in (15). In this way, Rooth's approach predicts the lack of the interpretation which the parallel dependencies requirement was meant to capture.

(15) *No Meaningless Coindexing:*
If an LF contains an occurrence of a variable *v* that is bound by a node α, then all occurrences of *v* in this LF must be bound by the same node α. (Heim 1997a, (24))

Rooth (1992a) and Fox (1999a) argue that sometimes the antecedent argument of ~, the index *n*, itself does not satisfy condition (9b), but an

entailment of $g(n)$ does (i.e., there is a β such that $g(n)(h)$ entails $\beta(h)$ and $\beta(h) \in [\![X]\!]^h_f$ for all assignment functions h).[6] In such cases, I will speak of *licensing via an entailment* in this paper. Consider Rooth's example in (16):

(16) First, John told Mary I was bad-mouthing her, and then Sue heard I was. (Δ = bad-mouthing her)

The representation of (16) for deletion licensing is the following:

(17) First σ_1 (Mary λ_z John told x I was bad-mounting x), and then $\sim_1$ (Sue_F λ_y y heard I was bad-mouthing y).

In (17), (9b) is not directly satisfied. But, the focus value of the complement of ~ contains the proposition *Mary heard I was bad-mouthing her*, which is an entailment of the argument of σ_1. Hence, licensing via an entailment is satisfied in (17).

Licensing via an entailment cannot apply freely, however. For example, deletion is not licensed in (18) though the first conjunct in (18a) entails that Kai read more than one book, and the first conjunct in (18b) entails that a cup broke.[7]

(18) a.* Kai has read almost every book, and Lina_F has Δ too. (Δ = read more than one book)
b.* First Kai broke a cup and then the vase did Δ. (Δ = break)

Rooth and Fox suggest different versions of a restriction on licensing via an entailment, both of which block cases like (18). Rooth's suggestion is that indirect identity is blocked whenever the lexical material in the scope of Δ is not identical to the corresponding material in the scope of σ. However, Fox shows – convincingly to my mind – that Rooth's lexical identity condition cannot be correct. Fox's own proposal is based on the intuition that indirect identity is licensed only in cases where the destressed material indicates which entailment needs to be drawn. However, Fox ends up with a very complicated condition. Furthermore, I am not convinced that only destressed material can indicate which entailment needs to be drawn, but rather believe that all overt material can trigger an entailment. Consider example (19):

(19) I don't have a red cent, but Kazuko_F might_F Δ. (Δ = have some money)

---

[6] The idea of licensing via an entailment can be extended to categories of all types that end in t, by existentially closing off open argument positions.

[7] Example (18b) is due to Chris Tancredi (p.c.).

In (19), the first clause contains the minimizer *a red cent*. Minimizers in the scope of negation have a stronger meaning than ordinary indefinites like *some/any money* (Krifka 1995 and others). The deleted VP, however, must contain an ordinary indefinite because there is no c-commanding negation. Hence, deletion in (19) can only be licensed by the entailment *I don't have some/any money* drawn from the first clause. But, this entailment is not indicated by destressed material, since the second clause does not contain any destressed material. I conclude therefore that in (19) the focused *might* indicates the kind of entailment that needs to be drawn.

Here then is my proposal for deletion licensing. The notion of a deletion variant in (20) captures what are possible completions of a deletion site.

(20) X' is a *deletion variant* of X iff X' is grammatical at LF and X and X' are identical except for material in the scope of Δ.[8]

The deletion licensing condition in (21) captures the intuition that whenever deletion is licensed via an entailment the completion of the deletion site must be the one closest to the antecedent. (I use the notation $\phi \rightarrow_C \psi$ to indicate that ϕ together with the common ground C entails ψ.)[9]

(21) $[\![\sim_n \mathrm{X}]\!]^g$ is defined iff there is a function α from assignments to propositions such that for all $h \in \mathit{domain}(g(n))$ all of the following hold:
a. $\mathit{domain}(g(n)(h)) \subset \mathit{domain}(\alpha(h))$
b. $\forall h \in \mathit{domain}(g(n))$: $g(n)(h) \rightarrow_C \alpha(h)$
c. $\mathit{domain}(\alpha) = \mathit{domain}[\![\mathrm{X}]\!]_f$
d. $\forall h \in \mathit{domain}(\alpha)$: $\alpha(h) \in [\![\mathrm{X}]\!]_f^h$
e. $\alpha = \max\{\beta | \forall h$: $g(n)(h) \rightarrow_C \beta(h) \wedge \exists \mathrm{X}'$: X' is a deletion variant of $\mathrm{X} \wedge \mathit{domain}(\beta) = \mathit{domain}[\![\mathrm{X}]\!]_f \wedge \forall h$: $\beta(h) \in [\![\mathrm{X}']\!]_f^h\}$

If defined, $[\![\sim_n \mathrm{X}]\!]^g = [\![\mathrm{X}]\!]^{g^{\backslash n}}$

Condition (21) generally allows licensing of ellipsis via the entailment α drawn from the antecedent $g(n)$, where I take entailment to be a proposition with weaker presuppositions ((21a)) and weaker truth conditions in

---

[8] The weak restriction to X' phrases that are grammatical at LF, rather than grammatical overall, was suggested to me by Chris Kennedy (p.c.) to make my proposal compatible with analyses that claim that PF-grammaticality conditions do not apply to PF-deleted material. For example, Merchant (2001) proposes such an account for cases where sluicing seems to make *wh*-extraction out of islands grammatical. For the following, the question whether a deletion variant must be grammatical overall or just at LF is of no relevance.

[9] For the maximum in (21e), I assume a partial order where α is greater than β iff α has stronger presuppositions and stronger truth conditions than β. Formally this means:

(i) $\alpha > \beta$ iff $\alpha \neq \beta \wedge \mathit{domain}(\alpha) \subset \mathit{domain}(\beta) \wedge \forall h \in \mathit{domain}(\alpha)$: $\alpha(h) \rightarrow \beta(h))$

the domain where the presuppositions of both of them are satisfied ((21b)). However, (21d) requires this entailment to be the maximal proposition for which the deletion can be completed at all.

Consider now examples (18) and (19) again. In (18a), the deletion variant *Lina$_F$ read almost every book* is closest to the antecedent by virtue of having a focus alternative that is equivalent to it. Since this deletion variant is also grammatical, it blocks all weaker deletion variants that might otherwise be licensed via an entailment. Similarly, in (18b) the deletion variant *[The vase]$_F$ broke a cup* blocks all other deletion variants. In (19), however, the completion *Kazuko$_F$ might$_F$ have a red cent* is ungrammatical, and therefore the next weaker deletion variant is the deletion that is licensed by (21).

Condition (21) shares with Fox's (1999a) proposal the problem that it predicts the examples in (22) to be grammatical: For example, the second conjunct in (22b) has the focus alternative *7 is equal to itself*, which is equivalent to the antecedent in the first conjunct. This problem, however, does not affect the discussion in the following as far as I can see. Possibly, its solution requires the addition of conditions to (21) that are sensitive to the syntactic content of the deleted phrase and its antecedent.

(22) a.*? First, Lina fed Miya and then Nana was Δ too. (Δ = fed by Lina)
b. * 7 is equal to 7 and 5 is Δ too. (Δ = equal to itself) (Rooth 1992a)

1.2. *Relative Clauses*

Two analyses of relative clauses play a role in this paper: the standard, externally headed one and the rivaling matching analysis of relative clauses. The externally headed analysis assumes that a null operator without any lexical content moves internal to the relative clause and leaves behind a trace that is interpreted as a plain variable. For the DP in (23a), (23b) illustrates the externally headed analysis.

(23) a. the lake Erik visited
b. the [lake Op_1 Erik visited t_1]

The alternatives to the head-external analysis are head-internal analyses of relative clauses. In Sauerland (1998), Sauerland (2003), and Hulsey and Sauerland (2003), it is argued that restrictive relative clauses are generally ambiguous between two structures, the raising and the matching structure (see also Bhatt 2002). The difference between the two is that raising involves actual movement of the relative clause head from a relative clause internal position, while matching involves phonological deletion of

an internal head licensed by the external head. For this paper, only the matching structure will be important. But in this section, I introduce both analyses and then argue for reasons why the raising structure is irrelevant for the remainder of the paper.

To state the matching analysis, a version of the σ-operator introduced in (11) that is not sensitive to indexation is needed. This is the μ-operator defined in (24):

(24) $[\![\mu_n\ X]\!]^g$ is defined if and only if $g(n)(h) = [\![X]\!]^g$ for some assignment function h. If defined, $[\![\mu_n\ X]\!]^g = [\![X]\!]^{g^{\backslash n}}$.

Now consider for example the DP in (25a): (25b) shows its raising structure, and (25c) its matching structure.

(25) a. the lake Erik visited

b. the lake [*Op t*] Erik visited *t*

c. the $\sigma_2(\mu_1$ lake) [*Op* $\sim_2(\Delta\ \mu_1$ lake)] Erik visited *t*

The raising structure (25b) is derived by two instances of movement. Underlyingly, a DP consisting of the silent relative clause operator with the NP *lake* as its complement occupies the object position in the relative clause. First, the entire DP moves to the edge of the relative clause. In a second step, the NP *lake* moves to the position of the NP that the relative clause is adjoined to. The matching structure (25c) involves only one step of movement, and one step of what Kennedy (2002) calls movement deletion.

Underlyingly, the representation of the relative clause also has the DP *Op lake* in the object position, which is then moved to the edge of the relative clause. However, a second instance of the NP *lake* occupies the position of the NP which the relative clause is attached to. The NP *lake* inside the relative clause is phonetically deleted, as indicated by the Δ-feature in (25c). The process triggering insertion of the deletion feature, movement deletion, applies obligatorily, and moreover ensures that the antecedent licensing deletion must be the NP which the relative clause is adjoined to. In (25c), this obligatory licensing dependency is captured by the coindexation of σ_2 and $\sim_2$.[10]

[10] In Hulsey and Sauerland (2003), we discuss the relation between the coindexation requirement on the μ-operators, on the one hand, and the corresponding ~ and σ operators, on the other, in more detail. In particular, we show (a) that both are needed, and (b) that coindexation of the μ-operators entails that ~ and σ must be in corresponding positions c-commanding both occurrences of μ.

Generally, movement deletion is the process stated in (26) (from Hulsey and Sauerland 2003). It follows from the deletion licensing requirements of Δ that the two phrases XP and YP which movement deletion applies to must be sufficiently similar in interpretation.

(26) *Movement deletion*:
For two phrases XP and YP, where XP c-commands YP, applying movement deletion is defined as:
(i) affix μ_i to XP (where *i* is an index that does not occur anywhere else except as inserted by (ii));
(ii) affix Δ and μ_i to YP.

One type of example motivating the matching analysis is illustrated by (27a) (see also section 5.3). In (27a), the idiomatic interpretation of *headway* is only licensed in a position outside the relative clause. Hence, the internal head of the relative clause must have a lexical content that has the same interpretation as *headway* does in the idiomatic context, but does not share its distributional restriction. I assume that *amount of progress* fills the internal head position of the relative clause.

(27) a. Bill made the amount of headway that Mary demanded. (Hulsey and Sauerland 2003, (47))
b. Bill made the σ_2 μ_1 (amount of headway) λ_x Mary demanded the$_x$ ($\sim_2$ Δ μ_1 (amount of progress))

The raising analysis is motivated in part by the existence of examples like (28), where a quantifier in the relative clause binds a variable that is part of the relative clause head like (28) (cf. Schachter 1973 and others).

(28) The picture of himself$_i$ that everybody$_i$ sent in annoyed the teacher.

In this paper, however, the raising analysis does not play a role, because raising relative clauses are generally incompatible with antecedent-contained deletion (ACD), as Wold (1995b) shows with examples like the following:

(29) *Sue likes every picture of himself$_i$ that every boy$_i$ does Δ. (Δ = like *t*)

Wold argues that binding of *himself* by *every boy* requires a representation like (30) where the relative clause head is interpreted only internal to the relative clause. But then, VP-deletion cannot be licensed in (30) because the trace position in the antecedent does not contain the material of the relative clause head (see also Hulsey and Sauerland 2003).

(30) [every λ_x every λ_y t_y like the$_x$ picture of y] λ_z Sue likes t_z
(deleted VP; antecedent)

Since I am almost exclusively concerned with ACD-structures in the following, I am going to assume that only the matching analysis of relative clauses is possible.

2. Double-Headed ACD

A major part of this paper concerns the analysis of a restriction on ACD. The restriction is demonstrated in (31), where VP-deletion is ungrammatical.

(31) * Polly visited every town in every country Erik did Δ. (Δ = visit *t*)

The restriction on ACD illustrated by (31) was first studied systematically by Kennedy (1994). The structurally similar examples in (32) show that the restriction is quite general ((32a) is repeated from (2a)).[11]

(32) a.* Polly visited every town that is near a lake Erik did Δ. (Δ = visit *t*)
b.* Polly visited every town in a country Erik did Δ. (Δ = visit *t*)
c.* Jon ordered a drink that's more expensive than the dish Sue did Δ. (Δ = order *t*)

All three examples contrast with standard ACD. For example, compare (31) with (33):

(33) Polly visited every town Erik did Δ. (Δ = visit *t*)

Kennedy (1994) argues furthermore that the type of ACD in (31) and (32) is blocked because it involves deletion – if the VP is not deleted, the sentence is grammatical, as shown in (34) in contrast with (31).

(34) Polly visited every town in every country Erik visited.

The following terminology is useful to talk about such examples: I call constructions like (31) *double-headed ACD* because two different NPs are involved in the ACD-structure. For the two NPs involved, I use the terms

[11] Example (32b) can marginally be construed as standard ACD with the relative clause construed with *town*. Example (i) does not allow this interpretation:

(i) Polly visited every town that's in a country Erik did Δ. (Δ = visit *t*)

relative clause head (RC-head) and QR-head. These two heads are indicated in (35) for (31). In normal ACD like (33) the QR-head and the RC-head are the same.

(35) *Polly visited every $\underbrace{\text{town}}_{\text{QR-head}}$ in every $\underbrace{\text{country}}_{\text{RC-head}}$ Erik did Δ.

The following three sections present three different types of LF-structures that have been proposed for ACD-structures: First, I present the structures proposed in the standard analysis of ACD of Sag (1976), May (1985), Larson and May (1990), Fiengo and May (1994), and others. I show that these do not predict any restriction on double-headed ACD. Then I present two proposals designed to do so: the proposal of Kennedy (1994) and Heim (1997a), which infers a restriction from the requirement that the indices of the two traces involved in ACD must be identical, and the proposal of Sauerland (1998), which deduces a restriction from the requirement that the lexical content of the two traces involved in ACD must be identical.

2.1. *The Problem of Double-Head ACD*

Why is the restriction on double-headed ACD introduced above surprising? To see this consider first the analysis of single-headed ACD. The problem of ACD within a theory of VP-deletion is that the antecedent VP seems to contain the deleted VP. But in that configuration, the deleted VP cannot be identical to the antecedent VP. Therefore, the antecedent VP cannot contain the deleted VP when deletion is licensed. As argued by several people (Sag 1976; May 1985; Larson and May 1990; Fox 2002, and others), the problematic antecedent containment in ACD must be resolved by movement. Specifically, the DP containing the ACD relative clause must move to a position outside of the antecedent VP. Consider again example (33), which is repeated in (36).

(36) Polly visited every town Erik did Δ. (Δ = visit *t*)

To resolve ACD in (36), the object of *visited* must move to a position outside of the VP headed by *visited*. Movement to a position between the subject and the verb would be sufficiently far, but I assume that the object moves to a position above the subject for ease of presentation. Furthermore, (37) adopts the head-external analysis of relative clauses. In (37), the deletion licensing condition (9) is satisfied: the deletion feature Δ is licensed by applying ~ and σ to the indicated constituents.

(37) [every town $\underbrace{\lambda_y \text{ Erik}_F \text{ did } [_\Delta \text{ visit } t_y]]}_{\sim_1}$ $\underbrace{\lambda_x \text{ Polly visited } t_x}_{\sigma_1}$

Now consider double-headed ACD as in (38) (repeated from (31) above).

(38) * Polly visited every town in every country Erik did Δ. (Δ = visit *t*)

Assuming movement of the object of *visited* in (38), the LF-representation of (38) is as given in (39).

(37) [every town in every country $\overbrace{\lambda_y \text{ Erik}_F \text{ did } [_\Delta \text{ visit } t_y]]}^{\sim_1}$ $\underbrace{\lambda_x \text{ Polly visited } t_x}_{\sigma_1}$

The constituents which ~ and σ apply to in (39) are identical to those marked in (37). Therefore, the deletion licensing condition (9) is satisfied in (39) just as it is in (37). However, this incorrectly predicts that deletion should also be licensed in (38), while it actually is ungrammatical.

Assuming that the standard account of deletion licensing as in section 1.1 is correct, the ill-formedness of double-headed ACD argues that the LF-representations (39) and possibly (37) as well are wrong. The question then is what aspect of these representations is wrong. The only aspect, as far as I can see, that might be wrong is the representation of the dependency of the traces and their antecedents. Any analysis of double-headed ACD along these lines, then, must abandon the standard, head-external analysis of relative clauses. And this is indeed the direction taken by the two approaches I discuss in the following: The *index identity* approach and the *copy identity* approach. The index identity approach, which Kennedy (1994) suggests and Heim (1997a) develops more fully, makes different assumptions about the indexation of traces and their binders. The copy identity approach, which I propose, assumes that traces are not bare variables, but have some internal content.

2.2. *The Index Identity Approach*

I have called the account of Heim (1997a) the index identity approach because it seeks to reduce the difference between single and double-headed ACD to a requirement that the indices of the traces in the deleted VP and

its antecedent must be identical. An important background assumption Heim uses is the prohibition on meaningless coindexing in (15). Recall that (15) requires any variable binder to have an index that no other binder or unbound variable has. However, both (37) and (39) in the previous section already satisfy restriction (15).

The core proposal of Heim (1997a) is a new syntax and semantics of variable binding that does not assume λ-operators for binding, but instead assumes that the quantificational determiners themselves bind variables. For example, the lexical entry for the universal quantifier Heim assumes is (40):[12]

(40) $\left[\begin{array}{c} \text{TP} \\ \text{DP} \quad \text{VP} \\ \text{every}_i \quad \text{NP} \end{array}\right]^g = 1$ iff. $\forall x$: $[\![\text{NP}]\!]^{g[i \mapsto x]} \rightarrow [\![\text{VP}]\!]^{g[i \mapsto x]}$

Both NP and VP in (40) would typically contain unbound variables with index i which are then bound by the quantifier *every*$_i$. Heim's (40) is intended to be just a short-hand notation for the more compositional lexical entry for *every*$_i$ in (41). The arguments R and S of *every*$_i$ must both be functions from assignment functions to truth values.

(41) $[\![\text{every}_i]\!]^g$ $(R)(S) = 1$ if and only if for every $x \in D_e$: if $R(g[i \mapsto x]) = 1$, then also $S(g[i \mapsto x]) = 1$.

With this change, Heim proposes (42) instead of (37) as the LF-representation of the single-headed ACD example (33).

(42) [every$_x$ town(x) Erik $\underbrace{[_\Delta \text{ visit } t_x]}_{\sim_1}$ Polly $\underbrace{\text{visited } t_x}_{\sigma_1}$

Because both variables are bound by the quantifier *every* they must be coindexed in (42), and therefore deletion is licensed with the domains indicated. Heim (1997a) does not discuss the syntax of relative clauses for her proposal in any detail. Perhaps it is most natural to assume that the relative clause head must always move from the relative clause internal trace position and leave behind a coindexed trace in that position.

---

[12] I use the notation $g[i \mapsto x]$ for the assignment function g' with $g'(j) = g(j)$ for any $j \neq i$, but $g'(i) = x$.

Now consider the representation of the double-headed ACD example (31) in (43).

(43) [every$_x$ town(x) in a$_y$ country(y) Erik [$_\Delta$ visited t_y]] Polly visited t_x

The variables corresponding to the traces in (43) are not coindexed and, because they have different binders, condition (15) does not permit two occurrences of the same variable in these two positions. Therefore, deletion licensing cannot be satisfied in the same way as in (42): the two VPs contain variables with different indices in the object positions. Nor is it possible to license deletion by applying ~ and σ to the constituents where these variables are bound: because the smallest constituent where t_x is bound in (43) includes the deleted VP, this type of deletion licensing is blocked by antecedent containment. In this way, Heim (1997a) makes the prediction that all double-headed ACD should be ungrammatical, while allowing the single-headed type.

2.3. *The Copy Identity Account*

My own approach to double-headed ACD is based on the idea that traces are not plain variables, but contain lexical material of the moved phrases. This assumption is inspired by work in syntax that assumes that the movement operation involves copying of a phrase, leaving an identical copy in the trace position (Chomsky 1993 and others). For this reason, I call my account the *copy identity account*. I propose that the lexical material in the traces is interpreted, and therefore deletion is only licensed when the two traces contain lexical material that has the same interpretation. Specifically, I claim that the core NP as defined in the introduction must be interpreted in the trace position. Recall that I defined the core NP of a DP as the NP-complement excluding all its adjuncts. Furthermore, I assume the matching analysis of relative clauses of section 1.2. Recall that the proposal there is that the relative clause operator must have an NP-complement that is deleted only phonologically. On the copy identity account, this unpronounced NP must be interpreted in the relative clause internal trace position. The essential idea of the copy identity account is that, by representing the QR-head and the RC-head in the trace positions, I predict that deletion will not be licensed unless the two are identical.

Again, consider single-headed ACD first ((44) is repeated from (33)).

(44) Polly visited every town Erik did Δ. (Δ = visit t)

I postulate (45) as the LF-representation. The syntactic operation *movement* in my proposal consists of three steps: Step one inserts an indexed λ-operator

λ_i at the landing site of movement. Step two copies the lexical material of the moving phrase to the landing site of movement. Step three replaces the determiner of the old copy in the trace position with an indexed determiner *the*$_i$ that is coindexed with the λ-operator inserted in the first step. Therefore, (45) has the following properties: Inside of the relative clause, the noun *town* is represented in the trace position inside of the deleted VP. Furthermore, *town* is also represented in the QR-trace inside the antecedent VP.

(45) every town [Op$_y$ Δ (town)] Erik$_F$ did Δ (visit the$_y$ town)]

$\sim_1$

σ_1

λ_x Polly visited the$_x$ town

Both occurrences of *town* in (45) are the complement of an indexed definite determiner. Recall the lexical entry in (46) (repeated from (8a)) which I assume for these.

(46) $[\![\text{the}_n]\!]^g(P)$ is defined if $P(g(n)) = 1$.
If defined, $[\![\text{the}_n]\!]^g(P) = g(n)$.

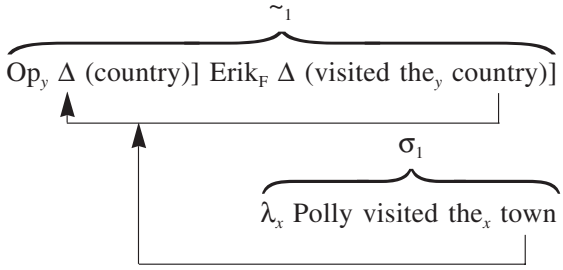
Therefore the two occurrences of *town* give rise to the presupposition that the variable *x* and *y* both denote a town. The arguments of $\sim_1$ and σ_1 in (46) both denote functions defined for the domain of towns. But since both the deleted VP and its antecedent have the same presupposition and their semantic interpretation is also the same in the relevant sense, deletion is licensed in (45).

Now consider the double-headed ACD example (47) (repeated from (31)).

(47) *Polly visited every town in every country Erik did Δ. (Δ = visit *t*)

My account entails that the LF-representation of (47) is (48). The question is whether the deletion licensing condition can be satisfied in (48) for any placement of the operators $\sim$ and σ. Specifically, consider the placement of $\sim$ and σ indicated in (48) that incorrectly allowed deletion in the standard representation (39).

(48) * [every town in every country



From the semantics of *the$_x$ town* it follows that the argument of σ_1 denotes a function that is defined only for individuals that are towns. But the argument of $\sim_1$ denotes a function that is defined only for countries. Therefore, the deletion licensing condition is not satisfied for the indicated placement of σ and $\sim$. But, then, deletion cannot be licensed in (48). In this way, the copy identity approach accounts for the ill-formedness of double-headed ACD examples like (47).

3. Arguments for the Copy Identity Account of Double-Headed ACD

There are three areas where the index identity account makes a wrong prediction whereas the predictions of the copy identity approach are borne out. I will discuss these in turn.

3.1. *Sloppy Identity in the Scope of a Binder*

Jacobson (1998) points out that the index identity account predicts deletion not to be licensed in examples like (49):[13]

(49) Every woman who John loves spoke to every woman who Bill does Δ. (Δ = love *t*) (Jacobson 1998, (33))

The LF-representation of (49) without deletion licensing operators on the index identity account is given in (50):

---

[13] Jacobson (1998) also claims that examples with pied-piping in the relative clause are problematic for the index identity account. This argument, though, is an argument not against index identity, but rather against an additional syntactic identity requirement on deletion proposed by Rooth (1992a). Fox (1999b) discusses Jacobson's pied-piping data from this perspective. Since I do not assume Rooth's syntactic identity requirement for deletion licensing, Jacobson's pied-piping data do not present a problem for either of the accounts of double-headed ACD that I laid out in the previous section.

(50) every$_x$ [woman(x) $\wedge$ John loves x] [every$_y$ [woman(y) $\wedge$ Bill$_F$ Δ (loves y)] [x spoke to y]]

Now consider whether deletion could be licenced. Since the variable y occurs in the scope of Δ and must be construed sloppily, the ~ licensing deletion must take scope over at least the quantifier *every*$_y$ that binds the sloppy variable. But consider the placement of ~ and σ in (51):

(51) $\overbrace{\phantom{*\text{every}_x\ [\text{woman}(x) \wedge \text{John loves } x]}}^{\sim_1}$
$*\text{every}_x\ [\text{woman}(x) \wedge \text{John loves } x]$
$\underbrace{[\text{every}_y\ [\text{woman}(y) \wedge \text{Bill}_F\ \Delta\ (\text{loves } y)]}_{\sigma_1}\ [x \text{ spoke to } y]]$

With the lexical entry for (41), it is clear that the presupposition of $\sim_1$ in (51) is not satisfied because the arguments of $\sim_1$ and σ_1 are different quantifiers. For example, the two yield different results if applied to the function defined in (52):

(52) λg: $g(x)$ is a woman and $g(y)$ is man

The argument of σ_1 is true of (52), whereas the argument of $\sim_1$ is false. But any placement of ~ and σ other than that in (52) stands no chance of licensing deletion because placing ~ and σ higher than in (52) results in overlap of their arguments, whereas placing them lower is ruled out by the sloppy variable. Therefore, the index identity account incorrectly predicts (50) to be ungrammatical. The problem is quite general: index identity incorrectly blocks sloppy deletion in all cases where the deleted phrase occurs in the scope argument of the phrase that binds the sloppy variable in the antecedent.

A possible way out of the problem created by (50) would be to stipulate a difference between the restrictor and the scope of the quantifier. This approach would rely on the lexical entry for *every*$_i$ in (53) rather than (41). In (53), only the restrictor argument of the quantifier is a function from assignments to truth values, whereas the scope argument is a function from individuals to truth values, which then can be an ordinary λ-abstract.

(53) $[\![\text{every}_i]\!]^g$ $(R)(S) = 1$ if and only if for every $x \in D_e$: if $R(g[i \mapsto x]) = 1$, then also $S(x) = 1$.

The representation in (54) for (49) explains why deletion is licensed.

(54) $\overbrace{\text{every}_x\ [\text{woman}(x) \wedge \text{John loves } x]\ \lambda_z}^{\sim_1}$
$\underbrace{[\text{every}_y\ [\text{woman}(y) \wedge \text{Bill}_F\ \Delta\ (\text{loves } y)]\ \lambda_v\ [z \text{ spoke to } v]]}_{\sigma_1}$

However, the different treatment of scope and restrictor in (54) is not independently motivated, as far as I can see. In fact, it creates a new problem with single-headed ACD: if only the lexical entry of *every* in (53) were available, the LF-representation predicted for single-headed ACD would be (55).

(55) $[\text{every}_x\ \text{town}(x)\ \text{Erik}\ \Delta\ (\text{visited } t_x)]\ \lambda_y\ \text{Polly visited } t_y$

Deletion in (55) cannot be licensed, and to explain single-headed ACD Heim's lexical entry for *every* in (41) is needed in addition to (54). It is undesirable, however, to have to assume two lexical entries for *every*. Hence, example (49) remains problematic for the index identity account.

On the copy identity account, on the other hand, (49) creates no problem. Because the sloppy variable is bound within the relative clause, the deletion licensing requirement is satisfied in (56).[14]

σ_1

(56) $\text{every}_x\ [\text{woman} \cap \lambda_x\ \text{John loves the}_x\ \text{woman}]$
$\lambda_z\ [\text{every}_y\ [\text{woman}(y) \cap \underbrace{\text{Bill}_F\ \Delta\ (\text{loves the}_y\ \text{woman})]}_{\sim_1}$
$\lambda_v\ [\text{the}_z\ \text{woman spoke to the}_v\ \text{woman}]]$

3.2. *Comparatives*

Heim (1997b) points out that index identity cannot account straightforwardly for deletion in comparatives. Consider example (57).

(57) Mike drives a more powerful car than Pete does Δ. (Δ = drive t)

---

[14] In Jacobson's example (49) the heads of the two relative clauses are lexically the same, which makes this kind of example particularly easy to account for on the copy identity approach. In section 4, I consider similar examples where the heads of the two relative clauses are not identical.

Comparatives like (57) involve ACD. The standard analysis of comparatives (see Wold 1995a; Lechner 1999; Heim 2000 for recent defenses of this view) assumes that ACD in comparatives is resolved by movement of the degree quantifier *more* and the *than*-clause to yield a structure like (58). Here, *more* can be understood as the following degree quantifier: $[\![\text{more}]\!](R)(S) = 1$ if and only if $\max\{d|R(d) = 1\} < \max\{d|S(d) = 1\}$.

(58) more $\overbrace{[\lambda_d \text{ Pete}_F \Delta \text{ (drives a } d \text{ powerful car)}]}^{\sim_1}$
$\underbrace{[\lambda_{d'} \text{ Mike drives a } d' \text{ powerful car}]}_{\sigma_1}$

On the basis of representation (58), comparative ellipsis is licensed by the placement of ~ and σ. This account is compatible with the copy identity account of ACD, independently of the question how the two traces of degree movement in comparatives, d and d' in (58), are interpreted. They could both be bare variables as in (58), or they could contain some additional lexical content that might, for example, be interpreted as a presupposition that d refer to degree of a certain type. Crucially, this lexical content will be the same in both trace positions.

Now consider the index identity account. For deletion to be licensed the degree variables in the matrix clause and the *than*-clause must be coindexed. Therefore, the index identity account is committed to representation (59):

(59) more_d [Pete_F $\overbrace{\Delta \text{ (drives a } d \text{ powerful car)}}^{\sim_1}$]
[Mike $\underbrace{\text{(drives a } d \text{ powerful car)}}_{\sigma_1}$]

With the lexical entry for $more_d$ in (60), the right interpretation is predicted for (59).

(60) $[\![\text{more}_d]\!]^g(R)(S) = 1$ iff. $\text{Max}\{\delta\text{: } R(g[d \mapsto \delta]) = 1\} < \max\{\delta'\text{: } S(g[d \mapsto \delta']) = 1\}$

However, there is a difference between (60) and Heim's lexical entry for $every_i$ in (41). While the two arguments of *every* in (41) are applied to the same modified assignment function, the two arguments of *more* in (60) are applied to two different modified assignment functions that may assign

different values to *d*. Hence, coindexation of the two occurrences of *d* in (59) does not reflect any semantic connection between these two variables, but is merely formal.

Note that formal coindexation could also be achieved in double-headed ACD, as in (61a) (repeated from (31)), if it is unconstrained. Consider the representation in (61b):

(61) a.* Polly visited every city in every country Erik did.
b. [every city in every]$_x$ [country(x) $\wedge$ Erik visited t_x] Polly visited t_x

A lexical entry for the part *[every city in every]*$_x$ of (61) that predicts the correct interpretation is given in (62).

(62) $[\![(\text{every city in every})_x]\!](R)(S) = 1$ iff $\forall a\ \forall b$: city(a) $\wedge$ a is in $b \wedge R(g[x \mapsto b]) = 1 \wedge S(g[x \mapsto a]) = 1$

But, (62) must be impossible since otherwise double-headed ACD is predicted to be always acceptable. This shows that the syntactic differences between *more* and *a city near a lake* are crucial for the success of the index identity account. Furthermore, the index identity account must assume that the co- and contraindexation of variables bound by an operator must be determined solely by the syntactic properties of the operator, independent of its semantics.

3.3. *The Effect of Lexical Content*

The most striking difference in the predictions of the index identity and copy identity accounts concerns the role of the lexical content of the two binders. On the index identity account, this is expected to have no effect on deletion licensing. On the copy identity account, on the other hand, an effect on deletion licensing is predicted.[15] In this section, I will argue that the

[15] Lappin (1984) proposes a more general condition on deletion licensing that would also predict (63). As I understand him, Lappin proposes that two traces or pronouns are identical for the purposes of deletion licensing if they "can be naturally interpreted as having the same intended range of possible values" (Lappin 1984, (10)). He does not discuss contrasts like (63), however. Furthermore, Lappin's general condition makes wrong predictions in many cases of the type which he actually does discuss, as was pointed out by Fiengo and May (1994, p. 229). For example, (ib) and (iib) are just as acceptable as Lappin's (ia) and (iia), but violate his condition.

(i) a. Here is the man who Bill saw, and her is the man who he did not ⟨see⟩. (Lappin 1984, (21b))
b. Here is the man who Bill saw, and here is the woman who he did ⟨see⟩.

prediction of the copy identity account is borne out. One contrast relating to the lexical content of the binders is (63), which is repeated from (2) in the introduction of the paper.

(63) a.* Polly visited every town that is near the lake Erik did Δ. (Δ = visit *t*)
b. Polly visited every town that is near the one Erik did Δ. (Δ = visit *t*)

Recall that with the notation introduced in (35) the QR-head in (63a) and (63b) is *town*, while the RC-head in (63a) is *lake* and in (63b) *one*. I claim that the generalization underlying (63) is as stated in (64).

(64) Double-headed ACD is acceptable whenever the QR-head and the RC-head have the same interpretation.

Single-headed ACD of course also satisfies generalization (64) because the QR-head and the RC-head are the same.

The facts presented in the subsections below corroborate generalization (64). After establishing (64), I shall go on to show that it supports the copy identity account.

3.3.1. *Partitives*

One additional piece of support for (64) is provided by partitives. Heim (1997b) already points out that the fact that partitives allow ACD is a problem for the index identity account. She considers partitive constructions like (65), where the noun *towns* is pronounced only once.

(65) Erik visited two of the towns Polly did Δ. (Δ = visit *t*)

Heim observes that the licensing of deletion in (65) can only be explained on the index identity account by assuming that *two of the* is a complex

---

(ii) a. [Every friend of John's]$_i$ wants Mary to kiss him$_i$, but [none of the little fellows]$_j$ believes that she will ⟨kiss him$_j$⟩. (Lappin 1984, (10))
b. [Every friend of John's]$_i$ wants Mary to kiss him$_i$, while [every friend of Bill's]$_j$ wants Sue to ⟨kiss him$_j$⟩.

Since the deleted VPs in (i) and (ii) contain not traces but bound pronouns, my copy identity account predicts no restriction on deletion here. I return to the discussion of examples where the deleted VPs contain traces other than double-headed ACD in section 4.

determiner.[16] Then (65) could be analyzed as single-headed ACD as shown in (66):

(66) [two of the]$_x$ [towns(x) ∩ Polly visited x] [Erik visited x]

Recall though that, while allowing (66), the index identity account must still rule out the analysis of *[every city near every]*$_x$ as a complex determiner in (62). What could the relevant difference be between these two complex determiners? Also relevant for this question are examples like (67), where the noun *towns* is repeated in the partitive.

(67) Erik visited no towns of all the towns Polly did Δ. (Δ = visit t)

In (67), ACD still is acceptable. Therefore, the index identity account would force us to assume that *[no towns of all the]*$_x$ is a complex determiner. I give a lexical entry for it in (68):

(68) $[\![[\text{no towns of all the}]_x]\!]^g(R)(S) = 1$ if defined if there is an individual t such $R(g[x \mapsto t]) = 1$ where defined: $\{t|\text{town}(t) \wedge S(g[x \mapsto t])\} = \varnothing$

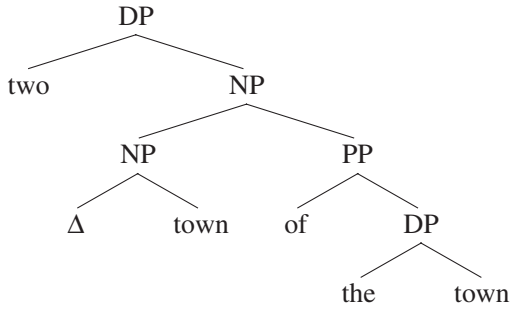
But it seems unlikely to me that *no towns of all the* is a determiner, while *a city near a* is not. Note also that, at least in German, the syntactic constituency of partitives argues against the complex determiner analysis assumed in (66) and (68). This is shown by the fact that German allows fronting of the PP *of the cities* in example (69).

(69) [Von welchen Städten]$_i$ hat Erik [zwei t_i] besichtigt?
of which towns has Erik two visited

In sum, it is unclear whether the index identity account can account for the possibility of ACD with partitives. Note that, on the other hand, ACD in partitives is predicted to be good by generalization (64). I assume that the *of*-phrase in partitives is always an adjunct to an NP that can be deleted. For example, (70) shows the structure of *two* Δ *of the towns*, with Δ = *towns*.

[16] Because indexed variable binders can only be introduced by lexical rules on Heim's proposal, it follows that *two of the* must be one lexical entry. Though I will not execute this here, it is probably possible to restate Heim's proposal such that binder indices have their own lexical entries which take three arguments; the quantifier, the restrictor, and its scope. On such a restatement, *two of the* would not have to be analyzed as a single lexical item, but it would still need to be a constituent.

(70)



In (70) *of the towns* forms a constituent, as suggested by the German example (69).

Since the QR-head and the RC-head are identical in interpretation even if one of them is deleted in the phonology, generalization (64) is satisfied.[17] Therefore, ACD with partitives is correctly predicted to be grammatical.

3.3.2. *More Double-Headed ACD*

Further, unequivocal support for generalization (64) comes from examples like (2), which I repeat in (71). The same effect is also present in (72), where the RC-head is part of a PP-modifier restricting the QR-head.

(71) a.* Polly visited every town that is near the lake Erik did Δ. (Δ = visit *t*)
b. Polly visited every town that is near the one Erik did Δ. (Δ = visit *t*)

(72) a.* John visited every town near a lake Mary did Δ. (Δ = visit *t*)
b. John visited every town near the one Mary did Δ. (Δ = visit *t*)

Both sentences in (72) marginally allow an interpretation of the deleted clause as *visited every town near*. This, however, is a form of single-headed ACD.

While the RC-head in (2) and (72) is *one*, most speakers also perceive a clear improvement in (73), where the same noun is repeated (vs. (72a)), and in (74a) (vs. (74b)).

(73) John visited every town near a town Mary did Δ. (Δ = visit *t*)

---

[17] NP-deletion in (70) needs to be licensed semantically by the ~-operator as discussed in section 1.1. Then, the QR-head in (70) is actually σ_1 *towns*, whereas the RC-head is σ_1 *towns*. This, however, is still consistent with generalization (64) because the interpretation of both heads is the same.

(74) a. John visited every town that's near a town Mary did Δ. (Δ = visit *t*)
b.* John visited every town that's near a lake Mary did Δ. (Δ = visit *t*)

The example in (75) corroborates that there is a contrast between, on the one hand, a *one*-anaphor in (75a) and repetition of the same head in (75b) and, on the other hand, different head nouns in (75c).

(75) a. Jon ordered a drink that's more expensive than the one Martin did Δ. (Δ = order *t*)
b. Jon ordered a drink that's more expensive than the drink Martin did Δ. (Δ = order *t*)
c.* Jon ordered a drink that's more expensive than the dish Martin did Δ. (Δ = order *t*)

For those speakers that accept (73), (74a), and (75b), these sentences require a particular pronunciation – namely, the RC-head must be destressed. This effect is predicted by my claim in (64) that the two heads must have the same interpretation in double-headed ACD: consider an example like (76), where focus on the second occurrence of the noun *town* is quite natural.

(76) Stuttgart is a town, but TOKYO_{F} is a TOWN_{F}.

The effect of focus in (76) is to contrast two concepts of town, one that includes Stuttgart and one that does not. A natural proposal for (76) would therefore be that there are two lexical entries for *town* corresponding to these two concepts. Since focus can only be licensed by means of these two lexical entries, it's expected that focus on one of the heads of double-headed ACD bars the two NPs from having the same interpretation, and therefore interferes with deletion licensing.

A minority of speakers I consulted does not accept examples with repetition of the same head ((73), (74a), (75b)), though they still accept the examples with *one*-anaphora ((2), (72), and (75a)). One possible reason might be that for these speakers repetition of the same noun is only possible when there is a semantic contrast between the two heads. This would interfere with deletion licensing in the way just discussed.

3.3.3. *Sensitivity to Interpretation*

This section presents further support for the claim that double-headed ACD is sensitive to the interpretation of the two heads involved. Specifically, I show that double-headed ACD improves when the two heads are close

in interpretation, even when they are not lexically identical. Consider the examples in (77) and (78):

(77) (Δ = live in *t*)
a. John lives in a city that's close to a city Mary used to Δ.
b. ?John lives in a city that's close to where Mary used to Δ.
c.?? John lives in a city that's close to a town Mary used to Δ.
d. *John lives in a city that's close to a castle Mary used to Δ.

(78) (Δ = order *t*)
a. Jon ordered a drink that's more expensive than the drink Sue did Δ.
b. ?Jon ordered a drink that's more expensive than what Sue did Δ.
c.?? Jon ordered a cocktail that's more expensive than the beer Sue did Δ.
d. * Jon ordered a drink that's more expensive than the dish Sue did Δ.

The examples in (77b), (77c), (78b), and (78c) show that double-headed ACD constructions are quite acceptable even when the two heads are not identical, but still close in meaning. As discussed above for the case of lexical identity, this observation only obtains if the relative clause head is not focused. The acceptability is slightly greater for (77b) and (78b) than for (77c) and (78c). This corresponds to another difference between the (b) and the (c) examples: the RC-head in (77b) and (78b) denotes a superset of the QR-head, whereas in (77c) and (78c) the RC-head does not denote a superset of the QR-head but a set from the same semantic field.

Although it is apparent that examples where identity of interpretation is not satisfied are somewhat degraded, the fact that these examples have an intermediate acceptability status supports the claim that licensing of double-headed ACD must take into account the interpretation of the two heads involved.

3.3.4. *Explaining the Effect of Lexical Content*

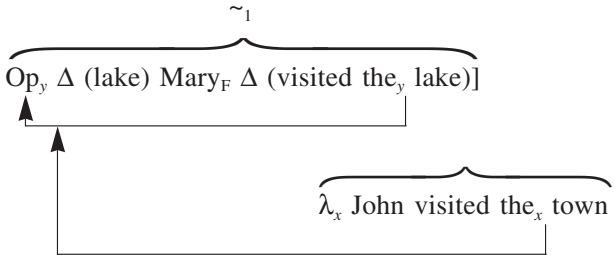
Generalization (64) provides a further argument for the copy identity account of double-headed ACD. As I will show now, the generalization follows straightforwardly from the copy identity account. The index identity account, on the other hand, has no way of accounting for (64), as far as I can see.

For this argument I use the two examples in (74), which are repeated in (79), as a concrete instance where generalization (64) is attested.

(79) a. John visited every town that's near a town Mary did Δ. (Δ = visit *t*)
b.* John visited every town that's near a lake Mary did Δ. (Δ = visit *t*)

Recall from section 2.3 how double-headed ACD in (79) is ruled out: the copy identity account relies on an LF-representation like (80) for (79a).

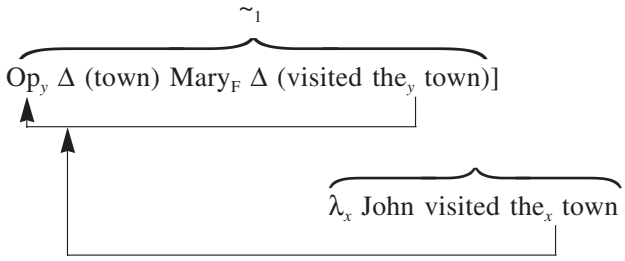
(80) * [every town that is near a town



I showed in section 2.3 that VP-deletion cannot be licensed for any placement of ~. In particular, for the placement of ~ and σ shown in (80), deletion licensing is not satisfied because the argument of σ is defined only for individual x with town(x) = 1, while any element of the focus value of the argument of ~ is defined for x if and only if lake(x) = 1.

The account of the well-formed (79a) is now straightforward. Consider the LF-representation of (79a) in (81):

(81) [every town that is near a lake



The only difference between (81) and (80) is that all occurrences of *lake* are replaced by *town*. But, this has the consequence that the elements of the focus value of the argument of ~ in (81) are defined for any individual x with town(x) = 1. Since, furthermore, *John* is a focus alternative to *Mary*, the deletion licensing condition is satisfied in (81).

Note that the argument of ~ and σ in (81) are the same as they are in single-headed ACD (see (45)). My account therefore predicts double-headed

ACD in examples like (79a) to be equally acceptable as single-headed ACD. This is indeed the case for the partitive examples in section 3.3.1 and the examples with *one*-anaphora. However, examples like (79a), where the same NP is repeated, seem to be slightly degraded even for those speakers that perceive a contrast with examples like (79b). I already suggested above that, for the minority of speakers who perceive no contrast in pairs like (79), the repetition of the head noun *town* is only possible if they assume two different lexical entries for the word *town*. Deletion licensing is predicted to be blocked in this case just as it is in (79b). Hence the markedness of examples like (79a) to all speakers should not be taken to argue against my analysis. Rather, it suggests that even for those speakers that allow repetition of the head noun with the same interpretation, this construal is in fact marginal.

The account offered for (79) also carries over to examples where one of the head nouns is deleted, as in partitives, or ones with a *one*-anaphor. In these cases, the two heads of the double-headed ACD construction have the same interpretation, and therefore the purely semantic deletion licensing condition (9) is satisfied just as it is in (81). This is independent of the the question whether *one*-anaphora are analyzed as NP-deletion or as NP-pronominals, since the resulting interpretation will be the same.

Finally, consider the examples in the previous section where the two head nouns are not identical, but close in meaning. Specifically, consider (78b) and (78c). In example (78b), the ACD-relative clause is a free relative. At present, the syntactic analysis of free relatives is controversial (see van Riemsdijk 2000). I consider one possible analysis for concreteness: I assume that free relatives only have an internal head which is not subject to movement deletion. For (78b), specifically, I assume that *what* can be analyzed as *which thing*. (82) shows the parts of the LF-representation of (78b) that are relevant for deletion licensing on this view.

(82) $\ldots \underbrace{\lambda_x \text{ Sue}_F \Delta \text{ (ordered the}_x \text{ thing)}}_{\sim_1} \ldots \underbrace{\lambda_y \text{ Jon ordered the}_y \text{ drink}}_{\sigma_1}$

The deletion licensing condition (9) is not directly satisfied because the argument of σ has a stronger presupposition than the argument of $\sim$ does. However, (82) is semantically different from (80): in (82), the argument of σ denotes a function that is defined for a proper subset of the domain that the functions in the focus set of the argument of $\sim$ are defined for. And since *Jon* is a focus alternative of *Sue*, this focus set contains one element that is an extension of the function that the argument of $\sim$ denotes. In (80), on the other hand, these two functions are defined for disjoint

domains. I would like to suggest that for this reason deletion in (82) is only slightly degraded.

In (78c), the QR-head *cocktail* and the RC-head *beer* are different, but from the same semantic field. Because there is no subset relationship between the two nouns, deletion cannot be licensed in the same way as just sketched for (78b). Since (78c) is judged less grammatical than (78b) by my consultants, this is a good result. Still, the difference between (78c) and the fully ungrammatical examples of double-headed ACD discussed at the beginning needs to be accounted for. I would like to tentatively suggest that speakers can marginally interpret the words *beer* and *cocktail* both as *some kind of alcoholic beverage*. This suggestion seems justifiable to me since we often do not care about the precise lexical content of what is said in a conversation. For example, imagine a barkeeper asking *Do you want a Caipirinha?*. Someone like me, who doesn't know what a Caipirinha is, would nevertheless understand the inquiry to be about some specific alcoholic beverage. Accepting this suggestion, the account of the marginal possibility of deletion is straightforward: deletion licensing is satisfied because under the marginal construal the two words have the same meaning.

4. Circumventing Copy Identity

The previous section has shown that traces contain lexical material and that this lexical material is important for deletion licensing. The evidence that led to this conclusion came from double-headed ACD. But a look at other constructions with a deleted constituent that contains a trace shows hat there are examples where deletion is licensed even when the traces contain different lexical material. Consider the two examples in (83):

(83) a. I know which cities you visited *t*, but I don't know which lakes you did Δ. (t = the$_x$ cities, Δ = visited the$_y$ lake)
b. I know the cities you visited *t* but I don't like the lakes you did Δ. (t = the$_x$ cities, Δ = visited the$_y$ lakes)

The history of discussion of *wh*-extraction out of a deleted VP is actually very confusing. In the first systematic studies of VP-deletion, Sag (1976, pp. 63–67) and Williams (1977, pp. 130–131) claim that *wh*-extraction from a deleted VP is impossible except for ACD. A number of years later, Evans (1988) and Jacobson (1992) show that extraction from deleted VPs is possible in many more cases. Since I adopt the latter position, I shall consider the arguments that were given for the former. Some of the examples from Sag's and Williams's work are given in (84). I would like to claim that deletion is ruled out in each of them for reasons not of concern here. In

(84a), since the auxiliary moves to Comp, the deleted VP is in the complement position of an empty head, which is generally impossible (cf. Lobeck 1995).[18] For (84b) and (84c), Fiengo and May (1994, p. 244) suggest a preference to delete as much material as possible if any material at all is deleted.[19] This suggestion has a problem with data like (83a) because with them, sluicing is also possible, and would delete *you did* in addition to the material already deleted. To rescue Fiengo and May's condition we stipulate that it applies only in case the phrase extracted from the two deleted phrases is identical.

(84) a. What did Harry take a picture of?
*What did Bill Δ? (Δ = take a picture of *t*) (Sag 1976, (1.3.18))
b. *John who Bill saw and who Bob did Δ, too. (Δ = see *t*) (Williams 1977, (93))
c.?? We finally got in touch with John, who my brother Al tried to visit, but who he couldn't Δ. (Δ = visit *t*) (Sag 1976, (1.3.22))

I conclude, therefore, that deletion in (84) is blocked for independent reasons that are of no concern for my present purposes. The examples in (83), however, are directly relevant for my claim that lexical material in the trace positions is relevant for deletion licensing. These examples contrast with the double-headed ACD cases discussed in the previous section. What is it that causes this difference?

In this section, I show why deletion in (83) is expected to be licensed. I will argue that the lexical material in the trace positions in (83) can be a bound variable, and therefore a sloppy interpretation is possible. Hardt (1999) and Schwarz (1999a, b) have argued that deleted lexical material can

---

[18] Note that (i) without auxiliary movement contrasts with (84a).

(i) I know what Harry took a picture of, but I don't know what Bill did Δ. (Δ = take a picture of *t*)

Shoichi Takahashi (p.c.) points out that VP-deletion is possible, however, in yes/no questions as in (ii):

(ii) a. Did you go to the party?
b. No. I didn't. Did you?

This interesting difference between (ii) and (84a) deserves further investigation. One possible analysis that comes to mind is that (iib) doesn't involve auxiliary movement, but instead movement of the TP *did* Δ, which leaves the deleted VP in the complement position of the auxiliary. In (84a), by contrast, such phrasal movement would block object extraction.

[19] I assume that (i) with across-the-board movement is the relevant alternative with maximal deletion that blocks (84b).

(i) John who Bill saw and Bob did too.

always receive a bound variable interpretation. Schwarz's argument is based on the following example:

(85) When I whistle you say I shouldn't Δ_1, but when I sing you don't Δ_2. (Δ_1 = whistle, Δ_2 = say I shouldn't sing)

I will furthermore argue that this type of sloppy interpretation is subject to the dependency parallelism condition of Fiengo and May (1994) (cf. section 1.1), as is expected in view of the analysis. This requirement, as I will show, blocks deletion in double-headed ACD. I also show that the effect of copy identity reemerges in some cases of non-ACD deletion.

The structure of this section is as follows: I first summarize Schwarz's (1999a) proposal in 4.1. Section 4.2 argues that his proposal predicts deletion to be licensed in (83). In section 4.3, I show how the parallelism requirement applies to sloppy deleted material, and argue that this requirement blocks sloppy deleted material to occur in double-headed ACD, as well as some other cases.

4.1. *Sloppy Deleted Material*

Consider example (85) in more detail. Schwarz calls attention to the interpretation of (85) where the first deleted VP is interpreted differently from the corresponding deleted material in the second deleted VP: the former contains *whistle*, the latter *sing*. The account of deletion licensing I adopted (see section 1.1) does not predict this interpretation to be available. To verify this, consider representation (86).[20]

(86) * When I whistle you say I shouldn't$_F$ Δ (whistle)
σ_1 $\sim_1$
σ_3
but when I sing$_F$ you don't$_F$ Δ (say I shouldn't Δ (sing))
σ_2 $\sim_2$
$\sim_3$

The deletion licensing condition is not satisfied in (86) for two reasons. For one, the argument of σ_3 in (86) is only defined for assignments *g* that

[20] There is no direct evidence that the verb *sing* in (86) must bear a deletion mark since it is in the c-domain of another Δ. It is easy to see, though, that even without this deletion mark and the licensing $\sim_2$, $\sim_3$ is not licensed.

assign the proposition that 'I sing' to index 1. The argument of $\sim_3$, on the other hand, is defined for any value of $g(1)$, but presupposes that $g(2)$ be the proposition that I sing. Therefore, the part of the deletion licensing condition that requires identical presuppositions is not satisfied. Secondly, the argument of σ_3 denotes a proposition that implies that you have said at some time that I should not sing. But, all the propositions in the focus set of the argument of $\sim_3$ entail that you talked about whether I should whistle, because the deleted occurrence of *sing* in (86) is not focused. Hence, it seems that the account of deletion licensing I adopted in section 1.1, together with representation (86), wrongly predicts (85) to not have this interpretation.

The problem raised by (85) has been discussed in three places in the literature. Kratzer (1991) first discovered a different version of the problem, but her account could probably be extended to (85). Hardt (1999, pp. 204–206) and Schwarz (1999a) independently propose accounts for (85) that are very similar to each other. In the following, I adopt Schwarz's account for concreteness, as it is the most simple to present here.[21] Schwarz's account leaves the deletion licensing condition unchanged, but assumes that there is a second kind of silent VP, namely, silent variables denoting VP-meanings. I indicate these variables as indexed VP-pro. Schwarz analyzes examples like (85) as cases of sloppy identity of such a VP-variable. The binder of the VP-variable must be the occurrence of the same VP in the *when*-clause. The VP does not c-command the VP-pro it binds, but Schwarz reminds us that sloppy readings are generally not constrained by c-command, and any explanation of that phenomenon should carry over to the case of VP-variables (see Hardt 1999; Tomioka 1999). For concreteness, I assume that the antecedent VP can move to a position where it c-commands the bound VP, as shown in (87).

(87) $\underbrace{\text{whistle } \lambda_4 \text{ when I } t_4 \text{ you say I shouldn't VP-pro}_4}_{\sigma_3}$
but $\underbrace{\text{sing}_F\ \lambda_5 \text{ when I } t_5 \text{ you don't}_F\ \Delta \text{ (say I shouldn't VP-pro}_5\text{)}}_{\sim_3}$

In (87), the deletion licensing condition is satisfied in the same way as in standard sloppy interpretations. Since at present I know of no real alternative proposal that can account for example (85), I adopt Schwarz's proposal. However, I believe that any proposal that can account for (85)

[21] In Sauerland (1998), I adopt an analysis closer to Kratzer's (1991) proposal.

is going to be compatible with the account of extraction out of deleted phrases that I propose in the next section.

One dissatisfying aspect of Schwarz's (1999a) proposal that he notes is the following (as far as I can see, Hardt's (1999) proposal has the same problem). The VP-pro, Schwarz postulates, behaves syntactically not like a pronoun, but like a full representation of the antecedent VP: extraction from VP-pro is possible, as (88a) shows.[22] Furthermore, (88b) shows that Condition B applies to the object.

(88) a. When I've reviewed a book I remember [which Δ_1] Δ_2, but when I've only read a book I don't Δ_3.
(Δ_1 = book, Δ_2 = I have reviewed *t*, Δ_3 = remember which book I have read)

b.* When you talk to me you say I shouldn't Δ_1, but when you email me you don't Δ_2.
(Δ_1 = talk t me, Δ_2 = say I shouldn't email me)

Since I am presently not aware of a solution to this problem, and have none to offer myself here, I simply note that the variables Schwarz postulates behave this way.[23]

4.2. *Sloppy Deleted Material in Traces*

I propose that the lexical content of traces can also be a silent variable denoting an NP-meaning, and I will use the notation NP-pro for this. Consider now how this proposal solves the problem raised at the beginning of this section. First consider example (83b), which is repeated in (89):

(89) I like the cities you visited *t* but I don't like the lakes you did Δ. (*t* = the$_x$ cities, Δ = visited the$_y$ lakes)

My proposal is that both object traces in (89) contain a silent NP-pro that has the same properties as Schwarz's VP-pro. In fact, Schwarz and Hardt both already note that silent variables of categories other than VP are motivated by their analysis. Schwarz points to the IP-deletion in sluicing

---

[22] Schwarz (1999a) gives an example of sluicing with *why*. I use the example in (88a) because it is sometimes claimed that *why* can be base-generated in Spec(CP) (see, for example, Tsai 1994).

[23] In a presentation at the SALT 10 conference, Mats Rooth formulated a proposal which in essence allows the variables motivated by Schwarz to range over entities that are as richly differentiated as syntactic structures. This seems to be just one possible direction to pursue.

cases such as (88), but (90) shows that NP-deletion too allows for sloppy deleted material.

(90) When Lina is eating one candy, Kai wants two Δ_1, but when she is eating one apple, he doesn't Δ_2. (Δ_1 = candies, Δ_2 = want two apples)

For example (89) I propose that, in both conjuncts, the head of the relative clause is the binder of NP-pro. Since I assume that binding is accomplished by covert movement of the antecedent NP, I arrive at the representation in (91) for (89).

(91) $\underbrace{\text{I like cities } \lambda_1 \text{ the } [t_1\ \lambda_x \text{ you visited the}_x \text{ NP-pro}_1]}_{\sigma_3}$,
but $\underbrace{\text{I don't}_F \text{ like lakes}_F\ \lambda_2 \text{ the } [t_2\ \lambda_y \text{ you did visit the}_y \text{ NP-pro}_2}_{\sim_3}$

Deletion licensing is satisfied in (91) with the indicated foci in the same way in which it was satisfied in (87).

The same idea can be applied to example (83a), where deletion applies in an indirect question. I repeat this example in (92).

(92) I know which cities you visited *t*, but I don't know which lakes you did Δ. (*t* = the$_x$ cities, Δ = visited the$_y$ lakes)

In (92), the NP-complement of *which* must be the binder of the NP-pro in the trace position in both conjuncts. Hence, I arrive at the representation in (93).

(93) $\underbrace{\text{I know cities } \lambda_1 \text{ which } t_1\ \lambda_x \text{ you visited the}_x \text{ pro}_1}_{\sigma_3}$,
but $\underbrace{\text{I don't know lakes}_F\ \lambda_2 \text{ which } t_2\ \lambda_y \text{ you visited the}_y \text{ pro}_2}_{\sim_3}$.

One interesting aspect of (93) is that the lexical material in the head and the tail of the chain formed by *wh*-movement is different: the head of the chain, *which* t_1, contains a trace, but the tail of the chain, *the*$_x$ *pro*$_1$, contains a NP-pro. However, the discussion of (88) established that a pro-XP of the type Schwarz postulates behaves syntactically as if a full XP was syntactically present. Given this generalization, the apparent violation of

chain-identity in (93) is then actually not a violation, but further evidence that NP-pro is syntactically not a pronoun.

In this way, the adoption of Schwarz's (1999a) silent pro analysis explains the apparent violations of copy identity in (83) from the beginning of this section. It furthermore seems to me that the connection with Schwarz's paradigm in (85) goes beyond the specifics of his analysis: whatever account of (85) somebody will come up with in the future will also explain (83). The remaining question at this point is the following: Why can copy identity not be overcome in the same way in double-headed ACD? This the topic of the next section.

4.3. *Dependency Parallelism and the Reemergence of Copy Identity*

In section 1.1, I referred to a restriction on sloppy interpretations which Fiengo and May (1994) discovered and have called *dependency parallelism*. This requires sloppy pronouns to be bound from isomorphic positions in both the deleted VP and its antecedent. Recall also that this constraint follows from Rooth's (1992a) analysis of sloppy reading, as I illustrated by means of example (12), repeated in (94a). Dependency parallelism blocks an interpretation of (94a) that is paraphrased as 'John's coach approved of John's contract and Bill approved of Bill's contract'. On Rooth's account, this follows because in (94b) the deletion licensing condition is not satisfied, as the focus set of the argument of ~ contains only propositions with binding of the pronoun by the subject.

(94) a. John's coach approved of his contract and Bill did Δ too.

b.* John λ_x. *x*'s coach approved of *x*'s contract σ_1 and Bill_F λ_y. *y* did Δ (approve of *y*'s contract) $\sim_3$

Dependency parallelism also seems to apply to sloppy binding of VP-pro. Consider example (95):

(95) When Mary tells me to whistle, you say I shouldn't Δ_1, but when Mary sings, you don't Δ_2.
a.* Δ_1 = whistle, Δ_2 = say I shouldn't sing
b. Δ_1 = whistle, Δ_2 = say I shouldn't whistle

As indicated in (95), the example does not allow an interpretation of the second deleted VP with *sing* as a part. In example (85), a comparable

interpretation was available and, following Schwarz, I understood this to demonstrate the presence of a sloppy interpretation of VP-pro. That the sloppy interpretation is blocked in (95) follows from Rooth's proposal in the same way it did for (94): in (96), deletion licensing is not satisfied because the argument of σ_3 is not an element of the focus set of the argument of $\sim_3$, nor does it have an entailment of the right form. In representation (87) above, on the other hand, dependency parallelism was satisfied.

(96) *$\underbrace{\text{whistle } \lambda_4 \text{ when Mary tells me to } t_4 \text{ you say I shouldn't VP-pro}_4}_{\sigma_3}$
$\underbrace{\text{but sing}_F\ \lambda_5 \text{ when I } t_5 \text{ you don't}_F\ \Delta \text{ (I shouldn't VP-pro}_5\text{)}}_{\sim_3}$

Dependency parallelism explains why there is a difference between double-headed ACD and the other cases of extraction from a deleted VP in (83). For the examples in (83), the representations in (91) and (93) show that dependency parallelism is easily satisfied. Now consider the double-headed ACD example (97) (repeated from 31)).

(97) *Polly visited every town in every country Erik did Δ. (Δ = visit *t*)

For a sloppy interpretation, the lexical content of the two relevant traces must be bound variable NP-pro. Representation (98) is one possibility for accomplishing this. In (98), both the QR-NP *town* and the RC-NP *country* have undergone covert movement to the lowest position where they c-command the restrictor position of the landing site of QR or the relative clause head respectively, and also each c-command the trace position.

(98) $\text{town}_F\ \lambda_1$ [[every t_1 in [$\text{country}_F\ \lambda_2$ [every $t_2\ \lambda_x$ Erik did Δ (visit the_x NP-pro_2)]]] λ_y Polly visited the_y NP-pro_1]

Deletion cannot be licensed in (98) for the following reason: a sloppy interpretation can only be licensed if ~ and σ both apply to a constituent that includes the binder of the sloppy variables. For this reason, σ must apply to either the entire clause in (98) or the constituent consisting of λ_1 and its scope. But then the scope of σ necessarily includes the deleted VP, forcing a configuration where the argument of σ contains ~. It is impossible to satisfy the deletion licensing condition (9) in this configuration. For this reason, a sloppy construal of the trace content is blocked in double-headed ACD.

We expect that dependency parallelism will also constrain the sloppy con-

strual of the trace content in cases other than double-headed ACD. This prediction is borne out. Consider the examples in (99):

(99) a.*? I like the house you're building, and I don't know which boat you are.
b.*? I don't know which house you're building, and I like the boat you are.

The examples in (99) point to a difference between *wh*-movement chains and relative clauses: extraction from the antecedent is of one type, while extraction from the deleted VP is of the other type. Two sets of contrasts are relevant in evaluating (99). On the one hand, (99) contrasts with the examples in (100), which have the same structure but where the two extracted phrases have the same core NPs. On the other hand, (99) contrasts with the examples in (101), which have different core NPs, but where the antecedent and the deleted VP are in structurally parallel environments.

(100) a.? I like the house you're building, and I don't know which house Bill is.
b.? I don't know which house you're building, and I like the house Bill is.

(101) a.? I like the house you're building, and I like the boat you are.
b.? I don't know which house you're building, and I don't know which boat you are.

The ill-formedness of (99) is predicted by the dependency parallelism requirement. Consider the representation of (99a) in (102):

(102) I like $\underbrace{\text{house } \lambda_1 \text{ [the } t_1\text{] } \lambda_x \text{ you're building the}_x \text{ NP-pro}_1}_{\sigma_3}$
and I don't know
$\underbrace{\text{boat}_F\ \lambda_2\ [\text{which}_F\ t_2]\ \lambda_x\ Q_{[+wh]}\ \text{you are}\ \Delta\ (\text{building the}_x\ \text{NP-pro}_2)}_{\sim_3}$

Because the core NPs of the two moved phrases are different, deletion licensing in (102) could only be made possible by sloppy binding of the core NPs, which requires the antecedent and focus domain to be at least as large as shown in (102). But, even if both *boat* and *which* are focused, deletion is not licensed because the argument of ~ contains the question complementizer $Q_{[+wh]}$, which, is not present in the argument of σ.

One remaining question at this point is why the ill-formedness of the

examples in (99) is less strong than that of the double-headed ACD examples. I would like to suggest that this is related to another difference between (98) and (102): namely, the argument of σ must contain ~ in (98), but in (102) there is no overlap between the arguments of σ and ~. Therefore, licensing of deletion via an entailment can be at least marginally available in (102), since the first conjunct in that representation suggests that it is also the case that 'I know which house you're building'.

As Rooth notes, his account of dependency parallelism predicts that deletion licensing via an entailment should be able to obviate the parallelism requirement when it applies. This is illustrated by the examples in (103); (103d) involves VP-sloppiness.

(103) a. Yesterday the guy John works for told him to shape up, and today Bill's boss did. (Rooth 1992a)
b. Those who live in New York hate its subway system, and people in Tokyo do, too. (Tomioka 1999, (16b))
c. First England withdrew its vote. Then France's representative did too. (Hardt 1999, (46))
d.? When you heard me whistle, you said I shouldn't, but when I sang you didn't.

5. The Lexical Content of Traces

In this section, I compare two tests for the lexical context of traces: deletion licensing, which I discussed in the previous three sections, and Condition C (Chomsky 1993; Fox 1999b, and others). I do so by investigating more precisely which parts of the moved phrase these tests show to be in the base position. The best result for both accounts would be that both tests support the same generalization as to what material is interpreted in the trace position. If this is the case, we can conclude that deletion licensing and Condition C apply at the same level of grammar, and that at this level of grammar some lexical content of moved phrases is represented in the base position. The evidence in this section indeed corroborates this best possible outcome.

The main empirical generalization I argue for in this section is the following: In A-bar movement of a DP, the core NP must be interpreted in the trace position. I will call this claim *core reconstruction*. I assume, following for example Chomsky (1981), that *wh*-movement and quantifier movement belong to the class of A-bar movements. The term *core NP*, as defined in (7) above, refers to the head noun with its argument, but excludes all adjuncts. My main claim does not say anything about reconstruction

in A-chains, and about material not in the core NP in A-bar chains. I show below that reconstruction is not obligatory in either of these two cases.

The two tests for reconstruction that I use in this section are Condition C reconstruction and deletion licensing in double-headed ACD. The best-known observation about Condition C reconstruction is the so-called argument/adjunct contrast (Freidin 1986; Lebeaux 1988).[24] This contrast is illustrated in (104).

(104) a.* [Which argument that $John_i$ was wrong]$_j$ did he_i accept j in the end?
b. [Which argument that $John_i$ had criticized]$_j$ did he_i accept j in the end?

If we assume that the presence of a Condition C effect indicates that the pronoun *he* c-commands *John* in the LF-representation, then (104) shows the following: The R-expression *John* in (104a), which is part of the core NP of the *wh*-phrase, must reconstruct for Condition C, whereas the R-expression *John* in (104b), which is not part of the core NP, need not reconstruct. Hence we can take (104) to show that reconstruction of the core NP is obligatory in overt *wh*-movement, but reconstruction of material that is not part of the core NP is not.

The copy identity account of double-headed ACD also argues for core reconstruction, as I have shown above. Consider again the contrast in (105) (repeated from (2)).

(105) a.* Polly visited every town that is near the lake Erik did Δ. (Δ = visit *t*)
b. Polly visited every town that is near the one Erik did Δ. (Δ = visit *t*)

I have argued in section 3 that (105) shows that reconstruction of the core NP is obligatory in both the QR-chain and the relative clause internal chain.

The examples in (104) and (105) constitute the first two arguments that the core NP is relevant for both Condition C and deletion licensing. However, the structure of the examples in (104) and (105) is greatly different: (104) involves overt *wh*-movement, whereas (105) involves a relative clause and covert movement for ACD resolution. The following sections corroborate my claim that the core NP is relevant for both Condition C

[24] Most of the examples discussed in the work of Freidin and Lebeaux have been argued to be poorly controlled (Lasnik 1998; Kuno 1997). However, controlled examples like (104) also corroborate the conclusions reached by Freidin and Lebeaux (Heycock 1995; Fox 1999b).

and deletion licensing, drawing on evidence from different types of movement.

5.1. *Argument/Adjunct Contrasts I: The Relative Clause Head*

In this section, I consider reconstruction in QR-chains. The basic double-headed ACD example in (105) shows that for deletion licensing reconstruction of the core NP, but not the adjoined relative clause, is required in these examples. Now, consider Condition C reconstruction in QR-chains. There's a difference between ACD-resolving QR and purely scopal QR with respect to Condition C. Namely, as Fiengo and May (1994, p. 274) and Fox (1995) show, only the movement that resolves ACD obviates Condition C for all speakers. A contrast between ACD-resolving QR and non-ACD resolving QR is shown in (106).

(106) a. Someone introduced him_i to everyone John_I wanted her to Δ. (Δ = introduce *t* him to)

b.*? Someone introduced him_i to everyone John wanted you to meet.

ACD resolution in (106a) requires QR of the *everyone*-DP to a position above VP. Similarly, QR to such a position must take place in (106b) at least when *everyone* takes scope above the subject. Nevertheless there's a difference with respect to Condition C for many speakers.

Both Fiengo and May (1994) and Fox (1995) propose for (106a) that Condition C applies to the representation generated after QR adjoins the *everyone*-DP to the VP and leaves a bare variable in the base position. (They differ though in their account for (106b).) This representation is shown in (107).

(107) [everyone λ_y John wanted her to Δ (introduce him to *y*)] λx someone introduced him to *x*

Representation (107) is not consistent with core reconstruction. But actually the Condition C evidence only shows that the relative clause is not represented in the trace position. A representation like (108) rather than (107) is also, therefore, compatible with the evidence in (106).[25]

(108) [everyone λ_y John wanted her to Δ (introduce him to *y*)] λx someone introduced him to the_x (one)

---

[25] I am not considering representations where the relative clause is split between the base position and the landing site, since I would not know how to interpret such representations.

In (108), the head noun is represented in the trace position, as is predicted by obligatory core reconstruction.

In fact, there is evidence from Condition C that a representation like (108) is correct. This observation was made independently by Merchant (2000) and myself (Sauerland 1998). Consider the contrast in (109).

(109) a.* In the end, I did ask him$_i$ to teach the book of David$_I$'s that Irene wanted me to Δ. (Δ = ask him to teach)

b. In the end, I did ask him$_i$ to teach the book of Irene's that David$_i$ wanted me to Δ. (Δ = ask him to teach)

The different between (109a) and (109b) is whether the R-expression *David* occurs in the ACD relative clause as in (109b) or somewhere else in the QR DP as in (109a). The fact that Condition C is not obviated in (109a) shows that the core NP must be represented in the base position of QR in ACD-examples, while the ACD-relative is not. The representation that (109a) has at the level where Condition C applies is shown in (110).

(110) [the book of David$_i$'s

$\underbrace{\lambda_y \text{ Irene wanted me to } \Delta \text{ (ask him}_i \text{ to teach the}_y \text{ (book of David}_i\text{'s))}}_{\sim_1}$

$\underbrace{\lambda_x \text{ I asked him}_i \text{ to teach the}_x \text{ (book of David}_i\text{'s)}}_{\sigma_1}$

Example (109) shows that there is in fact a complete match between the Condition C evidence and the deletion licensing evidence in the case of ACD-resolving QR. Left to explain is the difference between ACD-resolving QR and purely scopal QR with respect to Condition C in (106). Danny Fox, in a series of papers, makes two different suggestions to explain this difference which are both compatible with my assumptions. In the earlier papers (Fox 1995 and 1999b), he assumes that usually QR requires reconstruction of the entire moved phrase, but that ACD somehow blocks syntactic reconstruction of the relative clause. Specially, he assumes that QR involves copying of the entire DP to the landing site. Representations like (108) and (110) are then derived by deleting parts of the copied material.[26] Fox proposes that deletion of material is uneconomical, and

[26] Deletion alone is actually not enough. The determiner must be deleted in the base position, and replaced with a definite determiner that is coindexed with the binder which the moved phrase is the argument of. This predicts that Condition C should be obviated even with QR that does not resolve ACD if the R-expression is part of the determiner of the DP that

therefore only occurs in the case of ACD, where it is forced. Therefore, the entire moved DP is represented in the base position if QR does not take place for ACD-resolution. In later papers (Fox and Nissenbaum 1999; Fox 2002), Fox develops a different account of quantifier raising, within which the contrast (106) is analyzed in a different way. I refer the reader to Fox (2002).

The match between Condition C and Copy Identity can be strengthened even more in this case with further evidence from deletion licensing. In the examples of double-headed ACD considered so far, the core NP was always a single head noun. The contrast between (111a) and (111c) shows that double-headed ACD is also blocked when the two core NPs differ not in their head noun, but in the arguments of the head noun ((111b) is independently awkward because of the repetition of the complex core NP).[27]

(111) a. Bill gave a description of Mary that's similar to the one John did Δ. (Δ = give)
b.?? Bill gave a description of Mary that's similar to the description of Mary John did Δ. (Δ = give)
c.*? Bill gave a description of Mary that's similar to the description of Sue John did Δ. (Δ = give)

5.2. *Argument/Adjunct Contrasts II: Position of the Deleted VP*

The claim that core reconstruction of the deleted VP is forced has implications for ACD structures. Consider again example (112) of single-head ACD.

(112) Polly visited every town Erik did Δ. (Δ = visit)

Deletion licensing can be satisfied here, because the relative clause that contains the deleted VP is not reconstructed to the base position of QR.

---

undergoes QR. The contrast in (i) shows that this prediction is borne out (see also Sauerland 2001).

(i) a. Someone must've fed him_i $John_i$'s every move over earphones.
b. *Kasparov must've fed him_i $John_i$'s every move over earphones.

[27] The relative acceptability of (i) shows that different adjuncts of the two relevant core NPs do not block double-headed ACD. This might be taken to represent another correlation between copy identity and Condition C. However, (i) also involves items from the same semantic field and is therefore independently expected to be acceptable, as discussed in section 3.3.3.

(i) John visited a town near Madrid that had signs for the one near Rome Bill did ⟨visit⟩.

The account predicts that ACD is blocked whenever the relative clause is part of the core NP. The examples in (113) show one case where the prediction is borne out: ACD is impossible when the deleted VP occurs in an argument of the head noun, rather than a relative clause adjunct.

(113) a.*John found a proof that Bill never does Δ. (Δ = find a proof)
b.*Mary asked the question which question Sue did Δ. (Δ = ask *t*)

Example (113a) does not allow the indicated reading where the *that*-clause is an argument of *proof*. At the same time, (113a) allows an interpretation where the *that*-clause is a relative clause with the paraphrase 'that Bill never finds *t*'. Similarly, (113b) does not allow the indicated interpretation of the deleted VP.[28]

Further evidence of the relevance of the core VP comes from cases where the deleted VP is contained in an adjunct, but this adjunct is part of the core NP. This is shown by the examples of double-headed ACD in (114).

(114) *Jonathan visited every relative of the relative Danny did Δ. (Δ = visit *t*)

In (114), the head *relative* of the relative clause is itself an argument of the higher occurrence of *relative* and therefore part of the core NP of the DP that moves for ACD resolution. (115) makes the same point, showing a contrast between an argument and an adjunct to the higher occurrence of the noun *picture*.[29]

(115) a. *John is planning to paint many pictures of the one Dali is Δ. (Δ = planning to paint)
b. John is planning to paint many pictures showing the one Dali is Δ. (Δ = planning to paint)

Finally, the contrast in (116) makes the same point.

(116) a.*John is filing a report that Mary read the report Bill is Δ. (Δ = filing)
b. John is filing a report according to which Mary read the report Bill is Δ. (Δ = filing)

---

[28] I do not know of a discussion of examples like (113) in the literature. These examples also provide an (additional) argument against the account of ACD as deletion of just a verb, which was suggested by Cormack (1984). Deletion of the verb should not be sensitive to the argument/adjunct difference.

[29] For context to this judgment, imagine that John's profession is to paint pictures of Dali's pictures. One day, John meets Dali and Dali tells John about his plan for a new painting. John likes the plan a lot, and immediately makes his own plans based on Dali's plan.

The claim that adjuncts inside the core NP must reconstruct is not widely accepted for Condition C reconstruction. Mostly, the opposite is implicitly assumed. Consider for example the work of Takano (1995) and Heycock (1995) on reconstruction of predicatively used DPs. One of their arguments for obligatory reconstruction of predicates is based on examples like (117), where a Condition C effect is found.

(117) *The type of girl that hates John$_i$, he$_i$ said Mary is. (Takano 1995, (12d))

Takano and Heycock claim that the fronted DP in (117) must totally reconstruct to explain the presence of the Condition C effect since the R-expression *John* is contained in a relative clause. However, the relative clause containing *John* is not an adjunct to the core NP of the fronted DP, but rather adjoined somewhere inside of this core NP. Therefore, I would claim that (117) only provides evidence for obligatory reconstruction of the core NP of the fronted DP. To my knowledge, reconstruction of low adjuncts has only been explicitly discussed by Tada (1993), who argues that the claim is true for Japanese but has doubts about English.[30] The contrast in (118) indicates that Tada's claim also holds for English.

(118) a.* Which book of the woman Bill$_i$ admires did he$_i$ give to his$_i$ parents?
b. Which book about the woman Bill$_i$ admires did he$_i$ give to his$_i$ parents?

5.3. *Relative Clause Internal Traces*

The relation of relative clause internal movement to the relative clause head is a subject of great debate. Some relevant evidence from binding theory is discussed by Munn (1994), Safir (1999), and myself (Sauerland 2003).

---

[30] Tada (1993, fn. 25) doubts the existence of a difference between modifiers to the core NP and modifiers internal to the core NP for English because of (ia) (attributed to Noam Chomsky, p.c.). In (ia), Condition C is obviated even though the relative clause containing the R-expression *John* is adjoined to the lower NP *book*. This argument does not convince me, though, because the *for*-PP itself is a modifier to *book*. The obviation of Condition C in (iia) shows that the *for*-PP is an adjunct, as does the separability of the *for*-PP from the NP in the copular paraphrase in (iib) (see Schütze 1995 for argument/adjunct tests for DP-internal PPs).

(i) a. The award for the book that John$_i$ wrote, he$_i$ never received.
b. The award for the book that John$_i$ received, he$_i$ never cashed.

(ii) a. Which award for *Titanic*$_i$ did everybody agree it$_i$ deserved?
b. The award was for the book.

Specifically, I argue that Condition C reconstruction facts support the matching analysis of relative clauses that I summarized in section 1.2 above.[31]

At first, relative clauses seem to show a mismatch between Condition C and deletion licensing. Munn (1994) claims that Condition C provides no evidence for reconstruction into relative clauses, in contrast to *wh*-movement in questions. This is confirmed by the contrasts in (119) and (120).

(119) a. Which is the picture of John$_i$ that he$_i$ likes?
b.*Which picture of John$_i$ does he$_i$ like?

(120) a. I have a report on Bob's division he won't like. (Merchant 2000, fn. 5)
b.*Which report on Bob$_i$'s division will he$_i$ not like.

The absence of Condition C reconstruction in relative clauses is puzzling, given that deletion licensing indicates reconstruction of the relative clause head into an internal position.

However, there's also evidence from Condition C for reconstruction into relative clauses. Namely, Safir (1999) observes that in example (121) a strong crossover effect is observed.

(121) *Pictures of anyone$_i$ which he$_i$ displays prominently are likely to be attractive ones.

I assume with Chomsky (1981) and Reinhart (1997b) that strong crossover is an instance of Condition C (or the principles underlying Condition C) triggered by the trace of A-bar movement. I assume that *anyone* in (121) must QR. However, the trace left by *anyone* in (121) is not c-commanded by the pronoun *he* in the relative clause external position. It would only be c-commanded in the relative clause external trace position.

The apparent paradox is resolved by the matching analysis of section 5.3. Recall that, while in normal movement relationships material in a chain must be strictly identical, I assume that the material in the relative clause internal position need only be identical in meaning to other deleted material.

Consider now again the contrast between *wh*-movement and relative clauses. With *wh*-movement the moved material that is represented in the trace position must be exactly identical to the moved material. Since the

---

[31] I leave aside here cases with bound variable pronouns like (i) (see Sauerland 2003; Hulsey and Sauerland 2003).

(i) The picture of himself$_i$ every boy$_i$ brought annoyed his$_i$ mother.

NP is required to be represented in the base position, it must be represented as *picture of John* and triggers Condition C. In the relative clause internal trace position the NP must be represented as well; however, it's sufficient that a pronoun with the same interpretation occupy that position. I assume that the NP-pronoun *one* can refer to any one-place predicate, and therefore the lexical material inside the relative clause can consist of just *one*.

In Safir's example (121), however, the pronoun *one* is excluded from the relative clause internal trace position for the following reason. In (121), *anyone* takes scope outside of the relative clause head. After *anyone* undergoes QR, the relative clause head is *picture of x*, with *x* bound by *anyone*. Since there's no one-place predicate, the pronoun *one* cannot mean the same as the relative clause head in this example. Therefore, the material in the relative clause internal position must be *picture of x* or *one of x*, both of which are predicted to cause a Condition C/strong crossover violation as shown in (122).[32]

(122) *anyone λ_x [pictures of the$_x$ (person)
external head
which picture of the$_x$ (person)] λ_y he$_x$ displays prominently
internal head
the$_y$ (pictures of the$_x$ (person))] are likely to be attractive ones.
internal trace

For deletion licensing, it is irrelevant whether the lexical material in the relative clause is the RC-head or some material that is identical to the RC-head for the purposes of deletion licensing. Therefore, the view of relative clauses of Sauerland (2003) synthesizes the evidence from Condition C, strong crossover, and deletion licensing (see also Bhatt 2002 and Cresti 2000).

5.4. *Overt* Wh-*Movement*

The standard evidence showing the argument/adjunct distinction with Condition C concerns reconstruction of overt *wh*-movement. This section shows that there is corresponding evidence from deletion licensing showing

[32] I assume that movement of *anyone* leaves the core NP *person* in the trace position.

that reconstruction of the core NP is obligatory with overt *wh*-movement. Jacobson (1996) already points out that the ungrammaticality of example (2a) is also found in cases like (123a), where ACD is resolved by overt *wh*-movement rather than by covert movement. The improvement with (123b) and (123c) shows that it is the difference in lexical content that causes the ungrammaticality of Jacobson's example.

(123) a.* Do you know which town near a lake Mary did Δ John visited? (Δ = visit)
b. Do you know which town near a town Mary did Δ John visited? (Δ = visit)
c. Do you know which town near a one Mary did Δ John visited? (Δ = visit)

The account of (123) is straightforward: Overt *wh*-movement requires reconstruction of the core NP, and therefore the deletion licensing requirement enforces in particular identity of the core NPs of the *wh*-phrase and the relative clause head.

The examples in (124) show that judgments do not change if the other VP is deleted – the one that contains the trace of *wh*-movement (again, Jacobson (1998) already observes examples like (124a)).

(124) a.* Do you know which town near a lake Mary visited Δ John did? (Δ = visit)
b. Do you know which town near a town Mary visited Δ John did? (Δ = visit)
c. Do you know which town near a one Mary visited Δ John did? (Δ = visit)

5.5. *A/A-bar Contrasts*

In contrast to A-bar movement, A-movement does not require Condition C reconstruction if the moved DP takes scope in the landing side.[33] This is illustrated by the contrast in (125).

(125) a. [One relative of Kai_j's]$_i$ seemed to him_j to i like Kazuko
b.* [Which relative of Kai_j's]$_i$ did he_j say i likes Kazuko

Again, the same contrast is observed using deletion licensing as a test. The example in (126) shows a contrast between topicalization (A-bar-movement) and passivization (A-movement).

[33] Fox (1999b) and Romero (1997) show that the entire A-moved DP reconstructs when it takes narrow scope.

(126) a.? The town near the lake that was Δ seems to have been visited by vandals as well. (Δ = visited by vandals)
b.*The town near the lake they did Δ the vandals seem to have visited as well. (Δ = visit)

The same contrasts obtain if the other VP, the one in the matrix clause, is deleted. This is shown by (127).

(127) a. The town near the lake that was visited by vandals seems to have been Δ as well. (Δ = visited by vandals)
b.*The town near the lake they visited the vandals seem to have Δ as well. (Δ = visited)

The contrast between A- and A-bar movement in (126) and (127) matches the Condition C data. It is predicted by my account on the assumption that the traces of A-movement are bare variables, whereas core reconstruction applies to A-bar movement. However, Kennedy (1994) and Heim (1997a) report examples where deletion licensing seems to indicate obligatory reconstruction of A-movement. These are shown in (128).

(128) a.??Every man who said George would buy some salmon did Δ. (Δ = buy some salmon) (Kennedy 1994, (2b))
b.*A proof that God exists does Δ. (Δ = exist) (Wasow 1972, p. 93)

Kennedy and Heim assume that antecedent containment in (128) is resolved by A-movement of the subject from the VP-internal subject position to the overt subject position. But if this was the case, the Condition C evidence for reconstruction and the VP-deletion would point into different directions. Moreover, the fact that (126a) and (127a) are quite acceptable would need to be reconciled with any account of (128). Finally, Kennedy (1994, fn. 3) and Heim (1997a) point out that examples like (128a) improve for many speakers when a focus particle like *too, as well*, or *instead* is added. This improvement is not found in the A-bar movement examples, as the contrasts in (126) and (127) show.

Contrary to Kennedy and Heim, I argue that A-bar movement is required to resolve antecedent containment in the examples in (128). For (128b), which is the most clearly ungrammatical of all the A-movement examples considered, this is entailed by Diesing's (1992) analysis of existential indefinites. (128b) involves the subject of the predicate *exist*. Since the subject is naturally interpreted existentially, it must be interpreted in a VP-internal position, as Diesing (1992) and Kratzer (1989) argue. However, if the subject occupies a VP-internal position, A-movement has not actually resolved antecedent containment. To do so, QR of the existential subject from the VP-internal position would be necessary. If, as Diesing argues, QR of

existential indefinites is impossible, we need to consider only representation (129) to see whether deletion is licensed in (128b).

(129) [a proof that $\underbrace{\text{God exists}}_{\sigma_1}]_F$ does Δ (exist)

$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\sim_1}$

The argument of $\sim_1$ in (129) contains σ_1, however. As I pointed out above, it follows from condition (10) that this configuration has contradictory presuppositions. Recall that I proposed (10) to avoid Russell's paradox, which can arise if assignment functions can be, for some index, defined for themselves. (129), however, presupposes an assignment function that is defined for itself: the intention is that $g(1)$ be related to the interpretation of the argument of $\sim_1$ and σ_1. But since the argument of $\sim_1$ contains σ_1, the argument of $\sim_1$ must be defined for g, and therefore $g(1)$ must be defined for g, too. Hence, (129) must be ruled out to avoid Russell's paradox. In the definition of $\sim$ in (21), this was done by requiring that the argument of $\sim_1$ be defined for $g\backslash 1$. This presupposition of $\sim_1$ is violated in (129).

Even if QR of the subject were possible, the antecedent VP in (128b) would be contained in an argument of the subject, and therefore a copy of it would remain in the trace position as I argued in section 5.2 above. Thus, QR would derive the representation in (130). Since the scope of $\sim_1$ still includes one occurrence of σ_1, QR would not be able to rescue (128b).

(130) a $\overbrace{\text{proof that God exists}}^{\sigma_1}$

λ_x the$_x$ proof that $\underbrace{\text{God exists}}_{\sigma_1}$ does Δ (exist)

$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\sim_1}$

Now consider (130a). I suggest that the explanation of (130a) is related to factors determining the choice of the domain of deletion licensing. Rooth's account of deletion licensing is not constrained in this regard, but it may be too permissive (cf. Marti 2003). Tentatively, I suggest that the domain of deletion licensing must include a focused phrase in the phonological representation. For (130a), that requires that the surface position of the subject must be part of the deletion licensing domain (i.e., in the scope

of ~). Therefore, the subject in (130a) must QR from its surface position for resolution of ACD, and the copy identity account applies to (130a) just as to the examples of double-headed ACD.

(131) every man$_F$ who said George would $\overbrace{\text{buy some salmon}}^{\sigma_1}$
$\underbrace{\lambda_x \text{ the}_x \text{ man } \lambda_y\ y \text{ did } \Delta \text{ buy some salmon}}_{\sim_1}$

The ellipsis licensing requirement is not satisfied in (131): the argument of σ_1 is a function of type $\langle e, t\rangle$ that is defined for all individuals. By contrast, the argument of $\sim_1$ is a function of the same type that is defined only for individuals that are a man.

The idea that the domain of ~ must include a focus readily explains why example (128a) improves when a focus particle like *as well* is added to the VP. Because the focus particle itself is focused, ~ can take the VP as its argument, as in (132).

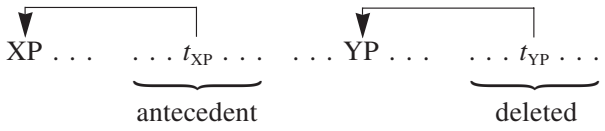
(132) every man$_F$ who said George would $\overbrace{\text{buy some salmon}}^{\sigma_1}$
$\lambda_x \text{ the}_x \text{ man } \underbrace{\lambda_y\ y \text{ did } \Delta \text{ buy some salmon [as well]}_F}_{\sim_1}$

The deletion licensing condition is satisfied in (132) because both the argument of σ_1 and that of $\sim_1$ are defined for all individuals.

6. Conclusion

The main result of this paper is the novel support it provides for obligatory reconstruction. Obligatory reconstruction is the idea of Chomsky (1993) that the base position in A-bar movement chains must contain lexical material of the moved phrase, which is interpreted in this position. Chomsky proposes that movement is really a copying operation followed by phonological deletion. Previous work arguing for obligatory reconstruction (Chomsky 1993; Fox 1999b) has relied on evidence from Condition C, and shown that complex data in this domain follow from a set of assumptions about the representation of lexical material in the base position of movement.

My paper has provided an entirely novel argument for obligatory reconstruction. The argument comes from cases of deletion where both the deleted phrase and its antecedent contain a trace, as sketched in (133).

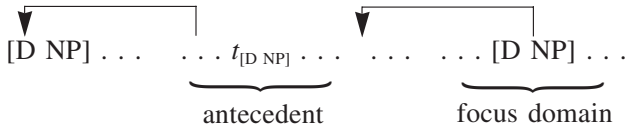
(133) 

Deletion is licensed in (133) if the deleted phrase has the same interpretation as its antecedent. If there is obligatory reconstruction of parts of the moved phrases into their base positions, it is predicted that the reconstructed material must be identical for the antecedent phrase and the deleted phrase. I show that this prediction is borne out in some instances of schema (133), namely the double-headed ACD cases discussed in section 2. And I furthermore show, in section 4, that in most other instances of structures like (133), independent factors obviate the effect, and that once this is controlled for the effect reemerges.

My new evidence, I furthermore argue, is closely related to the old evidence for obligatory reconstruction from Condition C. In section 5, I argue that the data from deletion and the data from Condition C establish exactly the same generalization as to what material must be reconstructed into the base position of movement. Namely, both tests argue that what I have called the core NP – the NP-complement of the D-head of a moving DP excluding all adjuncts to this NP – must be represented in the base position of movement.

Finally, the data in section 3.3.3 strongly support the claim that obligatorily reconstructed material is indeed interpreted in the base position of movement. The data show that the extent to which deletion is degraded in double-headed ACD is predictable from the semantic relationship of the syntactically reconstructed material. I believe to have established, therefore, that reconstructed material is interpreted in the semantic representation of A-bar movement, something that the Condition C evidence alone did not show clearly.

In another recent paper of mine (Sauerland 2001), I present a further argument for core reconstruction. The idea of that argument is to look at the configuration sketched in (134), in particular at the licensing of contrastive focus in the focus domain.

(134) 

If core reconstruction takes place in the antecedent, the trace should contain the same lexical NP that occurs in the focus domain, and therefore a narrow focus on D should be licensed in the focus domain. Without core reconstruction, on the other hand, only a wide focus on the entire DP should be licensed in the focus domain. As I show in my paper, the prediction of the core reconstruction approach is borne out. One interesting aspect of this line of reasoning is that it carries over to German, providing evidence for core reconstruction from a language other than English.

These three arguments together provide strong support for core reconstruction. This is, as far as I can see, an important advance for the study of syntax and semantics for several reasons. One important consequence is that A-bar traces must be present in linguistic representations. The existence of traces has frequently been cast into question as they were assumed not only to be inaudible, but also to be interpreted as bare variables without any lexical content. I have shown, however, that traces have lexical content that contributes to their interpretation in a way similar to definite DPs. In this way, my analysis establishes the existence of traces on purely semantic grounds.

Secondly, my results corroborate the copy theory of movement (Chomsky 1933). Recall that the copy theory of movement claims that the the syntactic operation 'movement' actually involves copying of a phrase with subsequent adjustments. The copy theory easily predicts that the lexical content of the traces is a part of that of the moved phrase. In this way, my analysis also corroborates the claim of *Syntactic Reconstruction* that reconstruction is represented in the syntax by actually having lexical material in the trace position. Syntactic reconstruction is frequently cast into doubt by work on pseudoclefts and also clitic left dislocation (Cecchetto 2001; Sharvit 1999; Sternefeld 2001). In this paper, I have not addressed these constructions since I believe them to be insufficiently well understood. However, I believe to have made the case for syntactic reconstruction for three core classes of A-bar movement (*wh*-movement in questions, *wh*-movement in relative clauses, and quantifier raising) considerably stronger. One might argue that two mechanisms are required, syntactic reconstruction in some cases and an enrichment of semantic mechanisms for the other cases. However, such a system would contain a severe degree of redundancy. My hope is, therefore, that this study will inspire new work extending

the reach of the syntactic reconstruction account to other reconstruction phenomena.

The third result of my analysis is that movement is a far more complex operation semantically than was previously assumed. One especially interesting aspect is that A and A-bar chains differ semantically at least with respect to the content of traces. I hope that this observation will make it possible to explain some distinctions between the two types of chains on semantic grounds. While I have not attempted to do so here, except for the discussion in section 5.5, in the appendix below I suggest some changes in the semantics of traces that then predict crossover phenomena for A-bar chains.

As always, my account also raises a number of issues that did not come up previously. For example, there is the question how constructions that have been thought to involve empty operators are analyzed. It seems more consistent with core reconstruction that all operators must have some lexical content, but it remains to be seen whether that can be maintained. The most important remaining question, though, is why core reconstruction is obligatory. This needs to be stipulated not only for what I have said above, but also in the work of Chomsky (1993) and Fox (1999b, 2002). In the appendix below, I offer an idea on how to derive the obligatoriness of core reconstruction from the semantics of quantifiers.

7. Appendix: A More General Semantics of Traces

In the main part of this paper, I adopt a semantics of traces based on the indexed definite determiner in (135) (repeated from (8a)).

(135) $[\![\text{the}_n]\!]^g(P)$ is defined if $P(g(n)) = 1$.
If defined, $[\![\text{the}_n]\!]^g(P) = (g(n)$.

Since the indexed definite description is a variable with an added presupposition, this proposal is probably the easiest way to incorporate the semantic contribution of the lexical material in the trace position. Neither the type of the variable is changed, nor is it necessary to change the mechanisms of variable binding and quantification. Furthermore, my paper has shown that core reconstruction combined with (135) explains a set of otherwise puzzling facts concerning the interaction of movement and deletion.

However, it is probably true that the facts in this paper follow from core reconstruction combined with any semantics of traces as long as a trace's lexical content makes a semantic contribution. In this appendix, consider an array of further facts that seem to be relevant for the seman-

tics of traces, other than the interaction of movement and deletion: quantifiers in questions, crossover phenomena, and the explanation of core reconstruction. Based on these facts, I shall conclude that the choice function alternative to (135), which I first suggested in my dissertation (Sauerland 1998), should be considered.

First consider quantifiers in questions. The semantics of traces based on (135) does not straightforwardly extend to questions with quantifiers like (136).

(136) Which relative of his did every student invite?

An LF-representation of (136) is shown in (137). In (137), core reconstruction is applied to the *wh*-phrase. Furthermore, the core NP of the *wh*-phrase is deleted in sentence-initial position since the core NP contains a variable bound only in the trace position of the *wh*-movement chain. Consider the interpretation predicted for the LF-representation (137), which is based on a simplified version of Karttunen's (1977) semantics of questions (Sauerland 1998), where a question corresponds to the set of possible answers.[34]

(137) $\lambda p\, \exists x \in D_e\colon p \rightarrow [\lambda_w\colon$ [every student] $\lambda_y\colon y$ invited(w) the$_x$ relative of y]

Since the variable x is unbound in the constituent consisting of the universal quantifier *every student* and its scope, the presupposition of the trace is projected universally, leading to a presupposition that x be a relative of every student. Therefore, (136) is predicted to presuppose that there be some person who is the relative of every person since otherwise the question would denote the empty set. This is wrong, though, because (136) can be used in other situations as well.

One way to overcome the problem with (137) is to assign *every student* wider scope, such that its scope includes the binder of the variable x. Assuming QR as the scoping mechanism, (138) is such a representation (for ease of presentation, I do not represent core reconstruction when it is not relevant, as with QR in (138)).

(138) λp [every student] $\lambda_y\colon \exists x \in D_e\colon p \rightarrow [\lambda_w\colon y$ invited(w) the$_x$ relative of y]

---

[34] In (137) and subsequently I omit world argument positions of nouns for ease of presentation. As far as I can see, including them would bring up the same problems I discuss by means of the individual variable in (136). Also, the word *which* is represented as an existential quantifier in (137), which I assume is its interpretation.

Representation (138) carries a weaker presupposition than (137); namely, that every student have at least one relative. Hence, (138) would be at least a good start towards a semantics of (136). It is widely assumed, however, that an account of (136) based on QR out of questions is syntactically implausible (Engdahl 1986; Chierchia 1993; Pafel 1999 and others). One problem, which I believe has not been discussed before, is posed by examples like (139), where the quantifier *every student* cannot take scope over *one of the letters*. It is, I believe, difficult to reconcile this scope restriction with the assumption that *every student* must QR to a position outside of the question.

(139) Which relative of his did one of the letters say that every student invited?

Therefore, I follow the majority of the literature in assuming that the problem raised by (136) is not in general solved by scoping quantifiers out of questions, and that some other account of (136) is still needed. Two other approaches to questions with quantifiers have been suggested: the Skolem function approach (Engdahl 1986; Chierchia 1993) and the choice function approach (Engdahl 1980).[35] I consider versions of these two adjusted to my general assumptions and, in particular, core reconstruction. (140) shows a representation for (136) within the Skolem function approach.

(140) $\lambda p\ \exists f \in D_{ee}$: $p \rightarrow$ [λ_w: [every student] λ_y: y invited(w) the$_{f(y)}$ relative of y]

The variable f in (140) has type $\langle e, e\rangle$. Therefore, *which* must be interpreted as existential quantification over functions of this type, which is already assumed in (140). In the trace position, f takes the individual variable y as its argument to yield an individual. Since the subscript of *the* in (140) is more complex than a bare variable, the lexical entry of the indexed definite determiner needs to be generalized as in (141). The trace in (140) then contributes a presupposition that $f(y)$ be a relative of y's to the sentence meaning and denotes $f(y)$. Consequently, (140) presupposes that there be a function f which assigns to any student a person that is related to that student.

(141) $[\![\text{the}_X]\!]^g(P)$ is defined if $P([\![X]\!]^g) = 1$.
If defined, $[\![\text{the}_X]\!]^g(P) = [\![X]\!]^g$.

---

[35] Choice functions have also been employed in the study of indefinites (Reinhart 1997a and others). As far as I can see, this use is unrelated to my present concerns.

The choice function approach, on the other hand, would on one version assume a representation like (142) for (136), where *which* has been interpreted as existential quantification over choice functions of type $\langle\langle e, t\rangle, e\rangle$.[36] Choice functions are functions that assign to a set (or its characteristic function) an element of that set.

(142) $\lambda p\, \exists f \in CF$: $p \rightarrow$ [λ_w: [every student] λ_y: y invited(w) f(relative of y)], where $CF = \{f \in D_{(et)e} | \forall P$: $P(f(P)) = 1\}$

The question denotation (142) will be the empty set unless at least one choice function f exists satisfying the scope of the interrogative. Therefore, (142) presupposes that for every student the set of people related to him be non-empty. In this way, the choice function approach predicts the right presupposition for the question.

The Skolem and the choice function approach have two interesting properties in common. Both require that the upper copy of the restrictor of *which* must be deleted when it contains a variable that is not bound in that position. In the main text, I have been assuming that both the upper and the lower copy of the restrictor are interpreted. This assumption made it possible to keep the normal interpretation of quantifiers throughout. It also played a role when applying matching in relative clauses in section 2.3 and in the details of the binding of NP-pro in section 4.2. I believe, nevertheless, that this tension is probably harmless. A second common property of both approaches is that they raise the question whether quantification over functions is also possible for non-interrogative quantifiers.[37] At first it may seem obvious that this will not work. Consider for example the representations (143b) and (143c) for (143a) (I use $\exists_2$ to represent the quantifier *there are two different*).

(143) a. Every girl invited two relatives of hers.
b. $\exists_2 f \in D_{ee}$: [every girl] $\lambda_x\, x$ invited the$_{f(x)}$ relative of x
c. $\exists_2 f \in CF$: [every girl] $\lambda_x\, x$ invited f(relative of x)

---

[36] Another possibility that probably should be explored is that the trace is represented as 'the$_{f(\text{relative of } y)}$ (relative of y)' with the lexical entry for *the* in (141). Then the restriction to choice functions need not be stipulated, since it is built into this representation. However, I do not know what kind of syntactic mechanism would come up with such a representation of the trace.

[37] Interestingly this question has, as far as I know, never been considered by the proponents of the functional analysis of questions. Similarly, Sternefeld (2001) and other proponents of the semantic reconstruction for binding – a generalization of the Skolem function approach – have not discussed semantic reconstruction of the restrictor of non-interrogative quantifiers. Nevertheless it seems to be a natural concern to me.

Sentence (143a) is false in a situation where one girl, 1, invited her relatives A and B, while the other girl, 2, invited only one relative of hers, C. But both (143b) and (143c) are true in this situation: $\{1 \mapsto A, 2 \mapsto C\}$ and $\{1 \mapsto B, 2 \mapsto C\}$ are two different Skolem function satisfying (143b). Similarly, $\{\{A, B\} \mapsto A, \{C\} \mapsto C\}$ and $\{\{A, B\} \mapsto B, \{C\} \mapsto C\}$ are two different choice functions satisfying (143c). Hence, it seems that both approaches need to adopt some stipulation to block or suitably constrain functional quantification in cases like (143). Actually, though, the issue of non-interrogative functional quantification is more difficult. We also have to consider the possibility that there are far fewer quantifiers than standardly assumed, for example the quantifier $\exists_2$ might not occur in linguistic representations (cf. Hackl 2000). Furthermore, problems like that of (143) could be resolved by further restricting the set of functions quantified over, e.g. by adopting a restriction that functions must be pointwise different (Sauerland 1998, 2000a). In any way, it seems that whatever resolution is adopted for these issues will be compatible with both approaches.

The two approaches also differ in a number of ways. With regards to syntax, Skolem functions are a conservative extension of the approach based on the indexed definite determiner (8a). The syntax of the choice function approach, however, diverges from that of the indexed definite determiner approach. However, it is possible to extend the choice function approach to the interpretation of non-interrogative traces (Sauerland 1998); this I will call the *generalized choice function approach*. I illustrate this approach by means of example (144) (repeated from (33)).

(144) Polly visited every town Erik did Δ. (Δ = visit *t*)

The LF-representation of (144) on the general choice function approach is shown in (145). Note that both the trace in the relative clause and the trace in the main clause are interpreted as choice functions applying to the core NP. Furthermore, I assume here that the core NP is not interpreted in the upper position of the chain, which will require some revisions of other parts of my account as I noted above.

(145) every $\underbrace{[\lambda_f \text{ Erik } \Delta \text{ (visited } f\text{(town))}]}_{\sim_1}$ $\underbrace{[\lambda_g \text{ Polly visited } g\text{(town)}]}_{\sigma_1}$

Assuming that *every* is interpreted as universal quantification over choice functions (von Stechow 2000), the correct truth conditions are predicted and ellipsis is also predicted to be licensed in (145).

Now, I would like to point out three advantages of the general choice

function approach over the Skolem function approach. The first advantage is that choice functions require only one lexical entry for *which* where it is interpreted as existential quantification over choice functions, while the Skolem function approach requires two lexical entries for *which*, as we saw above.

The second advantage of the choice function approach is that it predicts crossover, as Ruys (2000) and I (Sauerland 1998) independently discovered. Recall that crossover blocks binding of a pronoun from an A-bar position, as in (146) (Wasow 1972 and many others).

(146) a.*? Which girl$_i$ did her$_i$ father pick up late?
b.*? A friend of his$_i$ was talking to every boy$_i$.

Usually crossover must be blocked by a syntactic condition, since semantically an operator can bind any variable of the same type in its scope. But on the general choice function approach, operators in A-bar positions have a different type from pronouns. Consider the representation in (147) for (146a).

(147) *$\exists f \in CF$: her$_f$ father pick up late f(girl)

The illicit crossover configuration is one where the pronoun *her* is of the type of choice functions. But, since the pronoun occupies a position that must be occupied by a variable of the type of individuals, (147) is uninterpretable. In this way, crossover is ruled out by semantic considerations on the choice function approach. The Skolem function approach, however, only predicts crossover in those cases where the restrictor contains a bound variable, since otherwise the upper chain link can bind an individual variable. On the general choice function approach, binding in non-crossover configurations is still permitted by whatever mechanism makes possible the binding of unmoved arguments; for example, covert A-movement. Applying this mechanism to the trace of A-bar movement yields a valid representation for binding, illustrated for (148a) in (148b).

(148) a. Which girl smiled at her father?
b. $\exists f \in CF$ f(girl) λ_x x smiled at her$_x$'s father.

The third advantage of the general choice function approach is that it predicts that core reconstruction is obligatory. For example, consider what would happen in (148a) if instead of applying core reconstruction as in (148b), we don't. As (149) shows, the resulting representation contains an existential quantifier (the interpretation of *which*) that is restricted to the empty set: the intersection of the set of choice functions and a set of individuals.

(149) $*\exists f \in CF \cap \{x \in D_e|\text{girl}(x)\}\ f\ \lambda_x\ x$ smiled at her$_x$'s father.

This type-mismatch rules out representation (149), and therefore core reconstruction is required.[38] The Skolem function approach again does not make this prediction since it allows quantifiers to take predicates of individuals as their arguments. Therefore, the general choice function approach has three advantages over the Skolem function approach.

REFERENCES

Bhatt, R.: 2002, 'The Raising Analysis of Relative Clauses: Evidence from Adjectival Modification', *Natural Language Semantics* **10**, 43–90.

Carlson, G. N.: 1977, *Reference to Kinds in English*, PhD dissertation, University of Massachusetts, Amherst.

Cecchetto, C.: 2001, 'Syntactic or Semantic Reconstruction? Evidence from Pseudoclefts and Clitic Left Dislocation', in C. Cecchetto, G. Chierchia, and M.-T. Guasti (eds.), *Semantic Interfaces*, pp. 90–144. CSLI, Stanford.

Chierchia, G.: 1993, 'Questions with Quantifiers', *Natural Language Semantics* **1**(2), 181–234.

Chomsky, N.: 1981, *Lectures on Government and Binding: The Pisa Lectures*. Mouton de Gruyter, Berlin.

Chomsky, N.: 1993, 'A Minimalist Program for Linguistic Theory', in K. Hale and J. Keyser (eds.), *The View from Building 20, Essays in Linguistics in Honor of Sylvain Bromberger*, pp. 1–52. MIT Press, Cambridge, Mass.

Cormack, A.: 1984, 'VP Anaphora: Variables and Scope', in F. Landman and F. Veltman (eds.), *Varieties of Formal Semantics*, pp. 81–102. Reidel, Dordrecht.

Cresti, D.: 2000, 'Ellipsis and Reconstruction in Relative Clauses', in M. Hirotani (ed.), *Proceedings of NELS 30*. GLSA, University of Massachusetts, Amherst.

Diesing, M.: 1992, *Indefinites*. MIT Press, Cambridge, Mass.

Engdahl, E.: 1980, *The Syntax and Semantics of Questions in Swedish*, PhD dissertation, University of Massachusetts, Amherst.

Engdahl, E.: 1986, *Constituent Questions*. Reidel, Dordrecht.

Evans, F.: 1988, 'Binding into Anaphoric Verb Phrases', in J. Powers and K. de Jong (eds.), *Proceedings of ESCOL 5*, pp. 122–129. The Ohio State University, Columbus.

Fiengo, R. and R. May: 1994, *Indices and Identity*. MIT Press, Cambridge, Mass.

Fox, D.: 1995, 'Condition C Effects in ACD', in R. Pensalfini and H. Ura (eds.), *Papers on Minimalist Syntax, MITWPL 27*, pp. 105–119. MIT Working Papers in Linguistics, Cambridge, Mass.

Fox, D.: 1999a, 'Focus, Parallelism, and Accommodation', in T. Matthews and D. Strolovitch (eds.), *Proceeding of SALT 9*, pp. 70–90. CLC Publications, Cornell University, Ithaca, N.Y.

---

[38] In cases of successive cyclic movement, the general choice function approach requires core reconstruction only in the highest chain link (actually, this will depend on the treatment of intermediate traces). The contrast in (i) with Condition C would then follow, but more work is needed to see whether this prediction is desirable in general ((i) is from Huang 1993, (27)):

(i) a.*? How many pictures of John$_i$ does he$_i$ think that I like *t*?
b. ? How many pictures of John$_i$ do you think that he$_i$ will like *t*?

Fox, D.: 1999b, 'Reconstruction, Variable Binding and the Interpretation of Chains', *Linguistic Inquiry* **30**, 157–196.
Fox, D.: 2002, 'Antecedent Contained Deletion and the Copy Theory of Movement', *Linguistic Inquiry* **33**, 63–96.
Fox, D. and J. Nissenbaum: 1999, 'Extraposition and Scope: A Case for Overt QR', in S. Bird et al. (eds.), *Proceedings of WCCFL 18*, pp. 132–144. Cascadilla Press, Somerville, Mass.
Freidin, R.: 1986, 'Fundamental Issues in the Theory of Binding', in B. Lust (ed.), *Studies in the Acquisition of Anaphora*, Volume 1, pp. 151–188. Reidel, Dordrecht.
Hackl, M.: 2000, *Comparative Quantifiers*, PhD dissertation, MIT.
Hardt, D.: 1999, 'Dynamic Interpretation of Verb Phrase Ellipsis', *Linguistics and Philosophy* **22**, 185–219.
Heim, I.: 1997a, 'Predicates or Formulas? Evidence from Ellipsis', in A. Lawson and E. Cho (eds.), *Proceedings of SALT 7*, pp. 197–221. CLC Publications, Cornell University, Ithaca, N.Y.
Heim, I.: 1997b, 'Predicates or Formulas? Evidence from Ellipsis'. Presentation at LF reading group, MIT.
Heim, I.: 2000, 'Degree Operators and Scope', in B. Jackson and T. Matthews (eds.), *Proceedings of SALT 10*, pp. 40–64. CLC Publications, Cornell University, Ithaca, N.Y.
Heycock, C.: 1995, 'Asymmetries in Reconstruction', *Linguistic Inquiry* **26**, 547–570.
Huang, C. J.: 1993, 'Reconstruction and the Structure of VP: Some Theoretical Consequences', *Linguistic Inquiry* **24**, 103–138.
Hulsey, S. and U. Sauerland: 2003, 'Sorting Out Relative Clauses: A Reply to Bhatt', unpublished ms., MIT and UMass/Tübingen.
Jacobson, P.: 1992, 'Antecedent Contained Deletion in a Variable Free Semantics', in C. Barker and D. Dowty (eds.), *Proceedings of SALT 2*, pp. 193–213. OSU Working Papers in Linguistics, The Ohio State University, Columbus.
Jacobson, P.: 1998, 'Antecedent Contained Deletion and Pied-Piping: Evidence for a Variable-Free Semantics', in D. Strolovitch and A. Lawson (eds.), *Proceedings of SALT 8*, pp. 74–91. CLC Publications, Cornell University, Ithaca, N.Y.
Karttunen, L.: 1977, 'The Syntax and Semantics of Questions', *Linguistics and Philosophy* **1**, 1–44.
Kennedy, C.: 1994, 'Argument Contained Ellipsis', Linguistics Research Center Report LRC-94-03, University of California, Santa Cruz.
Kennedy, C.: 2002, 'Comparative Deletion and Optimality in Syntax', *Natural Language and Linguist Theory* **20**, 553–621.
Kratzer, A.: 1989, 'Stage-Level and Individual-Level Predicates', in *Papers on Quantification*, pp. 147–221. GLSA, University of Massachusetts, Amherst.
Kratzer, A.: 1991, 'The Representation of Focus', in A. von Stechow and D. Wunderlich (eds.), *Semantik: Ein internationales Handuch der zeitgenössischen Forschung* (Semantics: An International Handbook of Contemporary Research), pp. 825–834. de Gruyter, Berlin.
Krifka, M.: 1995, 'The Semantics and Pragmatics of Polarity Items', *Linguistic Analysis* **25**, 209–257.
Kuno, S.: 1997, 'Binding Theory in the Minimalist Program', manuscript, Harvard University.
Lappin, S.: 1984, 'VP Anaphora, Quantifier Scope, and Logical Form', *Linguistic Analysis* **13**, 273–315.
Larson, R. K. and R. May: 1990, 'Antecedent Containment or Vacuous Movement: Reply to Baltin', *Linguistic Inquiry* **21**, 103–122.
Lasnik, H.: 1998, 'Some Reconstruction Riddles', in A. Dimitriadis et al. (eds.), *Proceedings of the 22nd Penn Linguistics Colloquium*, pp. 83–98. University of Pennsylvania, Philadelphia.
Lebeaux, D.: 1988, *Language Acquisition and the Form of Grammar*, PhD dissertation, University of Massachusetts, Amherst.

Lechner, W.: 1999, *Comparatives and DP-structure*, PhD dissertation, University of Massachusetts, Amherst.
Lobeck, A.: 1995, *Ellipsis*. Oxford University Press, Oxford.
Marti, L.: 2003, '*Only*, Context Reconstruction, and Informativity', in M. Kadowoki and S. Kawahara (eds.), *Proceedings of NELS 33*, pp. 187–203. GLSA, University of Massachusetts, Amherst.
May, R.: 1985, *Logical Form: Its Structure and Derivation*, MIT Press, Cambridge, Mass.
Merchant, J.: 2000, 'Economy, the Copy Theory, and Antecedent-Contained Deletion', *Linguistic Inquiry* **31**, 566–575.
Merchant, J.: 2001, *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford University Press, Oxford.
Munn, A.: 1994, 'A Minimalist Account of Reconstruction Asymmetries', in M. Gonzàlez (ed.), *Proceedings of NELS 24*, pp. 397–410. GLSA, University of Massachusetts, Amherst.
Pafel, J.: 1999, 'Interrogative Quantifiers within Scope', *Linguistics and Philosophy* **22**, 255–310.
Reinhart, T.: 1997a, 'Quantifier Scope: How the Labor Is Divided between QR and Choice Functions', *Linguistics and Philosophy* **20**, 335–397.
Reinhart, T.: 1997b, 'Strategies of Anaphora Resolution', OTS Working Paper 97/07, Utrecht University, Utrecht.
Romero, M.: 1997, 'Problems for a Semantic Account of Scope Reconstruction', in G. Katz and H. Winhast (eds.), *Proceedings of the Tübingen Workshop on Reconstruction*, pp. 119–146. University of Tübingen, Germany.
Rooth, M.: 1992a, 'Ellipsis Redundancy and Reduction Redundancy', in S. Berman and A. Hestvik (eds.), *Proceedings of the Stuttgart Ellipsis Workshop*. Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 29, University of Stuttgart. Published by IBM Germany, Heidelberg.
Rooth, M.: 1992b, 'A Theory of Focus Interpretation', *Natural Language Semantics* **1**(1), 75–116.
Ruys, E.: 2000, 'Weak Crossover as a Scope Phenomenon', *Linguistic Inquiry* **31**, 513–539.
Safir, K.: 1999, 'Vehicle Change and Reconstruction in A-bar Chains', *Linguistic Inquiry* **30**, 587–620.
Sag, I.: 1976, *Deletion and Logical Form*, PhD dissertation, MIT.
Sauerland, U.: 1998, *The Meaning of Chains*, PhD dissertation, MIT.
Sauerland, U.: 2000a, 'No 'no': On the Crosslinguistic Absence of a Determiner 'no'', in H. Suzuki (ed.), *Proceedings of the Tsukuba Workshop on Determiners and Quantification*, pp. 415–444. University of Tsukuba, Japan.
Sauerland, U.: 2000b, 'Obligatory Reconstruction and the Meaning of Traces', Technical Report 160, Sonderforschungsbereich 340, University of Tübingen.
Sauerland, U.: 2000c, 'Two Structures for English Restrictive Relative Clauses', in M. Saito et al. (eds.), *Proceedings of the Nanzan GLOW*, pp. 351–366. Nanzan University, Nagoya, Japan.
Sauerland, U.: 2001, 'A Contrast to a Trace', in K. Megerdoomian and L. Bar-El (eds.), *Proceedings of WCCFL 20*, pp. 498–509. Cascadilla Press, Somerville.
Sauerland, U.: 2003, 'Unpronounced Heads in Relative Clauses', in K. Schwabe and S. Winkler (eds.), *The Interfaces: Deriving and Interpreting Omitted Structures*, pp. 205–226. John Benjamins, Amsterdam.
Schachter, P.: 1973, 'Focus and Relativization', *Language* **49**, 19–46.
Schütze, C.: 1995, 'PP Attachment and Argumenthood', in C. Schütze, J. Ganger, and K. Broihier (eds.), *Papers on Language Processing and Acquisition, MITWPL 26*, pp. 95–151. MIT, Cambridge, Mass.
Schwarz, B.: 1999a, 'Silent Verb Phrases as Bound Variables', manuscript, University of Massachusetts, Amherst.

Schwarz, B.: 1999b, *Topics in Ellipsis*, PhD dissertation, University of Massachusetts, Amherst.
Sharvit, Y.: 1999, 'Connectivity in Specificational Sentences', *Natural Language Semantics* **7**, 299–339.
Sternefeld, W.: 2001, 'Semantic vs. Syntactic Reconstruction', in C. Rohrer, A. Rossdeutscher, and H. Kamp (eds.), *Linguistic Form and its Computation*, pp. 145-182. CSLI Publications, Stanford, Cal.
Tada, H.: 1993, *A/A-bar Partition in Derivation*, PhD dissertation, MIT.
Takano, Y.: 1995, 'Predicate Fronting and Internal Subjects', *Linguistic Inquiry* **26**, 327–340.
Tomioka, S.: 1999, 'A Sloppy Identity Puzzle', *Natural Language Semantics* **7**, 217–248.
Tsai, W.-T. D.: 1994, *On Economizing the Theory of A-Bar Dependencies*, PhD dissertation, MIT.
van Riemsdijk, H.: 2000, 'Free Relatives'. Syncom Case Study 44, Tilburg University.
von Stechow, A.: 2000, 'Some Remarks on Choice Functions and LF-movement', in U. Egli and K. von Heusinger (eds.), *Reference and Anaphoric Relations*, pp. 193–228. Kluwer, Dordrecht.
Wasow, T.: 1972, *Anaphoric Relations in English*, PhD dissertation, MIT.
Williams, E.: 1977, 'Discourse and Logical Form', *Linguistic Inquiry* **8**, 101–139.
Wold, D.: 1995a, 'Antecedent-Contained Deletion in Comparative Constructions', manuscript, MIT.
Wold, D.: 1995b, 'Identity in Ellipsis: Focal Structure and Phonetic Deletion', manuscript, MIT.

Department of Linguistics, U-1145
University of Connecticut
337 Mansfield Road.
Storrs, CT 06269
USA
E-mail: uli@alum.mit.edu