Copies, Presuppositions and Conservativity

Amir Anvari
MIT

Last modified: May 2, 2026

Contents

1 Introduction

Quantification in natural language is relational but restricted. On the one hand, quantificational determiners have been argued to fundamentally express binary relations between sets of individuals (Barwise & Cooper 1981 and much subsequent work). On the other hand, it has been argued that not just any relation between sets is expressed by attested determiners (also Barwise & Cooper 1981 and much subsequent work). In and out of itself the latter claim is not surprising as any decent ontology for natural language is bound to be large enough to ensure that there are more relations between sets of individuals than there are words to express them.[1] The substance of the claim, of course, is that certain relations are lexically unexpressed *and inexpressible*. If so, one naturally wants to know what the relevant constraints are and why they hold. This paper is concerned with one aspect of this question.

One purported generalization about attested determiners is that they denote conservative relations (Barwise & Cooper 1981; Keenan & Stavi 1986; see also Higginbotham & May 1981).

(1) Q is conservative if and only if for any two sets A and B, $Q(A)(B) = Q(A)(A \cap B)$

[1]If words themselves are part of the ontology, this is necessarily the case. Compare with the case of sentential connectives in the classical setting with a total of sixteen possible meanings to choose from, at least in a logical sense.

Conservativity is puzzling. One might note that there is something unexpected about how it drives a wedge between relations that are intuitively cut from the same cloth. For example, subsethood ($\subseteq$) is conservative but supersethood ($\supseteq$), proper subsethood ($\subsetneq$) and set-identity ($=$) are not.

(2) a. For any A and B, $A \subseteq B$ if and only if $A \subseteq (A \cap B)$.
b. If $A = \varnothing$ and $B \neq \varnothing$, $A \supseteq B$ is false but $A \supseteq (A \cap B)$ is true.
c. If $A = \varnothing$ and $B \neq \varnothing$, $A \subsetneq B$ is true but $A \subsetneq (A \cap B)$ is false.
d. If $A = \varnothing$ and $B \neq \varnothing$, $A = B$ is false but $A = (A \cap B)$ is true.

Why should the mind be sensitive to a distinction of this kind? The question is particularly sharp with subsethood vs supersethood as the only difference between the two pertains to an aspect of their argument structure, namely the order in which they take their arguments. This sensitivity to argument structure happens to be a general fact about *Conservativity*: the reverse of any conservative relation is a non-conservative one unless the relation is symmetric to begin with (Zuber & Keenan 2019).[2] Why should the order of arguments matter like this? And even assuming that the order of arguments should for whatever reason matter, why should it matter in the left to right direction, so to speak, with the first argument being the one that 'sets the stage' and not the other way around — why is subsethood legal but supersethood illegal, and not vice versa?

Curiously, while there is a substantial literature on semantic universals as well as on presupposition projection, it has to my knowedge never been asked if the latter has consequences for the former. This is curious because theories of presupposition projection generally make strong assumptions about how projection profiles of various operators, including determiners, relate to their basic denotations. It is natural to wonder whether these assumptions have consequences for semantic universals, such as *Conservativity*.

This is the question that I'd like to take up in this paper. Doing so is of course not possible without making explicit assumptions about presupposition projection. I will adopt trivalent semantics for my treatment of presupposition failure and model presupposition projection by assigning trivalent functions to operators. This means that the set of all conceivable denotations for determiners coincides with the set of all *trivalent* functions of type `etett`. This set is quite large. Fortunately, we're uninterested in most of these functions. Rather we are only interested in those trivalent functions that are licensed by our favorite projection theory. A projection theory is a mapping from bivalent functions to trivalent ones.[3] The idea is that given any bivalent function as the basic denotation of a given operator, the projection theory uniquely determines a trivalent function of the same type that encodes how the function is supposed to behave if its arguments end up being trivalent objects. This means that the set of projection-compliant functions, i.e. those functions that are licensed by our projection theory, is no larger than the set of bivalent functions which, while still quite large, is *much* smaller than the set of trivalent functions as such.[4] The question then becomes whether projection-compliance has consequences for semantic universals. Conceivably, projection-compliance is rich enough a notion to guarantee certain other universal properties of determiner meanings. I will argue that given independent

[2]For any Q, let $Q^{-1} = \lambda A, B.\, Q(B)(A)$. If both Q and Q^{-1} are conservative then for any A and B, $Q(A)(B) = Q(A)(A \cap B) = Q^{-1}(A \cap B)(A) = Q^{-1}(A \cap B)(A \cap B) = Q(A \cap B)(A \cap B)$. This means that Q is symmetric, $Q = Q^{-1}$.

[3]Ideally, what we should expect of a projection theory is a recipe that allows us to map bivalent functions of type τ to trivalent functions of the same type for any complex type τ. In this paper I am only considered with presupposition projection from the second argument of generalized quantifiers, consequently I will not make any commitments regarding what this general recipe might look like.

[4]If the domain of discourse contains n individuals, there are $3^{3^n \times 3^n}$ possible trivalent functions of type `etett` compared to $2^{2^n \times 2^n}$ possible bivalent functions of the same type. This massive numerical difference is one indication of how a theory of presupposition projection makes strong assumptions about the lexical semantics of determiners.

assumptions about the syntax/semantics interface, this is indeed the case.

To see why a constraint holds one might ask what would go wrong if it didn't, keeping all else fixed. The idea I'd like to explore here is that presupposition projection is what goes wrong when it comes to non-conservative determiners. Specifically, the idea is that logical forms in which non-conservative determiners occur suffer from "systematic presupposition failure" (as suggested in Fox 2002, see footnote 6 below). If so, then *Conservativity* is an expected generalization about lexical semantics of determiners as non-conservative determiners yield pragmatically useless interpretations. More specifically, I will show that if we restrict attention to what I will refer to as well-behaved logical quantifiers,[5] a theory of presupposition projection based on Schlenker 2009 coupled with (a) copy theory of movement (see e.g. Takahashi 2010a,b and references therein), according to which traces have descriptive content, and (b) Fox's (2002; 2003) rule of "trace conversation", which turns the descriptive content of traces into presuppositions, ensures that logical forms in which non-conservative determiners occur are by and large infelicitous due to systematic presupposition failure.[6] A bit more formally, suppose Q is some well-behaved, logical and bivalent quantifier and suppose $Q^{\mathcal{P}}$ is its trivalent extension according to some projection theory $\mathcal{P}$. The main result of this paper is that there is at least one independently motivated projection theory,[7] based on Schlenker 2009, such that if Q is non-conservative then $Q^{\mathcal{P}}(A)(B)$ is undefined for any A and B such that B is undefined for at least one individual that A is not true of. Now if we assume copy theory plus trace conversion, then B is guaranteed to be undefined for *every* individual that A is not true of. Put the two together and non-conservative determiners become pragmatically doomed.

This result contrasts sharply with the findings of Romoli (2015) who explores a similar "structural" approach to *Conservativity* based on copy theory but without relying on presupposition projection in his treatment of traces. Specifically, Romoli comes to the conclusion that, mainstream assumptions notwithstanding, we cannot be certain that non-conservative quantifiers have not been lexicalized in natural language. We simply wouldn't know if they were. The reason for this startling conclusion is that the non-presuppositional treatment of copy-theoretic traces that Romoli assumes actually prevents us from deciding whether any given determiner denotes a conservative quantifier or a non-conservative quantifier that happens to project a clause that denotes the same proposition. In section 2, I will discuss Romoli's account in detail and I will argue based on evidence from late merge (Lebeaux 1988 and subsequent work) that this conclusion can't be right.

In contrast to Romoli's (2015) conclusions, once presupposition projection is brought into the picture as sketched above, lexically non-conservative determiners are indeed ruled out and the problem of late merge is solved along the way. This is the topic of section 3. However, this approach has two limitations that require discussion. The first problem pertains to local accommodation, the mechanism whereby presuppositions are "wiped out" in the course of semantic composition. Since on the present account the problem with non-conservative determiners boils down to the pathological way in which they project presuppositions that are triggered by traces, something must prevent these presuppositions from being locally wiped

[5] As discussed below, by "logical" I mean quantifiers that satisfy *Permutation Invariance* and *Uniformity*. For these as well as the conception and role of well-behavedness see the discussion in section 3 below.

[6] The idea that copy theory might have consequences for *Conservativity* dates back to a talk delivered by Gennaro Chierchia at Utrecht University in 1995. The handout of that talk appears to have been lost. Ludlow (2002) and Sportiche (2005) entertain ideas to the same effect. Fox 2002 is the only precedent I am aware of that suggests presupposition failure might be the critical factor (see e.g. footnote 8 of his paper).

[7] The sheer existence of projection theories that don't mesh with non-conservative quantifiers is trivial. One can easily handcraft a projection theory that maps non-conservative quantifiers to pathological objects. The significance of the claim made in the text is supposed to be that there are *independently motivated* theories of presupposition projection that have this property. For further discussion see the concluding section.

out as otherwise non-conservative DPs are predicted to be possible with the proviso that the presuppositions triggered by their traces ought to be locally accommodated. In section 4.1, I will sketch an analysis of local accommodation according to which trace-converted DPs are predicted to be "hard" presupposition triggers in the sense that the presuppositions they trigger can never be locally accommodated. While the conception of local accommodation that I will propose is *ad hoc*, it is not entirely implausible and I will point to some independent puzzles to which it might shed some light.

The second problem pertains to DPs that end up being vacuous after they are trace-converted. If we assume that there are nouns in language that are true of every individual in the domain (e.g. perhaps *thing* or *entity*), then the prediction is made that DPs with these restrictors can be headed by non-conservative determiners as the trace that such DPs leave behind fails to trigger a falsifiable presupposition. Essentially the same problem arises in the case of "wholesale late merger" (Takahashi & Hulsey 2009), that is, syntactic configurations in which the whole NP-restrictor of a DP is merged after the DP moves out of its base position. In the latter case, the NP need not be semantically vacuous itself but because it is late merged as a whole, the trace of the DP in its base position lacks any descriptive content and therefore, again, fails to trigger a falsifiable presupposition. In section 4.2, I suggest that the domain restriction variable R be treated not as the argument of the determiner (as von Fintel 1994 would have it) but as the sister of the restrictor NP composed via Predicate Modification. The idea is that NPs, including those that undergo wholesale late merger, will be merged as sisters of R which is itself *obligatorily* base-generated along with the determiner. Now if R is prevented from denoting the set of all individuals, this move provides a solution to the problem of "vacuous" traces as the domain restriction head ends up contributing some content to the otherwise vacuous presupposition that traces trigger in such cases. With this assumption in place, vacuous traces become an impossibility; every trace-converted DP at a bare minimum introduces a presuppositional restriction that excludes the pathological individual and, as weak as this presupposition is, we will see that it is nevertheless good enough from the perspective of the present proposal.

Before we proceed, a few remarks about some self-imposed restrictions on the scope of this paper are called for. First, there is a sizable literature on apparent counterexamples to *Conservativity* both in English and from a cross-linguistic perspective. Historically significant examples from English include the "reverse proportional" reading of *many* and the use of *only* in sentences like *only students are suspects*. By and large these apparent counterexamples have been reanalyzed in ways that are compatible with the letter of *Conservativity* (see, for example, Romero 2021 for *many* and Pasternak & Sauerland 2022 for *percent*) with *only* remaining a nuisance (von Fintel & Keenan 2018; Zuber & Keenan 2019). In this paper I take *Conservativity* at face value as a generalization about the lexical semantics of determiners to be explained and leave it to future work to investigate whether the account of *Conservativity* that is proposed here sheds light on, or faces a problem in, these apparent counterexamples. Second, it has been observed that *Conservativity* seems to hold of grammatical categories other than determiners, such as quantificational adverbs (von Fintel 1994). Since it relies on copy theory of movement, it is not obvious that the account proposed here will generalize to other grammatical categories. The claim does not strike me as altogether implausible that the syntactic lives of scope-taking expressions quite generally, including quantificational adverbs, share key features with that of quantificational DPs in the object position in that the former much like the latter ought to QR from their base positions to take scope. If so then the approach proposed here does in fact extend to quantificational adverbs and so forth. But investigating this question is far beyond the scope of the present paper and I will ignore the issue in the rest of this paper.

2 Copies and conservativity

Romoli's (2015) approach to *Conservativity* is based on an idea that goes back in the literature to the dawn of copy theory of movement (see footnote 6 above). The core assumption is that when a DP moves to take scope its base position is occupied by a syntactically complex expression which is, all else equal, identical to the moved DP.[8]

(3) a. Every student is a suspect
b. [*every student*]$_n$ [~~*every student*~~$_n$ *is a suspect*]

The resulting kind of structure requires special rules for interpretation. Romoli assumes that the scope-predicate is interpreted as a property that is false of individuals that do not satisfy the restrictor. That is to say, the descriptive content of the lower copy is interpreted *assertively*. For example, the derived sister of the DP in (3-b), that is, [~~*every student*~~$_n$ *is a suspect*], should denote a predicate that is false of non-students and true of a student x iff x is a suspect. This assumption is difficult to cash out compositionally, though. For the sake of concreteness, here's a rule that hard-wires the intended outcome (modelled after Fox 2003).

(4) In a structure [DP$_n$ [$_\phi$ ... ~~DP~~$_n$ ...]] formed by DP movement, ϕ is interpreted as a predicate that maps an individual x to true if and only if the restrictor of every DP in ϕ with index n is true of x and $\phi[x/n]$ is true as well, where $\phi[x/n]$ is the result of substituting every DP with the index n in ϕ with a pronoun referring to the individual x.

So for the sentences in (5), for example, we have the structures in (6).

(5) a. Every student is a suspect.
b. ??Every student is a suspect who is a student.

(6) a. [*every student*]$_n$ [$_{\phi_1}$ [~~*every student*~~]$_n$ *is a suspect*]
b. [*every student*]$_n$ [$_{\phi_2}$ [~~*every student*~~]$_n$ *is a suspect who is a student*]

Of interest is that the higher copy of the DP gets the same semantic argument in both structures. Given the rule in (4), both ϕ_1 and ϕ_2 are true of an individual x iff x belongs to the intersection of students and suspects. It follows that quite regardless of the interpretation of the determiner *every*, the two structures in (6) are bound to have the same truth-value simply because the function $[\![$ *every student* $]\!]$, whatever it is, is fed the same argument in both cases. Thus the truth-conditional equivalence between the sentences in (5) (which presumably leads to the oddness of the more verbose one) is captured in this framework *without* making any commitments about the semantics of the determiner beyond its semantic type. So long as the task at hand is to ensure the equivalence of sentences in (5) and other similar pairs, we do not need to assume *Conservativity* as copy theory given the rule in (4) guarantees the same result *even if the determiner is not conservative!*

Where does this observation leave us with respect to *Conservativity* viewed as a generalization about the lexical semantics of determiners? The key feature of the copy-theoretic approach as sketched above is that for any determiner D and restrictor A, the semantic argument of D(A) is bound to be some subset of A, namely that subset of A that consists of individuals that satisfy the VP whatever it may be. Now, the observation is that some non-conservative functions f are

[8] In some cases, the DP in question is in a syntactic position that necessitates movement (e.g. for type reasons). But in other cases there is no independent reason to motivate movement or to prevent a moved DP from being reconstructed. I will assume that the grammar of quantification requires DPs to move to establish their scope, even if this movement is not strictly necessary to avoid type mismatch. This assumption is necessary for Romoli's account to work, as he indicates, and my own proposal as well.

such that $f(A)(B)$ yields the same truth-value for any A and B so long as $B \subseteq A$. For example, consider the functions below.

(7) a. $X = \lambda A, B.A \subsetneq B$
b. $Y = \lambda A, B.A \supseteq B$

No set is a proper subset of any of its subsets and every set is a superset of all of its subsets. It follows that, for any A and $B \subseteq A$, $X(A)(B) = 0$ and $Y(A)(B) = 1$. But then it follows that non-conservative functions like X and Y are not suitable denotations for natural language determiners as a determiner that denotes one of these functions will project a semantically trivial clause given copy theory and the rule in (4).

Encouraged by this initial observation, one might entertain the claim that *all* non-conservative functions are similarly ruled out in the copy-theoretic approach. In other words, the idea would be that *Conservativity* as a generalization about the lexical semantics of determiners is simply a side-effect of copy theory (plus a pragmatic pressure against triviality). This is the claim that Romoli (2015) shows to be incorrect. Specifically, Romoli points out that there are many non-conservative functions that yield perfectly informative interpretations even in the framework of copy theory. Here's an example.

(8) $Z = \lambda A, B.A = B$

The truth-value of $Z(A)(B)$ varies depending on A and B even if $B \subseteq A$. This is because even if B is a subset of A, there is still a distinction to be drawn based on whether B is identical with A (in which case $Z(A)(B) = 1$) or B is a proper subset of A (in which case $Z(A)(B) = 0$). It follows that triviality considerations do not rule Z out as a possible denotation for determiners despite the fact that Z is non-conservative. In conjunction with the point made in the previous paragraph, this means that only some non-conservative determiners are ruled out in the copy-theoretic approach on the charge of triviality.

Does *Conservativity* remain an open problem even in the copy-theoretic approach, then? Interestingly, Romoli argues that this is the wrong lesson to learn from the previous discussion. It is not that *Conservativity* remains unaccounted for, rather the very empirical foundation which motivated *Conservativity* to begin with is now undermined. This is because, given the assumptions sketched above, the question whether the function that a determiner denotes is conservative or not is made undecidable: We no longer have any reason to think that determiners are lexically conservative to begin with. To see this, suppose we are interested in discovering the denotation of *every* in English and suppose we are considering two hypotheses, (9-a) (according to which *every* is conservative) and (9-b) (according to which *every* is not conservative).

(9) a. $[\![\textit{every}]\!] = \lambda A, B.\, A \subseteq B$
b. $[\![\textit{every}]\!] = \lambda A, B.\, A = B$ (= Z from above)

In the classical framework, it is easy enough to argue that (9-b) is not correct. We simply have to consult our intuitions about the truth-conditions of a sentence like *every student is a suspect* and observe that this sentence does not entail that every suspect is a student. Since the classical framework is based on descriptively impoverished traces, this observation constitutes a good reason to discard (9-b). In the copy-theoretic approach, however, truth-conditional judgments of this kind are useless. For if *every student* is fed the set of suspects who are students, and not the set of suspects in general, then on both hypotheses in (9) the sentence *every student is a suspect* has the same truth-conditions. Formally, this is because if $B \subseteq A$ then $A = B \Leftrightarrow A \subseteq B$. Consequently, if we assume copy-theoretic LFs in which QR leaves behind a full-fledged copy of the moved DP along with the interpretation rule in (4), we have no truth-conditional evidence to distinguish between (9-a) and (9-b)!

It is worthwhile to state this observation in general terms. The key is that every function f of type etett has a conservative counterpart f^{CC} as defined below.

(10) $f^{CC} = \lambda A, B. f(A)(A \cap B)$.

f is conservative if and only if $f = f^{CC}$. Note that the conservative counterpart of equality is subsethood.

(11) $Z^{CC}(A)(B) \Leftrightarrow Z(A)(A \cap B) \Leftrightarrow A = (A \cap B) \Leftrightarrow A \subseteq B$

Now, if we want to single out the function that a determiner, say *every*, denotes, we must be able to somehow distinguish between non-conservative functions and their conservative counterparts as candidates for $[\![$ *every* $]\!]$. In the copy-theoretic framework, we cannot do this via the usual truth-value judgments because DPs in this framework are evaluated for subsets of their restrictors, as pointed out above, and by definition any determiner and its conservative counterpart yield the same truth-value if their second argument is a subset of the first.

(12) For any f, if $B \subseteq A$ then $f^{CC}(A)(B) = f(A)(A \cap B) = f(A)(B)$.

In other words, to distinguish between a non-conservative function f (such as (9-b)) and its conservative counterpart f^{CC} (such as (9-a)) as possible denotations for some determiner such as *every*, we need to be able to evaluate DPs of the form *every* A for some B that is *not* a subset of A. If this is not possible in the copy-theoretic approach then it follows that either neither function is a truth-conditionally plausible candidate or both are. Contra the textbooks, then, *every* may very well denote equality rather than subsethood; we simply can't be sure.[9]

Let me recap. The assumptions that Romoli (2015) adopts result in a framework which in principle allows for lexically non-conservative determiners with the proviso that clauses that these determiners project are "structurally conservative", after QR, due to the semantic contribution of the lower copy as given in (4). This structural conservativity, in turn, acts as a mask that prevents us from knowing whether the determiner in isolation is conservative or not. The only thing that we can be reasonably sure of is that those non-conservative functions that yield pathological interpretations (such as X and Y discussed above), are not lexicalized. My goal in the next section is to show that if we replace (4) with a rule according to which traces are interpreted presuppositionally then lexical conservativity can be made to follow from structural conservativity given rules of presupposition projection. Before that, though, let me first emphasize a major insight of Romoli's proposal, one that the proposal in the next section builds on, and then discuss a problem with it which the proposal in the section will overcome.

The major insight of Romoli's proposal (which goes back in the literature) is that it offers a principled explanation for what is arguably the fundamental puzzle of *Conservativity*, namely, the fact that it is sensitive to argument structure. Recall from the previous section that *Conservativity* depends on order of arguments; paradigmatically, subsethood is conservative but supersethood is not. The copy-theoretic approach yields an explanation of this. Argument structure matters because there is a syntactic difference between the arguments of a determiner that has a semantic consequence. The restrictor predicate has a syntactic copy in the nuclear-scope predicate, en-

[9]Another way to put the issue is as follows. Define the relation $\sim$ between functions of type etett so that two functions stand in this relation together iff they have the same conservative counterpart, i.e. $f_1 \sim f_2 \Leftrightarrow f_1^{CC} = f_2^{CC}$. The relation $\sim$ is an equivalence relation and therefore induces a partition on the class of all etett-functions. The claim is that while we can truth-conditionally distinguish between functions that belong to different cells in this partition, we cannot do so when the functions are cell-mates (if copy theory as sketched above is adopted). The functions $X = \lambda A, B.A \subsetneq B$ and $Y = \lambda A, B.A \supseteq B$ discussed above are special only in that their conservative counterparts are the two trivial functions $\lambda A, B.A = A$ and $\lambda A, B.A \neq A$ respectively.

suring that in the canonical cases the denotation of the latter is a subset of the denotation of the former via (4). This semantic fact in turn enters into Romoli's account of *Conservativity*. The proposal in the next section follows the same template, except that there the semantic consequence of copy theory is presuppositional and this difference will be crucial.

The discussion so far has been based on the assumption that a moved DP and its trace have the same descriptive content; we cannot truth-conditionally distinguish between a non-conservative function and its conservative counterpart *if* the second argument of the function is *guaranteed* to be a subset of the first and this guarantee stems from the assumption that DPs have the same descriptive content as their traces plus the rule in (4). But this assumption is challenged in the literature on copy theory. For example, it has been argued (Fox 2002, a.o.) that antecedent-contained deletion should be analyzed as involving the movement of the DP and late merging the relative clause that hosts ellipsis in the landing site. On this analysis, the relative clause restricts the head noun in the higher DP but not in the lower one, resulting in a mismatch in the descriptive contents of the two copies at LF. This mismatch between the two copies in turn opens a gap that we can truth-conditionally peek through to get a sense of the lexical denotation of the determiner.

(13) a. I visited every country that John did.
b. [every [country that John did visit t]]$_n$ I visited [~~every country~~]$_n$

For example, suppose we are interested in the denotation of *every* on the assumption that it is of type `etett` and suppose we are still trying to decide between the two options in (9) repeated below.

(9) a. $[\![\textit{every}]\!] = \lambda A, B.\, A \subseteq B$
b. $[\![\textit{every}]\!] = \lambda A, B.\, A = B$

In (13-b), the first argument of $[\![\textit{every}]\!]$ is the set of countries that John visited and its second argument is the set of countries that I visited. If (9-b) were correct, we would predict that (13-b) should entail that every country that I visited is one that John visited as well. This is of course not an attested entailment of (13-b). On the other hand, (9-a) allows us to predict the correct truth-conditions, namely that there is no country that John visited that I did not visit as well. As far as I can see, this point can be made with every single LF whose derivation has been argued to involve late merge; over and over again, the attested truth-conditions of these LFs are compatible with the claim that the determiner denotes a conservative function and not a non-conservative one that just happens to be related to it in a particular way.

The problem that late merge raises for Romoli, then, is that (i) in certain configurations involving late merge we can truth-conditionally distinguish between non-conservative functions and their conservative counterparts as candidates for denotations of determiners and (ii) all available evidence points to *Conservativity* as a legitimate generalization about the lexical semantics of determiners.[10] In the next section we will see that this problem disappears if traces are interpreted presuppositionally (though another problem in the vicinity, the problem of vacuous traces, poses a challenge regardless of whether traces are interpreted presuppositionally or not, see section 4.2 below).

Before I move on, let me add a note of clarification. Romoli's assumption that traces are

[10] In his paper, Romoli addresses a different problem that is raised by late merge namely that those non-conservative determiners that yield trivial interpretations are predicted to yield non-trivial interpretations if their restrictors are late merged. The problem that I am raising in the text is more general in that it indicates that, evidently, in *all* cases where late merge allows us to sneak a peek at the lexical semantics of determiners, what we see is a conservative function. I'm grateful to Danny Fox for stressing that, despite Romoli's discussion in his paper, late merge remains a problem for his account.

interpreted assertively along the lines of (4) appears to be no more than a simplification.[11] Specifically, he does not argue that traces *must* be interpreted in this way nor am I aware of any argument to this effect in the literature. Indeed, as alluded to above, this assumption is not easy to cash out compositionally to begin with. Rather Romoli proceeds on the basis of this assumption as a way to bypass the complications that would arise if traces were interpreted presuppositionally, assuming that his conclusions do not hinge on this simplification. This assumption was premature, however, as we will now see.

3 Copies, presuppositions and conservativity

I'd now like to lay out a particular theory of presupposition projection, inspired by Schlenker (2009) but implemented in the framework of trivalent semantics, and point out that given independently motivated principles of syntax/semantics interface, specifically copy theory of movement and Fox's (2002; 2003) rule of trace conversion, non-conservative determiners can *hardly ever* be used felicitously. The "hardly ever" part is important as I will also point out that while this approach goes a long way towards the goal of ruling out non-conservative determiners, it nevertheless allows them to sneak in under certain conditions. Two of these conditions are addressed in the next two sections.

Let me begin by outlining the framework just alluded to. I will make some specific, and strong, assumptions for the purpose of this paper without argument. The goal is to see whether conservativity can be adequately dealt with if presupposition projection is understood along certain lines. I leave open the extent to which what I have to say about the matter depends on the assumptions I'm about to make. For all I know, the same result (or perhaps an improvement on it) can be implemented under very different assumptions. I hope future work can address this question. Furthermore, I should add that Schlenker's theory of presupposition projection is by no means uncontroversial. It has been convincingly argued that it makes predictions that are systematically too strong, especially in the case of quantificational DPs. The obvious question to ask is if the fundamental result of this paper can be maintained if Schlenker's theory is replaced with a more adequate one, e.g. one based on "strong kleene" (**?**). I leave this important question to future work.

I will adopt trivalent semantics and I will assume that presupposition-failure is modelled with the third truth-value. The burden of presupposition projection is on the meanings of various operators, which therefore will have trivalent denotations. Just how a sentence with a certain kind of presupposition ought to be used in discourse is an issue that I will leave aside here save for an assumption that is uncontroversial as far as I know, namely that sentences with contradictory presuppositions are infelicitous in every context. I assume the 'projection profile' of any given operator can be fully recovered from its classical (bivalent) meaning. The necessary ingredient is the theory of presupposition projection $\mathcal{P}$ which I assume is a mapping from bivalent functions F to trivalent functions $F^{\mathcal{P}}$ of the same type. A trivalent function F' is $\mathcal{P}$-compliant if there is some bivalent function F such that $F' = F^{\mathcal{P}}$. Not any mapping from bivalent to trivalent functions is a projection theory, however. Inherent to the idea of a projection theory is that $F^{\mathcal{P}}$ agrees with F on all bivalent inputs; the projection theory is meant to fill in the gaps that result from trivalent or partial inputs, keeping all else fixed. In the case of quantifiers Q of type `etett`, bivalent and trivalent functions are defined as expected and a projection theory $\mathcal{P}$

[11] Romoli notes, "For the sake of the discussion, I will make some simplifications concerning the semantics of chains, but nothing hinges on these assumptions. The proposal here is compatible with a class of semantics of chains, as long as somehow the NP part of the copies is interpreted at the tail and at the head of the chain".

is a mapping from the former to the latter such that $Q^{\mathcal{P}}$ agrees with Q on all bivalent inputs.[12,13]

(14) a. $Q \in \mathcal{Q}^2$ is a bivalent quantifier on E if $Q : 2^E \to (2^E \to 2)$ where $2 = \{0, 1\}$.
b. $Q \in \mathcal{Q}^3$ is a trivalent quantifier on E if $Q : 3^E \to (3^E \to 3)$ where $3 = \{0, \#, 1\}$.
c. $\mathcal{P} : \mathcal{Q}^2 \to \mathcal{Q}^3$ is a projection theory iff $\forall A, B \in 2^E : Q^{\mathcal{P}}(A)(B) = Q(A)(B)$.

As advertised, in this section I will be primarily concerned with a certain projection theory that is inspired by Schlenker's (2009) "local-contexts theory" although it is implemented in the framework sketched above. Some differences with Schlenker's original implementation are worth listing immediately for the sake of orientation. His implementation is context-dependent, therefore intensional, unlike the present approach which is extensional. The semantics that he assumes is classical with presuppositions ear-marked for pragmatic and not semantic purposes, while of course here I will assume the trivalent treatment of presuppositions. Finally, his theory of presupposition projection is sensitive to certain aspects of syntactic representation (scope and binding) and phonological representation (linear order) that the present implementation does not rely on at all. Despite these differences the two implementations are descriptively equivalent as far as presupposition projection from the nuclear-scope of quantificational structures is concerned, which is my sole focus here.[14]

I will first outline the theory with broad strokes. Take *every student smiled*. This sentence is true if all students smiled and false if there is at least one student who didn't. To verify this sentence, therefore, we need not 'tag' every individual there is with the truth-value that *smiled* assigns to that individual; we can safely ignore all non-students, which are irrelevant to the truth-value of the sentence, and look merely at how *smiled* behaves on the set of individuals that in principle matter, that is, the set of students. Now take *every student stopped smoking*. The idea of our projection theory is that we proceed as before. We ignore individuals that do not matter, i.e., non-students, and we consult the truth-value that *stopped smoking* assigns to every individual that in principle matters, i.e., students. If *stopped smoking* assigns a determinate truth-value (true or false) to all students, then we have all we need to decide if the sentence as a whole is true or not, regardless of whether *stopped smoking* is undefined for some non-students as the case may be. However if *stopped smoking* is undefined even for a single student, then we throw in the towel: the sentence is undefined in that case. This is the intuition that is fleshed out below. Note that this informal description of the projection theory under consideration makes it look as though *Conservativity* is a crucial *assumption* of the theory. But we will see that in fact we can coherently and non-trivially flip the order of explanation so that *Conservativity* becomes a *consequence* of the theory.

Let us get clear about the key notion, 'the set of individuals that in principle matter', in the bivalent setting. As a matter of policy, I will switch to set-theoretic notation when dealing with bivalent functions hoping that this will make it easier for the reader to keep track of where we are. To explicate this notion of 'individuals that in principle matter' we can use the *live-on* relation

[12] E is the domain of discourse and 2^E is the set of all functions from E to 2. Similarly for 3^E.

[13] Note that I am excluding from consideration quantifiers that are themselves presupposition triggers, e.g. *both*. This exclusion is innocuous, though, and the set of bivalent quantifiers can be defined as $2^E \to (2^E \to 3)$ without any substantive change to the discussion below as far as I can see.

[14] Schlenker is more ambitious than I am (here), as he is interested in spelling out a projection theory that applies across the board, not just to quantificational structures. (See also footnote 3 above.) Consequently, the observation in the text should not be construed as the claim that the theory discussed here is more parsimonious than Schlenker's because it may very well turn out that assumptions pertaining to scope, binding and linear order are unavoidable ingredients for an adequately general theory of presupposition projection. Whether that's the case or not is an open question at this point, I believe, though not one that I aim to take a stance on here. It *is* true that the conception of projection theory that I am working with in this paper is most naturally viewed as strictly lexical, making the role of syntactic and phonological considerations at best indirect.

introduced in Barwise & Cooper (1981).

(15) $Q(A)$ lives on C if and only if $\forall B : Q(A)(B) = Q(A)(C \cap B)$

If $Q(A)$ lives on C then $Q(A)$ lives on any superset of C as well,[15] so in general $Q(A)$ lives on several sets. However, often enough we can simply focus our attention on the *smallest* live-on set of $Q(A)$, which I will notate as $\Sigma(QA)$.

(16) $\Sigma(QA) = C$ if and only if,
 a. $Q(A)$ lives on C and
 b. For any C', if $Q(A)$ lives on C' then $C \subseteq C'$.

Now, the reason that smallest live-on sets constitute a decent explication of 'the set of individuals that in principle matter' is the following lemma.[16] In a nutshell, to say that an individual belongs to the smallest live-on set of $Q(A)$, all we have to do is to find two predicates that differ on the value they assign to that individual but are otherwise identical and show that $Q(A)$ distinguishes between them. Stated differently, if an individual x does not belong to the smallest live-on set of $Q(A)$, then the lemma below (see A for proof) tells us that x can be safely ignored in the sense that, for any B, the truth-value of $Q(A)(B)$ does not depend on whether B is true of x or not.

(17) **Lemma.** For any Q and A, if $\Sigma(QA)$ exists then,

$$x \in \Sigma(QA) \Leftrightarrow \exists B : Q(A)(B) \neq Q(A)(B \cup \{x\})$$

For any determiner *det*, what Schlenker takes to be the "local context" of e.g. *stopped smoking* in *det student stopped smoking* in context c is in effect the intensional property of being a member of the smallest live-on set of *det student* relative to c, (18). This is the fundamental reason why the projection theory that I'm about to introduce is descriptively equivalent to Schlenker's theory as far as presupposition projection from the second arguments of determiners is concerned.

(18) $\lambda w, x.\ w \in c \,\&\, x \in \Sigma([\![\textit{det student}]\!]^w)$

Importantly, $\Sigma(QA)$ sometimes does not exist. This happens if $Q(A)$ traffics in infinity. For example, *all but finitely many*, viewed as a complex determiner, lacks a smallest live-on set (von Fintel & Keenan 2018) and so does *infinitely many* (Schlenker 2009).[17] Specifically, the following

[15]Suppose $Q(A)$ lives on C and let $C \subseteq C'$. For any B, $Q(A)(C' \cap B) = Q(A)(C \cap C' \cap B)$, because $Q(A)$ lives on C by assumption. Since $C \subseteq C'$, it follows that $Q(A)(C \cap C' \cap B) = Q(A)(C \cap B)$. Again, since $Q(A)$ lives on C, it follows that $Q(A)(C \cap B) = Q(A)(B)$. Therefore, for any B, $Q(A)(C' \cap B) = Q(A)(B)$.

[16]In his lecture notes on presupposition projection, Danny Fox has used the lemma in (17) to directly define the set of individuals that in principle matter, bypassing smallest live-on sets.

[17]Consider the domain of all, say, natural numbers and the hypothetical DP *infinitely many* of type `ett` with the following semantics. Recall that any set of natural numbers S has a bottom element, $min(S) \in S$.

(i) $[\![\textit{infinitely many}]\!] = \lambda B. \neg\exists n \in \mathbb{N} : \forall m \in B : n \geq m$

Define,

(ii) a. $\mathbb{N}^0 = \mathbb{N} = \{1, 2, \dots\}$
 b. $\mathbb{N}^{n+1} = \mathbb{N}^n \setminus \{min(\mathbb{N}^n)\}$

For example, $\mathbb{N}^1 = \{2, 3, \dots\}$ and $\mathbb{N}^2 = \{3, 4, \dots\}$, etc. Let $n \geq 0$. For any set B, either B is true of only finitely many natural numbers or it is true of infinitely many natural numbers. In the former case, both *infinitely many*(B) and *infinitely many*$(\mathbb{N}^n \cap B)$ are false and in the latter case both are true. Therefore, *infinitely many* lives on $\mathbb{N}^n$ for any $n \geq 0$. But there is no set of natural numbers that is a subset of $\mathbb{N}^n$ for all $n \geq 0$. Therefore, *infinitely many* does not have a smallest live-on set.

lemma is easy to establish[18] and it entails that if $Q(A)$ lacks a smallest live-on set then every set that it lives on is infinite. This is the sense in which $Q(A)$ "traffics in infinity".

(19) $Q(A)$ does not have a smallest live-on set iff, for any set C, if $Q(A)$ lives on C then there is a proper subset C' of C (i.e. $\varnothing \subsetneq C' \subsetneq C$) such that $Q(A)$ lives on C' as well.

That said, in the discussion below quantifiers that lack smallest live-on sets will be ignored. This is more an indication of my limitations than the substance of the results — I have simply failed to prove the key lemmas in the most general form that would include cases where the smallest live-on sets do not exist. I hope future work can address this limitation. For the time-being, let me say that a bivalent quantifier Q is *well-behaved* if for any set A, $Q(A)$ has a smallest live-on set and let me restrict the discussion below to well-behaved quantifiers.

We are now in a position to spell-out what kind of a projection theory we get on the basis of these considerations. Call this projection theory S. Let Q be some bivalent quantifier and let us denote by Q^{S} the trivalent extension of Q according to S. As a first pass, we can define Q^{S} as follows.

(20) $Q^{\mathrm{S}} = \lambda A, B : \Sigma(QA) \subseteq \Delta B.\, Q(A)(\nabla B)$ (*to be revised*)

I'm relying on two auxiliary notions in this formulation. First, for any trivalent predicate B, ΔB is B's domain, that is, the set of individuals that B maps to either true or false. Second, for any trivalent predicate B, ∇B is the set of individuals that B maps to true.

(21) a. $\Delta B = \{x : B(x) \neq \#\}$
b. $\nabla B = \{x : B(x) = 1\}$

So, according to (20), to evaluate $Q^{\mathrm{S}}(A)(B)$, where A and B are trivalent predicates, we first check whether the smallest live-on set of the bivalent quantifier Q given restrictor A, i.e., $\Sigma(QA)$, is a subset of B's domain (B is defined for every individual that in principle matters). If the condition is met, we proceed to simply return the truth-value of $Q(A)(\nabla B)$, with the presupposition of B wiped out. Otherwise, $Q^{\mathrm{S}}(A)(B)$ is undefined.

Two notes. First, as stressed above, if the condition $\Sigma(QA) \subseteq \Delta B$ is met, it doesn't really matter how we deal with individuals that B is undefined for (since those individuals will be outside of $\Sigma(QA)$). In (20) I have chosen to map these to false using ∇. I could just as well have mapped them to true. Second, there is a problem with (20) as it only makes sense if A is bivalent. In that case, it tells us how to deal with B if it happens to be partial. But what if A itself is partial? Nothing that I will say here depends on how presupposition projection from restrictors is supposed to work; all that matters is presupposition projection from nuclear-scopes. Consequently, I will simply assume that any presupposition triggered in the restrictor is automatically wiped out (again, any other assumption will do just as well as far as the results below are concerned). The result, (22), is the formalization of the projection theory S for my purposes here. Note that I am following my policy of working with sets in the bivalent setting but with functions in the trivalent setting. Thus, bivalent quantifiers Q take sets as arguments but their trivalent extensions Q^{S} take functions of type `et` as arguments. This is merely for expository purposes of course.

[18] Every quantifier lives on the domain of discourse, therefore it cannot be the case that $Q(A)$ does not have a smallest live-on set because it does not live on any set. Furthermore, it is easy to see that if $Q(A)$ lives on C and C' then it must live on $C \cap C'$ as well. Therefore, it cannot be that $Q(A)$ does not have a smallest live-on set because it has at least two live-on sets that are not ordered with respect to subsethood. The only remaining option, then, is that $Q(A)$ lacks a smallest live-on set because for any live-on set of $Q(A)$ there is another, smaller live-on set of $Q(A)$ *ad infinitum*, as seen in the example discussed in fn. 17.

(22) $Q^{\text{S}} = \lambda A, B : \Sigma(Q\nabla A) \subseteq \Delta B.\, Q(\nabla A)(\nabla B)$

With the projection theory S in place, the next step is to have a characterization of the projection-profile of various kinds of DPs, including non-conservative ones. What we need is a characterization of what $\Sigma(QA)$ amounts to depending on whether or not Q lives on A. Now, the following fact follows immediately from the relevant definitions.

(23) $Q(A)$ lives on A if and only if $\Sigma(QA) \subseteq A$.

But beyond (23), $\Sigma(QA)$ can in principle be just about any set. To get a sense of why this is, consider two imaginary determiners.

(24) a. $[\![\textit{every-minus-John}]\!] = \lambda A, B.\, A \setminus \{\textit{John}\} \subseteq B$
b. $[\![\textit{every-plus-John}]\!] = \lambda A, B.\, A \cup \{\textit{John}\} \subseteq B$

It is easy to see that the smallest live-on set of *every-minus-John student* is the set of students minus John and the smallest live-on set of *every-plus-John student* is the set of students plus John. There is obviously a lot of room for creativity here and this makes it difficult to formulate a crisp characterization of what sort of set $\Sigma(QA)$ might be in the general case. But the reason why this is difficult to achieve pertains largely to the 'non-logical' part of the meaning of Q (e.g., reference to John in the above examples). If following much of the literature we assume that determiners denote logical quantifiers, we can sharpen (23) quite a bit. Specifically, one ingredient of 'logicality' is often assumed to be *Permutation Invariance* (Gamut).

(25) a. Q is permutation invariant if and only if for all permutations π of the domain of individuals E and all $A, B \subseteq E$, $Q(A)(B) = Q(\pi A)(\pi B)$.
b. For any permutation π of E and $A \subseteq E$, $\pi A = \{\pi(x) : x \in A\}$.
c. π is a permutation of E if and only if it is a bijection from E to E.

As it happens, for permutation invariant quantifiers there are merely four possibilities for what the smallest live-on set of $Q(A)$ might be, depending on Q and A (see the appendix for proof).

(26) **Lemma.** If E is the domain of discourse and Q is a permutation invariant, well-behaved quantifier on E, then for any $A \subseteq E$,
a. If $Q(A)$ lives on A then $\Sigma(QA)$ is either $\emptyset$ or A,
b. If $Q(A)$ does not live on A then $\Sigma(QA)$ is either E or $E \setminus A$.

Let us see some examples that instantiate these four possibilities. In general, if the value of $Q(A)(B)$ does not depend on what B is, $\Sigma(QA)$ is the empty set. Thus in both examples below the smallest live-on set is empty. Note that these examples are also vacuously conservative.

(27) a. $\lambda A, B.\, A = A$
b. $\lambda A, B.\, |A| = 2$

The case where the smallest live-on set coincides with the restrictor is the familiar one involving all conservative determiners. Putting these aside, we now enter the realm of the non-conservative. Take Z again, i.e. set-equality.

(28) $Z = \lambda A, B.\, A = B$

To see that the smallest live-on set of $Z(A)$ for any A is just E, take x to be some individual. There are two possibilities, either x is in A or it is not. If the former, then consider A and $A \setminus \{x\}$. These sets differ only on x. Obviously, $Z(A)(A) = 1$ and $Z(A)(A \setminus \{x\}) = 0$. So, by (17), x must be in $\Sigma(ZA)$. Suppose x is not in A. In this case, consider A and $A \cup \{x\}$. Again, $Z(A)(A) = 1$ and $Z(A)(A \cup \{x\}) = 0$, therefore, by (17), x must be in $\Sigma(ZA)$. We conclude that $\Sigma(ZA) = E$, for any $A \subseteq E$.

Finally, consider supersethood.

(29) $Y = \lambda A, B.\, B \subseteq A$

Because statement $B \subseteq A$ is equivalent to $B \cap (E \setminus A) \subseteq A$, $Y(A)$ lives on $E \setminus A$. To see that the latter is the smallest set that $Y(A)$ lives on, suppose C is a set that $Y(A)$ lives on and suppose x is in $E \setminus A$ but not in C. Consider $Y(A)(A \cup \{x\})$. Since $x \notin A$, $Y(A)(A \cup \{x\}) = 0$. Now consider $Y(A)(C \cap (A \cup \{x\}))$. Since $x \notin C$, $C \cap (A \cup \{x\}) = C \cap A$ is some subset of A, therefore $Y(A)(C \cap (A \cup \{x\})) = 1$. But this means that $Y(A)$ does not live on C. We conclude that $\Sigma(YA) = E \setminus A$, for any $A \subseteq E$.

So, (26) clears things up as far as what sort of object $\Sigma(QA)$ might be depending on whether or not $Q(A)$ lives on A assuming that Q is permutation invariant and well-behaved (i.e. has a smallest live-on set). We need one more ingredient before we move on to the LF side. A conservative function Q, of course, is such that $Q(A)$ lives on A for any A. Now, non-conservative functions come in two varieties. Some are *totally* non-conservative and some are only *partially* non-conservative.

(30) Q is totally non-conservative iff for any $A \subsetneq E$, $Q(A)$ does not live on A.

Most non-conservative functions that readily come to mind are totally non-conservative in this sense; we have seen several examples repeatedly above. But there are also partially non-conservative functions. If Q is such a quantifier, then it is possible to find sets $A, A' \subsetneq E$ such that $Q(A)$ lives on A but $Q(A')$ does not live on A'.[19] Here's an example.

(31) $\lambda A, B.\, (|A| \neq 2 \wedge A \supseteq B) \vee (|A| = 2 \wedge A \subseteq B)$

Intuitively, this function does two different things depending on what its first argument is. If A is true of exactly two individuals, it just computes subsethood; otherwise, it computes supersethood. So for any $A \subsetneq E$, $Q(A)$ lives on A if and only if $|A| = 2$. That's why this function is only partially non-conservative (assuming that E contains more than one individuals). The account that I will discuss in a moment effectively predicts that e.g. *det student smiled* can only be used felicitously if *det student* lives on the set of students. So, it rules out all totally non-conservative determiners, but partially non-conservative determiners are in principle allowed with the proviso that they can only be used felicitously if they live on their first argument. (31), then, is in principle a possible a denotation for a natural language determiner, but it can only be used if its restrictor happens to be true of only two individuals.[20] Ideally, we'd like to get rid of partially non-conservative functions as well as totally non-conservative ones.

Now, it seems to me that there are independent reasons to exclude partially non-conservative determiners. For one thing, partially non-conservative functions raise a serious challenge for learnability. If these functions are in principle allowed, then one cannot infer on the basis of any amount of less than total evidence that a given determiner is conservative since it is quite possible that the determiner is non-conservative with respect to arguments not yet encountered. More generally, Gamut (p. 255 vol. 2) point out that all natural language determiners are 'uniform',

[19] If E is the domain of discourse, all quantifiers on E live on E by definition. That's why to see if a quantifier is partially non-conservative we need to look at *proper* subsets of E.

[20] Jad Wehbe and Benjamin Spector have pointed out (p.c.) that perhaps there are ways to put partially non-conservative functions to good use. For example, could it be that something like (31) is actually the denotation of English *both*? This is an interesting idea but I suspect it is wiser to leave this particular can of worms closed.

intuitively in the sense that what they do does not depend on what their arguments are.[21] The concept is difficult to formalize as one would need to make explicit what a determiner does in the form of, say, a program or an automaton. Be that as it may, I believe *Uniformity*, however it is cashed out, amongst other things rules out partially non-conservative functions given their inherently disjunctive nature and it seems to me that even very abstract considerations pertaining to learnability point in the same direction as well. Without dwelling on the details any further, then, let me assume that partially non-conservative functions are already excluded from consideration by *Uniformity*. To avoid clutter, I will subsume *Uniformity* along with *Permutation Invariance* under the label "logical".

Let us now spell out the assumptions needed on the LF side. First, all DPs, including those that can in principle be interpreted *in situ*, must undergo QR to take scope. Second, QR is to be understood in copy-theoretic terms in that what's 'left behind' by QR is a copy of the moved DP and not merely a descriptively impoverished variable. For example, the LF for a sentence like *det student smiled* is something like this.

(32) [*det student*] λn [[~~*det student*~~]$_n$ *smiled*]

The third assumption pertains to the interpretation of the lower copy. I assume, following Fox (2002, 2003) that the lower determiner is replaced with an indexed definite article and the scope-predicate, therefore, gets a presuppositional interpretation.[22]

(33) a. [*det student*] λn [[*the$_n$ student*] *smiled*]
b. $[\![\textit{the}_n]\!]^g = \lambda P_{et} : g(n) \in P.\, g(n)$
c. $[\![\lambda n\ \textit{the}_n\ \textit{student smiled}]\!] = \lambda x : \textit{x is a student}.\ \textit{x smiled}$

We are finally in a position to see what goes wrong with non-conservative determiners. Suppose $[\![\textit{det}]\!] = Q$ assuming that Q is logical and well-behaved. To compute the truth-value of *det student smiled* given the LF in (32) we have to evaluate $Q^{\mathrm{S}}(\textit{student})(\textit{student} : \textit{smiled})$, where *student* is short for $[\![\textit{student}]\!]$ and *student* : *smiled* is short for the denotation of the scope-predicate, (33-c). Given (26) and assuming that not all individuals are students, it is now easy to see how things will play out. If Q is uniformly non-conservative, then $Q(\textit{student})$ does not live on *student* and $\Sigma(Q\,\textit{student})$ is either E or $E \setminus \textit{student}$. Either way $\Sigma(Q\,\textit{student}) \not\subseteq \textit{student}$ and $Q^{\mathrm{S}}(\textit{student})(\textit{student} : \textit{smiled})$ will be undefined. Stated in general terms, given the projection theory S, trace conversion triggers the presupposition that the smallest live-on set of the DP must be a subset of its restrictor, which amounts to the presupposition that D must live on its restrictor. If D is logical and well-behaved, this means that D must be conservative.

This concludes the discussion of the present proposal. The core assumptions were copy theory plus trace conversion, and a theory of presupposition projection inspired by Schlenker's (2009) "local-contexts theory" along with auxiliary assumptions about the denotations of determiners, namely that they should be well-behaved, permutation invariant and uniform. In the next section I will stress-test the proposal in various ways. Before that, though, let me quickly point out that the present proposal does not face the problem with late merge as Romoli's (2015) proposal did.

Consider again the sentence below this time using trace conversion.

(34) a. I visited every country that John did.
b. [every [country that John did visit t]] λn I visited the$_n$ country

[21] A similar constraint, which I'd subsume under *Uniformity*, is *Extension* which requires quantifiers to 'do the same thing' in every domain on which they are defined.

[22] Just how the so-called 'trace conversion rule' ought to be implemented does not matter for my purposes; all that matters is that the scope-predicate somehow comes to be defined only for individuals in the restrictor.

Note that the argument of the DP *every country that John did ~~visit t~~* is a predicate that is only defined for countries. If *every* denotes a non-conservative function, then the smallest live-on set of *every country that John did ~~visit t~~* is either the set of all individuals or the set of all individuals minus the countries that John visited. Either way, non-countries belong to the smallest live-on set of *every country that John did ~~visit t~~* and the nuclear-scope predicate is undefined for these individuals. It follows that if *every* is non-conservative, this LF is predicted to be undefined (again, so long as there are individuals in the domain that are not countries).

More generally, we have established the following result in this section which, coupled with copy theory and trace conversion, guarantees that so long as the restrictor of the trace of a DP is entailed by (or is a superset of) the restrictor of the moved DP itself, and is not vacuous, D has to be conservative for the structure to avoid presupposition failure.

(35) Suppose Q is a bivalent quantifier that is well-behaved, permutation invariant and uniform. If Q is non-conservative then $Q^{\mathrm{S}}(A)(B)$ is undefined for any A, B so long as there is at least one individual x such that $A(x) \neq 1$ and $B(x) = \#$.

4 Further thoughts

4.1 On local accommodation

There is evidence that the grammar is equipped with some mechanism or other to 'wipe out' presuppositions. For example, for the sentence in (36) to receive its intuitively attested, coherent interpretation there must be some way to collapse the presupposition-assertion distinction as otherwise standard assumptions about presupposition projection in disjunction one way or the other predict an incoherent interpretation.

(36) (Why is he so jittery?) Either he has started smoking or he has stopped smoking.

A standard approach is to posit a 'local accommodation operator', with a semantics rather similar to the meta-linguistic ∇-operator defined in (21-a) above, that gets strategically inserted at LF as a rescue mechanism to avoid presupositional incoherence. Within trivalent semantics, this operator standardly gets the bare-bones, truth-functional analysis below.

(37) $[\![\text{A}]\!] = \lambda t.\, t = 1$

For example, if the sentence in (36) is parsed with A in both disjuncts it comes to mean that either he never used to smoke but smokes now or he used to smoke but doesn't smoke now, which is intuitively the correct reading for this sentence.

(38) [A he has started smoking] or [A he has stopped smoking]

The problem for any attempt to reduce *Conservativity* to how presupposition projection works is that if such an operator exists, then something must prevent it from applying to presuppositions that are triggered by trace-converted DPs. Otherwise, non-conservative determiners can sneak back in the lexicon but come with the requirement that their nuclear-predicate must be parsed with A.

(39) a. [*det student*] λx A [[*the*$_x$ *student*] *smiled*]
b. $[\![$ λx A *the*$_x$ *student smiled* $]\!]$ = λx. *x is a student and x smiled*

In fact, the result of adding A in such structures is that we are now back to the situation as it was envisioned by Romoli (2015) since the scope-predicate no longer presupposes the restrictor but asserts it.

What seems to be needed is a way to make sure that local accommodation cannot apply to presuppositions triggered by traces. Let me spell out one way in which this can be done with potentially interesting consequences. Suppose that, *contra* standard assumptions, A is an intensional operator: it takes a proposition p and a world w and returns true if p is true in w and returns false otherwise.

(40) $[\![\text{A}]\!] = \lambda w.\lambda p.\, p(w) = 1$ (*to be revised*)

Now, if A is an intensional operator then it stands to reason that it should generate *De Dicto* / *De Re* ambiguities. Stated differently, it is natural to think that if a certain presupposition trigger is interpreted *De Re* with respect to A, i.e., its world argument is not locally bound by A, then the presupposition that it triggers should not be affected by A. The entry above does not allow this but we can implement the idea if we assume that, somewhat counter-intuitively, A itself is a presupposition trigger; namely, it presupposes that its prejacent should be 'informative' in the sense that it can possibly be true and it can possibly be false.

(41) $[\![\text{A}]\!] = \lambda w.\lambda p : \exists w' : p(w') = 1 \,\&\, \exists w' : p(w') = 0.\, p(w) = 1$

Let us see how this works with a very simple example. Consider the sentence *the student smiled*. Without A, the sentence carries the usual uniqueness presupposition. If we parse the sentence with A, assuming that intensional dependencies are syntactically explicit, there are two possible parses depending on whether the world argument of the restrictor is locally bound by A, as in (42-a), or not, as in (42-b).

(42) a. $\lambda w_@ \, \text{A}_{w_@} \, \lambda w' \, \textit{the student}_{w'} \, \textit{smiled}_{w'}$
b. $\lambda w_@ \, \text{A}_{w_@} \, \lambda w' \, \textit{the student}_{w_@} \, \textit{smiled}_{w'}$

Given the entry in (41), these LFs are *not* synonymous. (42-a) is the one in which the uniqueness presupposition of the definite is accommodated and the structure as a whole is false in worlds in which there are more or less than one students. What about the LF in (42-b)? The prejacent of A in this LF denotes the proposition in (43). The presupposition of this proposition depends on $w_@$, so what the proposition in (43) amounts to depends on what kind of world $w_@$ is.

(43) $\lambda w' :$ *there is exactly one student x in* $w_@$. *x smiled in* w'

If in $w_@$ there is exactly one student, then the presupposition of (43) is satisfied and the sentence is true or false depending on whether that one student smiled or not. But if there are more or less than one students in $w_@$, then (43) is the 'null' proposition, that is, the proposition that is always undefined. Given the entry in (41), A itself presupposes that its prejacent must not denote the null proposition. It follows that if there are more or less than one students in $w_@$, then the sentence as a whole is undefined. In other words, in (42-b) the local accommodation operator is simply redundant because the only presupposition trigger in its domain is interpreted *De Re* with respect to it.

In light of this example, consider again (39-a) this time with world variables represented explicitly. In this LF, I have assumed that both copies of the restrictor should be bound by the same world-binder and therefore, since the higher copy is not c-commanded by $\lambda w'$, both copies are bound by $\lambda w_@$ as the only viable option.

(44) $\lambda w_@ \, [\textit{det student}_{w_@}] \, \lambda x \, \text{A}_{w_@} \, \lambda w' \, [[\textit{the}_x \, \textit{student}_{w_@}] \, \textit{smiled}_{w'}]$

Now, what does the derived argument of the DP denote? By the same reasoning as above, it turns out that since the restrictor in the lower copy is not world-bound by A, the presupposition that it triggers projects through A: the scope-predicate is still undefined for individuals who are not students (in $w_@$). In other words, the entry in (41) in conjunction with some auxiliary assumptions about world variables, the most crucial being that copies of the same DP must be evaluated with respect to the same world variable, guarantees that the presupposition triggered by the trace of DP can never be locally accommodated below the DP.[23]

Of course, this is all quite *ad hoc*. Ideally, one would want to have an independent motivation for the entry in (41). Without engaging in the details, as I cannot possibly do justice to them, let me sketch a few predictions that are testable. To begin with, this analysis predicts that presuppositions that are not dependent on the world of evaluation can never be locally accommodated. Examples include indexical presuppositions, e.g. first- and second-person features on pronouns (e.g. Heim 2005). If Schlenker (2007) is on the right track, expressives fall in the same category as do, for example, presuppositions that pertain to there being contextually salient objects of appropriate kinds (e.g. additive presuppositions) and presuppositions pertaining to non-emptiness of contextually constrained sets, be they focus-alternatives or comparison classes. None of these presuppositions can be locally accommodated according to the current analysis, which seems about right. Furthermore, even presuppositions that are dependent on the world of evaluation can in principle be distinguished into those that resist *De Dicto* interpretations and those that do not. For example, Sudo (2012) has argued that gender features on pronouns resist both *De Dicto* interpretations and local accommodation. If correct, on the present analysis the former generalization entails the latter. On the standard assumption about the semantic of the local accommodation operator, however, the two generalizations are not obviously related. On the other hand, world-dependent presuppositions that can be easily read *De Dicto* (e.g. factive presuppositions) are predicted to be easy to locally accommodate. This too seems about right. See also Doron & Wehbe (2022) for a constraint on local accommodation that the analysis above might shed some light on. In a nutshell, Doron & Wehbe argue that if an expression triggers a presupposition that entails its assertive content in the local context of the expression, then this presupposition cannot be locally accommodated. This generalization can be made to follow from the analysis of local accommodation above on the assumption that the presupposition of A involves existential quantification over worlds in its local context.

4.2 On vacuous traces

Suppose the noun *thing* is true of every individual in the domain.

(45) $[\![\text{thing}]\!] = \lambda x.\, x = x$

On this assumption *thing* creates a problem for the account of *Conservativity* that I have been discussing. The reason is that trace-converted DPs whose only descriptive content is contributed by *thing* do not trigger a falsifiable presupposition: the derived sister of *det thing* below is defined for every individual as far as the trace is concerned because *thing*, by assumption, is true of every individual.

(46) [det thing] λx [… [the$_x$ thing] …]

But if that's the case then nothing prevents *det* from denoting a non-conservative function; since the nuclear-scope fails to trigger a presupposition due to trace conversion (and other presupposi-

[23]The other assumptions that are needed are these: (i) every sentence must be parsed with a global world binder, (ii) all world variables must be bound either locally or by the global binder, and (iii) 're-binding' must be impossible (i.e., $* \ldots \lambda w_i \ldots \lambda w_i \ldots$).

tions triggered in the scope-predicate, if any, can in principle be locally accommodated), there is no problem of presupposition failure in such structures.

Perhaps there are no nouns that are true of every individual in the domain, then. But fundamentally the same problem arises with "wholesale late merger" (Takahashi & Hulsey 2009). For reasons that we need not get into here, Takahashi & Hulsey propose that A-movement may leave behind a content-less trace that consists merely of the determiner with the restrictor entering the picture higher up in the landing site via late merge. Assuming that the stranded determiner after trace conversion is interpreted as a mere variable bound by the DP (details do not matter here), in this kind of structure as well the derived sister of the DP fails to trigger a hard and falsifiable presupposition. So, again, the prediction is made that in such configurations the determiner may very well denote a non-conservative function.

(47) a. Det argument seems to be correct.
b. [det argument] λx seems to be [the$_x$] correct

Ideally the problem of vacuous traces should be solved in the same way across the board. Here I'd like to suggest one conceivable line of attack. It is common ground that the domain of quantification is restricted both overtly by linguistic material in the restrictor and covertly through contextual domain restriction. The most prominent approach to the latter (von Fintel 1994) is to postulate a variable of type `et` that merges directly with the determiner creating a generalized quantifier of type `etett`. This means that strictly speaking the determiner itself is of type `etetett`. For example,

(48) $[\![\text{every}]\!] = \lambda R, A, B.\, R \cap A \subseteq B$

An alternative is to attach the domain restriction variable directly to the NP, i.e. to have (49-b) instead of (49-a), on the assumption that the domain restriction variable composes with the NP via Predicate Modification.

(49) a. [every R] student
b. every [R student]

If so, then the problematic structures above should be analyzed as follows.

(50) a. [det [R thing]] λx [. . . [the$_x$ [R thing]] . . .]
b. [det [R argument]] λx seems to be [the$_x$ R] correct

If it is assumed that domain restriction is never vacuous, that is, R necessarily denotes a set that excludes some individuals, then even the traces in (50) trigger falsifiable presuppositions; that is, the scope predicates are now undefined for individuals that do not belong to the denotation of R. But what prevents R from being vacuous? Unfortunately, I have to leave this question open here.

5 Conclusion

My goal in this paper was to explore a way of thinking about semantic universals that has so far been neglected (to my knowledge). The common attitude in the literature is to ignore presuppositions altogether and focus on the classical, bivalent meanings of the relevant lexical items. My point here was that this simplifying assumption might be a mistake. Specifically, it might be the case that some semantic universals, in this case *Conservativity*, follow from how presupposition projection works (in conjunction with independently motivated assumptions about the syntax-semantics interface as well as other semantic universals). More specifically, this paper raises questions that I hope future work can address. Stated in broadest possible way, the question is which universals follow from which theories of presupposition projection given which auxiliary assumptions about the grammar? I hope the work reported here provides some motivation for future work in this direction.

A Proof

(51) **Lemma (17).** For any Q and A, if $\Sigma(QA)$ exists then,

$$\mathbf{x} \in \Sigma(QA) \Leftrightarrow \exists B : Q(A)(B) \neq Q(A)(B \cup \{\mathbf{x}\})$$

Proof. First we show that if $\mathbf{x} \notin \Sigma(QA)$ then $\forall B : Q(A)(B) = Q(A)(B \cup \{\mathbf{x}\})$. Take an arbitrary $B \subseteq E$. $Q(A)(B \cup \{\mathbf{x}\}) = Q(A)(\Sigma(QA) \cap (B \cup \{\mathbf{x}\})) = Q(A)((\Sigma(QA) \cap B) \cup (\Sigma(QA) \cap \{\mathbf{x}\})) = Q(A)(\Sigma(QA) \cap B) = Q(A)(B)$. Next we show that if $\mathbf{x} \in \Sigma(QA)$ then $\exists B : Q(A)(B) \neq Q(A)(B \cup \{\mathbf{x}\})$. Let $X = \Sigma(QA) \setminus \{\mathbf{x}\}$. We show that if for any B, $Q(A)(B) = Q(A)(B \cup \{\mathbf{x}\})$ then $Q(A)$ lives on X, which contradicts the assumption that $\Sigma(QA)$ is the *smallest* set that $Q(A)$ lives on. Take an arbitrary B. If $\mathbf{x} \notin B$ then $X \cap B = \Sigma(QA) \cap B$. Therefore $Q(A)(B) = Q(A)(\Sigma(QA) \cap B) = Q(A)(X \cap B)$. If $\mathbf{x} \in B$, then $B = C \cup \{\mathbf{x}\}$ where $\mathbf{x} \notin C$. $Q(A)(X \cap B) = Q(A)(X \cap (C \cup \{\mathbf{x}\})) = Q(A)((X \cap C) \cup (X \cap \{\mathbf{x}\})) = Q(A)(X \cap C) = Q(A)(\Sigma(QA) \cap C) = Q(A)(C) = Q(A)(C \cup \{\mathbf{x}\}) = Q(A)(B)$.

(52) **Lemma (26).** If E is the domain of discourse and Q is a well-behaved, permutation invariant quantifier on E, then for any $A \subseteq E$,

a. If $Q(A)$ lives on A then $\Sigma(QA)$ is either $\emptyset$ or A,

b. If $Q(A)$ does not live on A then $\Sigma(QA)$ is either E or $E \setminus A$.

Proof for (52-a): Suppose $Q(A)$ lives on A and $\Sigma(QA)$ is not empty. Let $\mathbf{x} \in \Sigma(QA)$. Since $Q(A)$ lives on A, $\Sigma(QA) \subseteq A$ by definition. We now show that $A \subseteq \Sigma(QA)$. Suppose, for contradiction, $\mathbf{y} \in A \setminus \Sigma(QA)$. Since $\mathbf{x}$ belongs to $\Sigma(QA)$ there is $B \subseteq E$ such that, $Q(A)(B) \neq Q(A)(B \cup \{\mathbf{x}\})$. We show that this can't be if Q is permutation invariant. Let π be a permutation that maps every individual to itself except that $\pi(\mathbf{x}) = \mathbf{y}$ and $\pi(\mathbf{y}) = \mathbf{x}$. Note first that $Q(\pi A)(\pi B) = Q(A)(\pi B)$ because $\pi A = A$ (since both $\mathbf{x}$ and $\mathbf{y}$ are in A). By definition, $Q(A)(\pi B) = Q(A)(\Sigma(QA) \cap \pi B)$. Next note that $Q(\pi A)(\pi(B \cup \{\mathbf{x}\})) = Q(A)((\pi B) \cup \{\mathbf{y}\}) = Q(A)(\Sigma(QA) \cap ((\pi B) \cup \{\mathbf{y}\})) = Q(A)(\Sigma(QA) \cap \pi B)$. The last step is because $\mathbf{y} \notin \Sigma(QA)$ therefore $\Sigma(QA) \cap \{\mathbf{y}\} = \emptyset$. It follows that $Q(\pi A)(\pi B) = Q(\pi A)(\pi(B \cup \{\mathbf{x}\}))$, contradiction. Therefore if $\Sigma(QA)$ is non-empty then $\Sigma(QA) = A$. **Proof for (52-b):** Suppose $Q(A)$ does not live on A. There must be some member of $\Sigma(QA)$ that does not belong to A, as otherwise $\Sigma(A)$ would be a subset of A which would entail that $Q(A)$ lives on A. Let $\mathbf{x} \in \Sigma(QA)$ and $\mathbf{x} \notin A$. First we show that $E \setminus A \subseteq \Sigma(QA)$. Suppose, for contradiction, that $\mathbf{y} \in E \setminus A$ but $\mathbf{y} \notin \Sigma(QA)$. Let π be a permutation on E that swaps $\mathbf{x}$ and $\mathbf{y}$ but maps every other individual to itself. Since $\mathbf{x}$ and $\mathbf{y}$ do not belong to A, it follows that $\pi A = A$. Since $\mathbf{x} \in \Sigma(QA)$ it follows that there is some B such that $Q(A)(B) \neq Q(A)(B \cup \{\mathbf{x}\})$. We show that this can't be if Q is permutation invariant. $Q(A)(B \cup \{\mathbf{x}\}) = Q(\pi A)(\pi(B \cup \{\mathbf{x}\})) = Q(A)(\pi B \cup \{\mathbf{y}\}) = Q(A)(\Sigma(QA) \cap (\pi B))$. On the other hand, $Q(A)(B) = Q(\pi A)(\pi B) = Q(A)(\pi B) = Q(A)(\Sigma(QA) \cap (\pi B))$. Therefore $Q(A)(B) = Q(A)(B \cup \{\mathbf{x}\})$, contradiction. Therefore, $E \setminus A \subseteq \Sigma(QA)$. Next we show that if there is some $\mathbf{y} \in \Sigma(QA) \cap A$ then $A \subseteq \Sigma(QA)$. Suppose, for contradiction, that $\mathbf{y} \in \Sigma(QA) \cap A$ and $\mathbf{z}$ is a member of A that doesn't belong to $\Sigma(QA)$. Let π be a permutation on E that swaps $\mathbf{y}$ and $\mathbf{z}$ must maps every other individual to itself. Since $\mathbf{y}$ and $\mathbf{z}$ both belong to A, it follows that $\pi A = A$. Since $\mathbf{y} \in \Sigma(QA)$, there is a set B such that $Q(A)(B) \neq Q(A)(B \cup \{\mathbf{y}\})$. We show that this can't be if Q is permutation invariant. $Q(A)(B \cup \{\mathbf{y}\}) = Q(\pi A)(\pi(B \cup \{\mathbf{y}\})) = Q(A)(\pi B \cup \{\mathbf{z}\}) = Q(A)(\Sigma(QA) \cap (\pi B))$. On the other hand, $Q(A)(B) = Q(\pi A)(\pi B) = Q(A)(\pi B) = Q(A)(\Sigma(QA) \cap (\pi B))$. Therefore, $Q(A)(B) = Q(A)(B \cup \{\mathbf{y}\})$, contradiction. Therefore, if $\Sigma(QA)$ overlaps with A then it includes A as a whole. We have already established that $(E \setminus A) \subseteq \Sigma(QA)$. Therefore, either $\Sigma(QA) = E \setminus A$ or $\Sigma(QA) = E$.

References

Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2). 159–219. doi:10.1007/BF00350139.

Doron, Omri & Jad Wehbe. 2022. A constraint on presupposition accommodation. In *Proceedings of the 23rdA msterdam Colloquium*, https://shorturl.at/e28u8.

von Fintel, Kai. 1994. *Restrictions on quantifier domains*: University of Massachusetts, Amherst PhD Dissertation.

von Fintel, Kai & Edward L Keenan. 2018. Determiners, Conservativity, Witnesses. *Journal of Semantics* 35(1). 207–217. doi:10.1093/jos/ffx018.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33(1). 63–96. doi:10/cbq2fh.

Fox, Danny. 2003. On Logical Form. In Randall Hendrick (ed.), *Minimalist Syntax*, 82–123. Oxford, UK: Blackwell Publishing Ltd. doi:10.1002/9780470758342.ch2.

Heim, Irene. 2005. Features of Pronouns: The Semantics of Number, Gender, and Person.

Higginbotham, James & Robert May. 1981. Questions, quantifiers and crossing 1(1). 41–80. doi:10.1515/tlir.1981.1.1.41. Publisher: De Gruyter Mouton Section: The Linguistic Review.

Keenan, Edward L. & Jonathan Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9(3). 253–326. doi:10.1007/BF00630273.

Lebeaux, David. 1988. *Language acquisition and the form of grammar*: University of Massachusetts, Amherst dissertation.

Ludlow, Peter. 2002. LF and Natural Logic. In Georg Peter & Gerhard Preyer (eds.), *Logical Form and Language*, 132–168. Oxford University Press.

Pasternak, Robert & Uli Sauerland. 2022. German measurement structures: case-marking and non-conservativity. *The Journal of Comparative Germanic Linguistics* 25(2). 221–272. doi:10.1007/s10828-022-09134-y.

Romero, Maribel. 2021. The many readings of many: POS in the reverse proportional reading. *Linguistics and Philosophy* 44(2). 281–321. doi:10.1007/s10988-019-09288-1.

Romoli, Jacopo. 2015. A structural account of Conservativity. *Semantics - Syntax Interface* 2(1). 28–57.

Schlenker, Philippe. 2007. Expressive presuppositions. *Theoretical Linguistics* 33(2). doi:10.1515/TL.2007.017.

Schlenker, Philippe. 2009. Local Contexts. *Semantics and Pragmatics* 2. 1–78. doi:10/bwt.

Sportiche, Dominique. 2005. Division of labor between Merge and Move: Strict locality of selection and apparent reconstruction paradoxes.

Sudo, Yasutada. 2012. *On the Semantics of Phi Features on Pronouns*: Massachusetts Institute of Technology PhD Dissertation. http://hdl.handle.net/1721.1/77805.

Takahashi, Shoichi. 2010a. Traces or Copies, or Both – Part I: Characterizing Movement Properties. *Language and Linguistics Compass* 4(11). 1091–1103. doi:10.1111/j.1749-818X.2010.00255.x.

Takahashi, Shoichi. 2010b. Traces or Copies, or Both – Part II: Introducing Further Considerations. *Language and Linguistics Compass* 4(11). 1104–1115. doi:10.1111/j.1749-818X.2010.00256.x.

Takahashi, Shoichi & Sarah Hulsey. 2009. Wholesale Late Merger: Beyond the A/Ā Distinction. *Linguistic Inquiry* 40(3). 387–426. https://www.jstor.org/stable/40284323.

Zuber, Richard & Edward L Keenan. 2019. A Note on Conservativity. *Journal of Semantics* 36(4). 573–582. doi:10.1093/jos/ffz007.