Predicate nominals and distributivity

September 26, 2017

Nicholas Longenbaugh[1]

1 Introduction

It is widely held that natural language makes available a covert alternative to the adverbial distributive operator *each*, which I will refer to here as *Dist*. On an informal level, *Dist* combines with a predicate and a plural DP and asserts that each atomic element in the extension of the DP satisfies the predicate. A canonical example involving this operator is provided in (1), which, on the most plausible reading, asserts that each of several men wore a different tie. In this case, *Dist* combines with the plural DP *several men* and the associated predicate *wear a bolo tie*.

(1) a. Several men wore a bolo tie.
b. [[several men] [*Dist* [wore a bolo tie]]]
c. ≈ several men each wore a different bolo tie.

Assuming this basic treatment of the facts, (1) teaches us that *Dist* can felicitously combine with a plural subject and its associated predicate. Against this backdrop, Dotlačil (2011) observes the following puzzle: *Dist* is apparently unable to combine with plural subjects of predicate nominals (PNs). If we grant that the singular PN *a doctor* denotes the set of all atomic individuals who are doctors, applying Dist to *several women* and *are a doctor* should license license the semantically sensible reading that each of several women is a doctor. This reading is, however, unavailable, so (2) is anomalous.

(2) a. #Several women are a doctor.
b. *[[several women] [*Dist* [are [a doctor]]]]

This is a robust phenomenon: it extends to small clauses (see (3c)), to predicative definites (see (4)), and to overt distributive operators (see (5)), although Dotlačil (2011) does not discuss these cases.[2,3]

(3) a. #Several women I know are a doctor.
b. #My friends are a really good teacher.
c. #I consider all my students a good listener.
(Dotlačil 2011)

(4) a. #These men are the best candidate in their districts.
b. #I consider these books the longest volume on some topic.

(5) a. #These buildings (each/both/all) are (each/both/all) a warehouse.
b. #These men (each/both/all) are (each/both/all) a good teacher.

Dotlačil considers and rejects two solutions to this problem. The first is that the infelicity displayed above reflects a syntactic agreement mismatch between the PN and its subject, and is independent of distributivity. The problem with this approach is twofold. First, English PNs

[1]I would like to thank Roger Schwarzschild, Danny Fox, Irene Heim, David Pesetsky, Sabine Iatridou, Daniel Margulis, Itai Bassi for helpful discussion and insights on this project.

[2]Although Dotlačil (2011: 330) suggests in passing that floated *each* should distribute over PNs, counter to fact.

[3]The determiner use of *each* does license distributivity here. I return to this issue in Section 3.3.4.

(i) Each linguist here is also a doctor.

need not agree in number with their subjects. Both collective nouns and pluralia tantum, which are semantically singular but syntactically plural, can serve as the subjects of singular PNs.

(6) a. These students are a great team.
b. These scissors are a great tool.

Second, in languages with gender agreement morphology, PNs need not show syntactic agreement for gender, casting doubt on the involvement of syntactic agreement in these cases.

(7) On byl bab-a.
he.SG.M was.SG.M witch.SG.**F**
"He was a coward." (*Czech*; Dotlačil 2011: 317)

The second rejected analysis is to simply say that *Dist* cannot apply to singular PNs. While this may or may not be true, Dotlačil points out that it is not the only source of the anomaly in the cases reviewed above. In particular, if the plural subject of a singular PN undergoes subject-to-subject raising, *Dist* likewise cannot apply to the derived predicate created by raising the subject. This is despite the fact that *Dist* can ordinarily apply to the derived predicate formed by raising a plural DP (see (9)).

(8) a. #These students seemed to be a genius.
b. [These students [*Dist* [1 [seemed [t_1 to be a genius]]]]]

(9) a. These students seemed to at least one professor to be smart.
b. [These students [*Dist* [1 [seemed to at least one professor [t_1 to be smart]]]]]

We can summarize the core puzzle as in (10).

(10) **The anti-distributivity of PNs**:
The plural subject of a PN may not be distributed over that PN.

To account for this behavior, Dotlačil argues that the semantics of the distributive operator must be rethought: rather than distributing over the scope of the entire predicate it combines with, he proposes that *Dist* is "fastidious" and selectively distributes over arguments in its scope. The canonical example in (1) then gets an LF as below, where *Dist* selectively distributes over only the co-indexed indefinite *a bolo tie*.

(11) a. Several men wore a bolo tie.
b. [[several men] [$Dist_x$ [wore [a bolo tie]$_x$]]]

The impossibility of distributing over PNs, Dotlacil proposes, then follows if we assume that distributivity is only possible over *argument* DPs.

(12) *Dist* may be coindexed only with DPs in argument position. (Dotlačil 2011: 321)

This notion of fastidious distributivity, augmented with the argument-only proviso in (12), captures the facts so far discussed.

That said, there are reasons to suspect this approach may not be on the right track. To begin, while (12) is an adequate description of the phenomenon, as Dotlačil (2011) himself acknowledges, it is unclear why such a principle should exist, or how to independently derive it. More pressingly, however, there appears to be two related classes of counter-examples to (12), that is, cases where a plural DP distributes over a PN.

The first case involves plural subjects of PNs that appear as the head of a curious type of agreement-mismatched relative clause, as in (13) (see Longenbaugh 2017 for arguments that the bracketing is as indicated). In this case, the distributive reading, where, e.g., each of few linguists is also a doctor, is readily licensed.

(13) a. John is one of [the few [linguists [who is also a doctor]]].
b. Sally is one of [many [mathematicians [who is a genius]]].

The second case involves plural DPs that bind the pronominal subject of a PN.[4] The distributive reading is again licensed in such cases.

(14) a. ?Several of the medical students think they're already a doctor.
b. ?Many the graduate students act like they're already a professor.
c. These men all think they are they only one in the room. (Heim 2008)
d. These candidates each think they are a/the winner.

It would seem that these data warrant a new approach to Dotlačil's puzzle.

To this end, I argue that the exceptional cases share in common a core property to the exclusion of those examples that do exhibit anti-distributivity, suggesting a possible first step towards an eventual analysis. Specifically, in those cases where distributivity is licensed, the syntactic subject of the PN at LF is not semantically plural. In the former case, Longenbaugh (2017) argues that the agreement mismatched relative clauses, roughly speaking, involve a movement chain where the higher and lower copies differ in terms of their number feature: the lower copy contains a syntactically active and semantically interpretable singular feature, in contrast to the plural feature in the higher copy.

(15) a. John is one of [the few [linguists [who is a doctor]]].
b. [the few [linguists.PL_1 [*Dist* [$_{\text{RC}}$ [~~linguist~~.SG_1 is a doctor]]]]]

Likewise, in the latter case, if we follow Kratzer (1998), Schlenker (2003), von Stechow (2003), Heim (2008) in assuming that the plural feature on bound pronouns need not be interpreted, the bound pronominal subject of the PN is syntactically but not semantically plural.

(16) a. These men each think they are a/the winner.
b. [[these men].PL_1 [*Dist* [think [they.$\emptyset_1$ are a genius]]]]

In the canonical cases of anti-distributivity, in contrast, the underlying subject of the PN is either the semantically plural DP itself or the plural lower copy left behind by moving this DP.

(17) a. #These students seem to be/are a genius.
b. *[these students].PL_1 [*Dist* [seem to be/are [[~~these students~~].PL_1 [a genius]]]]

It therefore appears that the relevant factor in licensing distributivity may be whether the subject of the PN at LF, be it the lower copy in a movement chain or a bound pronoun, is semantically

[4] For some speakers, these examples are most felicitous in the presence of an overt distributive operator (cf. **??**, **??**). On the analysis I develop here, this can be attributed to the fact that third person plural pronouns most readily tolerate readings where they range over atomic individuals when they are in the scope of an overt distributive operator. The contrast, in other words, extends beyond the specific case of predicate nominals.

(i) a. These are the students who think they're tall. (? *distributive*)
b. These candidates think they're highly capable. (? *distributive*)

singular. If so, distributivity is possible. If not, it isn't.

In the remainder of this article, I develop this observation into a full fledged account of the anti-distributivity effect. I begin, in Section 2, by arguing that under the traditional view of plural number, applying *Dist* to the predicate derived by moving a plural DP results in a presupposition failure if the lower copy is plural.

(18) #[DP.PL$_1$ [*Dist* [...~~DP~~.PL$_1$ ...]]] (*presupposition failure*)

In order for movement to feed distributivity, it is therefore necessary for the lower copy not to contain a plural number feature. I argue that this can be accomplished via Takahashi and Hulsey's (2009) mechanism of *wholesale late merge* (WLM), which allows for reduced lower copies lacking a number feature.

(19) [DP.PL$_1$ [*Dist* [...$\overbrace{\text{DP.}\emptyset_1}^{\text{WLM}}$ ...]]] (✓ *distributivity*)

In Section 3, I then argue that the syntax of PNs independently blocks WLM of the subject, deriving anti-distributivity. If the lower copy/variable is non-plural, no WLM is needed and distributivity is possible, capturing the exceptional cases.

In Section 4, I extend the analysis, arguing that linking distributive readings of moved DPs to WLM accounts for the well-known difficulty of inverse-distributive readings in examples like (20). This follows from the independently needed proposal that WLM is difficult to impossible in A′-chains.

(20) A speech inspired these students. (?? *distributive*)

Finally, Section 5 reviews remaining issues and possible extensions.

2 Proposal

2.1 Background assumptions

2.1.1 The syntax and semantics of number

To begin, I adopt an ontology that associates each model with a domain of individuals, D_e, a partial order on the domain, $\sqsubseteq$, and a structure such that for any $x, y \in D_e$, $x \neq y$, there is a unique individual $z = x \oplus y$, the mereological sum of x and y, such that $x, y \sqsubseteq z$ and for all w, $x, y \sqsubseteq w \rightarrow z \sqsubseteq w$. An individual $x \in D_e$ is said to be *atomic* if and only if for all y, $y \sqsubseteq x$ entails that $y = x$. Those individuals that are not atomic will be said to be *pluralities*. Let ATOM be defined as the function on individuals that returns true if and only if its argument is atomic and SUM be the function on individuals that returns true if and only if its argument is a plurality.

I will further assume, following Krifka (1992), Landman (1996), Kratzer (2008), a.o., that lexical predicates in natural language are cumulative – if a, b satisfy the predicate, then so does $a \oplus b$ – and hence defined for both atoms and pluralities (number neutral). Following standard practice, I express the denotation of a lexical predicate in terms of the $*$-operator, which closes a set under sum formation, applied to the primitive intuitive meaning of the corresponding predicate, which I express in small caps.[5]

(21) $[\![\text{boy}]\!] = \lambda x.\ *\text{BOY}(x)$

[5] More needs to be said about predicates with arity greater than two. I set aside such cases here.

Given that lexical nouns do not themselves distinguish number, I assume that this task falls to the number features, SG, PL. For now, let's grant that there is a dedicated head, Num, that lies between D and NP (Ritter 1991; Zamparelli 1995; Nelson and Toivonen 2000; Longobardi 2001; Matushansky 2006; Watanabe 2006; Danon 2011), and that serves as the syntactic and semantic locus of number features (I refine my assumptions concerning DP structure in Section 3.1). I will then adopt the semantics for number features along the lines in (22), where both singular and plural function to restrict the predicate they combine with the be defined only over atoms or pluralities, respectively.[6] These features can then combine via predicate modification (Heim and Kratzer 1998) with the number neutral NP to give a singular or plural NP.[7]

(22) a. $[\![\text{SG}]\!] = \lambda x : \text{ATOM}(x).\ x = x$
b. $[\![\text{PL}]\!] = \lambda x : \text{SUM}(x).\ x = x$
c. $[\![\text{NUM NP}]\!] = \lambda x : [\![\text{NUM}]\!](x).\ [\![\text{NP}]\!](x)$

Finally, I will assume the presence of a covert adverbial distributive operator, *Dist*, which can adjoin at least to TP (we will refine the syntactic behavior of *Dist* in later sections) and which distributes indiscriminately over its scope.

(23) $[\![\text{Dist}]\!] = \lambda P.\lambda X.\ \forall x[(\text{ATOM}(x) \wedge x \sqsubseteq X) \rightarrow P(x)]$

This operator allows us to capture the distributive reading of, e.g., (1) along the lines below.

(24) a. Several men wore a bolo tie.
b. LF: [[several [PL men]] [*Dist* [1 [[a [SG [bolo tie]]] [2 [t_1 wore t_2]]]]]]
c. $\exists X[|X| = \text{SEVERAL} \wedge *\text{MAN}(X) \wedge \forall x[\text{ATOM}(x) \wedge x \sqsubseteq X \rightarrow \exists y[\text{ATOM}(y) \wedge *\text{TIE}(y) \wedge *\text{WEAR}(y)(x)]]]$

My assumptions concerning number can be summarized as in (25).

(25) **Assumptions concerning semantic number and plurality**
a. Lexical predicates (N, V, Adj, etc.) are number-neutral and cumulative
b. There is a dedicated head in the DP, Num, that is the syntactic and semantic locus of number features, and that sits between D and NP.
c. The singular and plural morphemes restrict the NP they combine with to be true only of atoms and pluralities, respectively.
d. There is a covert adverbial distributive operator.

2.1.2 Movement and Late Merge

The account will also rely on the copy theory of movement (Chomsky 1995; Fox 2002; a.o.), augmented with Fox's (2017) trace conversion algorithm and Takahashi and Hulsey's (2009) (T&H) theory of WLM.

[6] The presuppositionality of these features is not crucial to the account of (10), which carries through even if the atomicity/plurality requirement is part of the assertional content of the features. As we will see in Section 3.3.3 my primary motivation for making these features presuppositional is to account for the infelicity of sentences like (i).

(i) Mary is doctors.

[7] There is significant controversy in the literature over the proper treatment of plural number (Chierchia 1998; Sauerland et al. 2005; Spector 2007; de Swart and Farkas 2010; Grimm 2012; Kiparsky and Tonhauser 2012; Marti 2017; a.o.). My assumptions concerning plural, in particular, go against the claims of many of these authors. I set aside this issue until Section 5

Syntactically speaking, the copy theory holds that movement involves re-merging the target XP in the structure, resulting in two identical copies at the head and foot of the chain. I then assume that movement chains are interpreted according to Fox's (2017) rule of trace conversion, with the result that the lower copy is interpreted as a restricted variable.

(26) **Trace Conversion**:
If YP=Merge(ZP, XP), with XP is contained in ZP, then, $[\![\text{YP}]\!]^g = [\![\text{XP}]\!](\lambda z.[\![\text{ZP}(i/\text{X}^0)]\!]^{g[i/z]}$, with i a numerical index that does not appear in ZP and ZP(i/X^0) a modification of ZP differing only in that the highest position of the head of XP in ZP bears the index i

In the case of DP movement, this amounts to interpreting the sister of a moved phrase with the determiner in the lower copy of DP replaced with the indexed version of that determiner. Following Fox (2017), Elbourne (2005), I assume indexed determiners are interpreted as in (27).

(27) $[\![\text{D}_i]\!]^g = \lambda P : P(g(i)) = 1.\ g(1)$

For an explicit example, (28a) has the LF in (28b), granting that the subject is generated vP internally and undergoes EPP-movement to Spec(TP). The interpretation of this LF is given in (28c), following the rule of TC above. Crucially, the presupposition introduced by the number feature in the lower copy projects to the predicate that forms the sister of the moved DP.

(28) a. Several boys laughed.
b. LF: [$_{\text{TP}}$ [several [PL boy]] [$_{v\text{P}}$ [several [PL boy]] laughed]]
c. $[\![\text{TP}]\!] = [\![\text{[several [PL boy]]}]\!](\lambda z.\ [\![\text{[[several}_1\ \text{[PL boy]] laughed]}]\!]^{[1/z]})$
$= [\![\text{[several [PL boy]]}]\!](\lambda z : \text{SUM}(z) \wedge *\text{BOY}(z).\ *\text{LAUGHED}(z))$
$= \exists X[|X| = \text{SEVERAL} \wedge *\text{BOY}(z) \wedge *\text{LAUGHED}(z)]$

TC as conceived of here does not formally alter the lower copy in any way, nor does it depend on movement introducing a binding index. That said, in the remainder of this paper, I will represent the LF of movement examples with an indexed determiner in the lower copy and an explicit binding index, for the sake of readability.

With the copy theory in place, I turn my attention now to T&H's theory of WLM. One of the main empirical arguments for the copy theory (Romero 1998; Fox 1999, 2002; Sauerland 1998, 2004; Sportiche 2005 for additional empirical arguments) is that it straightforwardly captures the fact that A′-movement tends not to bleed Condition C. In (29a), the R-expression *John* in *the claim that John is guilty* is c-commanded by the co-referential pronoun *him*, causing a Condition C violation. As (29b) shows, A′-movement of the DP containing *John* leads to only a marginal improvement. This follows straightforwardly if the higher and lower copies are syntactically identical, as the presence of *John* in the lower copy will induce a Condition C violation.[8]

(29) a. *He$_1$ bought several pictures of John$_1$.
b. ??Which pictures of John$_1$ did he$_1$ buy?
c. [[which picutres of John$_1$] [2 [he$_1$ bought [which$_2$ pictures of John$_1$]]]]
✗ Condition C

That said, there are a variety of well-known instances where movement *does* bleed Condition C. The case relevant here is subject-to-subject raising: it is possible to raise a DP containing an R-expression over an experiencer PP containing a co-referential pronoun (Sauerland 1998;

[8]Something else must be said to account for the fact that for many speakers, at least, (28b) is not as degraded as (28a). I return to this question below.

Lebeaux 1998; Fox 1999; Takahashi and Hulsey 2009). If the lower copy in (30b) is a full copy, as in (30c), we incorrectly predict that the sentence should induce a Condition C violation. The simple copy-theory therefore under-generates the set of acceptable structures.

(30) a. *It seems to him_1 that those pictures of John_1 were stolen.
b. Those pictures of John_1 seem to him t to have been stolen.
c. [[those pictures of John_1] [2 [seem to $\underbrace{\text{him}_1 \ldots [\text{those}_2 \text{ pictures of } \text{John}_1]}_{\text{✗ Condition C}} \ldots]]]]

T&H propose a solution to this problem based on Lebeaux's (1988) mechanism of *late merge*.[9] The explicit proposal is that in A-movement chains, the NP, or in our case NumP, component of the DP can be *late merged* in the higher copy of the chain. The lower copy is therefore merged into the structure as a bare D head, which is then copied and moved according to the basic copy theory sketched above. After movement has taken place, NumP is then merged to the higher copy only. T&H refer to this as *wholesale late merge* (WLM). The resulting LF for an example like (29) is then as in (31).

(31) [[those $\overbrace{\text{[pictures of John}_1\text{]}}^{\text{WLM}}$]] [2 [seem to him_1 [to have [those_2 $\emptyset$] been stolen]]]]

To interpret (31), T&H assume that the trace conversion procedure interprets bare determiners as unrestricted variables. Given our assumptions, this amounts to the proposal that indexed determiners also have the option of being interpreted as unrestricted variables, as below. I will therefore take indexed determiners to be ambiguous along these lines, with the choice among interpretations being free up to interpretability.

(32) $[\![\text{D}_1]\!]^g = g(1)$

Finally, to account for the standardly assumed differences between A- and A′-movement, T&H propose that WLM is only available in A-chains. They derive this result as a corollary of case theory, assuming (i) that NP (or NumP) has a case feature that must checked in the course of the derivation, and (ii) that case cannot be checked in A′-positions (Chomsky 1981; 2000; but see also Kayne 1989; Rezac 2013; van Urk 2015). This rules out WLM of NP to the top of an A′-chain. I will mostly be concerned with A-movement in the next two sections, so I set aside this aspect of of T&H's proposal, and its present consequences, until Section 4.

My assumptions concerning movement are summarized below.

(33) **Assumptions concerning movement**:
a. Movement proceeds according to the copy theory, and is interpreted via Fox's (2017) version of TC, which does not formally alter the lower copy.
b. There is a process of WLM, whereby the sister to D in a given DP can be merged counter-cyclically in the higher copy, after movement has taken place
c. WLM is licensed in A-chains.

2.2 Proposal

With these assumptions in place, my proposal is that distributing over the derived predicate created by moving a plural DP is contingent on WLM.

[9] Lebeaux (1988) is primarily concerned with certain asymmetries in the bleeding of Condition C under A′-movement which need not concern us here.

(34) **No distributing over plural traces**:
Given a movement chain involving a plural DP, [D NumP], the derived predicate created by movement may be distributed over if and only if NumP has been late merged in the higher copy.

This follows as a direct consequence of the framework sketched in the previous two subsections. I illustrate via the following example, where a plural DP undergoes A-movement from its *v*P-internal basic merge site to Spec(TP). Suppose that we elect not to WLM the complement of D, and moreover that we insert a distributive operator over the predicate derived by movement. The resulting LF is given below.

(35) a. Several girls laughed.
b. [[several [PL girls]] [*Dist* [1 [... [[several$_1$ [PL girls]] laughed]]]]]

Following the conventions above, the predicate derived via movement is then interpreted as the function that holds of all individuals who laughed, with a domain restricted to those individuals in the extension of the plural NP *girls*.

(36) $[\![[1\ [\ldots[[\text{several}_1\ [\text{PL girls}]]\ \text{laughed}]]]]\!] = \lambda X : \text{SUM}(X) \wedge *\text{GIRL}(X).\ *\text{LAUGH}(X)$

The crucial observation is that this function is undefined on atoms. Recalling the denotation we are assuming for *Dist*, repeated in (37a), it is possible to view this operator as a lexically restricted universal quantifier whose sister functions as its nuclear scope (see (37b)).

(37) $[\![\text{Dist}]\!] = \lambda P.\lambda X.\forall x[(\text{ATOM}(x) \wedge x \sqsubseteq X) \rightarrow P(x)]$
$= \lambda P.\lambda X.\ [\![\text{every}]\!](\lambda x.\ \text{ATOM}(x) \wedge x \sqsubseteq X)(P)$

Under most theory of presupposition projection (Karttunen and Peters 1979; Heim 1983; Schlenker 2007), universal quantifiers presuppose that every element of their restriction satisfies the presuppositions of its scope. In the case at hand, the restriction denotes a subset of all atomic individuals (see below), so if the nuclear scope is undefined for atoms, applying Dist to this predicate yields a presupposition failure.

(38) $[\![[\text{Dist}\ [1\ [\ldots]]]]\!] = \lambda X.\ [\![\text{every}]\!](\lambda x.\ \text{ATOM}(x) \wedge x \sqsubseteq X)(\lambda X : \text{SUM}(X) \wedge *\text{GIRL}(X).\ *\text{LAUGH}(X))$

This result is fully general: as long as the presupposition introduced by the lower copy, namely that the value the assignment function maps the abstraction index to must be a plural individual, projects to the level of the distributive operator, a presupposition failure will always result.

WLM offers a way to avoid this presupposition failure. Assuming that the Num head is contained in the sister of D that can be late merged, the lower copy in a WLM derivation need not contain any number information, which removes the problematic presupposition and allows the predicate derived via movement to combine with the Dist operator. In a derivation with WLM, the example from above gets the LF in (39a), and the predicate derived via movement is interpreted as in (39b). This can then felicitously combine with the distributive operator, as in (39c).

(39) a. [[several [PL students]] [Dist [1 [... [[several$_1$ $\emptyset$] laughed]]]]]
b. $[\![[1\ [\ldots[[\text{the}_1\ \emptyset]\ \text{laughed}]]]\!] = \lambda x.\ *\text{LAUGHED}(x)$
c. $[\![[\text{Dist}\ [1\ [\ldots]]]]\!] = \lambda X.\forall x[(\text{ATOM}(x) \wedge x \sqsubseteq X) \rightarrow *\text{LAUGHED}(x)]$

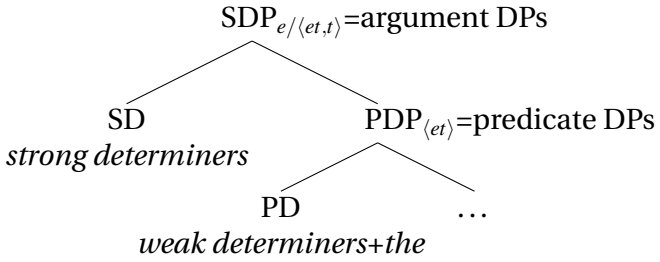
3 Predicate nominals

In this Section, I return to the core topic of this paper, the impossibility of distributing a plural DP subject of a PN (PN) over the corresponding PN. Building on (39), my proposal here is that WLM is blocked for plural subjects of PNs by independent syntactic factors, deriving anti-distributivity. I begin with some requisite syntactic assumptions concerning the structure of DP and of PNs. I then propose the analysis and explore several predictions, show specifically how the approach captures the possibility for distributivity over PNs in the exceptional cases in the introduction.

3.1 Syntactic Assumptions

I begin with some prerequisite assumptions about the syntax and semantics of predicate and argument nominals. I adopt the view, which goes back to at least Stowell (1981), that argument nominals are "built" by adding extra structure to predicate nominals.[10] This is a common position in the semantics literature, but the specific forumulation that I adopt is due to Zamparelli (1995), with modifications as proposed by Coppock and Beaver (2015). Zamparelli's core proposal is that the traditional DP-layer can split into two categories, the *strong determiner phrase* (SDP) and the *predicational determiner phrase* (PDP), where SD^0 takes PDP as its complement (see (40)). On this view, PDP is the highest projection in the nominal system that has predicative type, $\langle et \rangle$, and the SD head then shifts this into an argument meaning, type e or $\langle et,t \rangle$. Predicate and argument nominals differ, on this view, only in that the former lacks the SDP layer, so that its subject is merged in Spec(PDP).

(40) **DP structure**:



The category SD in this system comprises those determiners that cannot be used in PNs, which roughly overlaps with Milsark's (1974) class of "strong determiners." The definite article, which has predicative uses (see (41c,d); Zamparelli 1995: 5.3; Fara 2001; Winter 2001; Coppock and Beaver 2015: 2.1), is a notable exception.[11]

(41) a. *John, Bill, and Ted are every boy.
b. *Sue and Sally are most doctors.
c. Bill considers Sue the best doctor in Walla Walla. (Coppock and Beaver 2015: 381)

[10] I eschew, in particular, the view that there is a special "PredP" projection that transforms an argument nominal into a PN (Bowers 1993, 2001), which in addition to being semantically unwarranted has recently been argued independently to be syntactically suspect as well (Matushansky 2015).

[11] Other arguments that definites can be used predicatively are that they: (i) can be conjoined with adjectives and other clearly predicative types, unlike proper names; (ii) can participate in so-called appositional conjunction, unlike proper names. These are illustrated consecutively below.

(i) a. John is tall, handsome, and the love of my life.
b. A great man and the best magician in New Jersey has passed away. (Coppock and Beaver 2015)

The category PD comprises those determiners that *can* be used in PNs, including the definite article, the indefinite article, cardinal numerals, and other "weak" determiners in Milsark's sense.

Given this basic structure, elements of category PD can naturally be thought of as functions of type $\langle et, et \rangle$. For most weak determiners, this matches common assumptions (see, e.g., Ionin and Matushansky 2006 on cardinal numerals, Hackl 2000 on *many*, and so on). I discuss these cases in much more detail in Section 3.4, but for now the example denotations below suffice.

(42) a. $[\![\text{several}]\!] = \lambda P.\lambda x.\ |x| = \text{SEVERAL} \wedge P(x)$
b. $[\![\text{many}]\!] = \lambda P.\lambda x.\ |x| = \text{MANY} \wedge P(x)$
c. $[\![\text{three}]\!] = \lambda P.\lambda x.\ |x| = 3 \wedge P(x)$

As the definite and indefinite article also both allow for predicate uses, I will also assume they are underlying type $\langle et, et \rangle$, as below (see especially Coppock and Beaver 2015 for independent arguments in favor of this treatment).[12]

(43) a. $[\![\text{a}]\!] = \lambda P.\ P$
(Heim and Kratzer 1998: 62; Winter 2001: 146; Coppock and Beaver 2015: 401)
b. $[\![\text{the}]\!] = \lambda P : \exists! x[P(x)].\ P$
(Fara 2001; Winter 2001; Coppock and Beaver 2015)

A basic PN involving, e.g., *the*, then has the structure in (44), where the surface subject is merged in Spec(PDP), then raised to Spec(TP)

(44) a. Mary is the hero of my youth.
b. [$_{\text{TP}}$ Mary is [$_{\text{PDP}}$ ~~Mary~~ [$_{\text{PDP}}$ the [hero of my youth]]]]

This treatment of weak determiners raises the obvious question of how the DPs they head come to be used in argument position. Here I follow the basic proposal of Coppock and Beaver (2015) and assume that the class of SD elements included two silent determiners, IOTA and EX, modelled after Partee's (1987) type-shifters of the same name. The argument use of weak DPs then involves a structure where SD is occupied by one of these silent determiners. The traditional readings of DPs headed by *the* and *a* then arise in the configurations below.[13,14]

[12]Extending this denotation for *the* to capture definite plural predicative NPs is slightly more complicated. One option is to follow Winter (2001) in positing the obligatory presence of the MAX_SORT operator, defined as below, between *the* and Num. With this, we can maintain the denotation for *the* above and IOTA below. There are other possibilities as well. I set this complexity aside here, as it is orthogonal to the main argument.

(i) $\text{MAX\_SORT} = \lambda P_{et}.\lambda X.\ P(X) \wedge \forall Y[X \sqsubset Y \rightarrow \neg P(y)]$

[13]If we say nothing else, this system predicts that IOTA and EX can combine with any PDP. We therefore predict readings of DPs headed by *a* with uniqueness presuppositions, and conversely readings of DPs headed by *the* without them. Coppock and Beaver (2015) argue that the latter but not the former are attested, so that we only need to block IOTA from combining with a PDP headed by *a*. They propose to achieve this with a modified version of Maximize Presupposition along the lines of Percus (2007). More generally, the present system correct predicts that overt elements of SD and PD can be combined in a single DP in many cases, as in (i). There are also many unattested combinations, which must be ruled out on independent grounds (see Zamparelli: 1995: 4.1.4 and Coppock and Beaver 2015: 3.3 for discussion and explicit proposals to do just this).

(i) a. Every three girls
b. At most four students

[14]This has non-trivial but generally unproblematic consequences for the interpretation of the lower copy under TC, at least for weak DPs. Given the TC rule above, and given that we are assuming weak DPs are headed by an overt

(45) a. $[\![\text{IOTA}]\!] = \lambda P_{et} : \exists!x[P(x)].\ \iota x[P(x)]$
b. $[\![\text{EX}]\!] = \lambda P_{et}.\lambda Q_{et}.\ \exists x[P(x) \wedge Q(x)]$

(46) a. the boy
b. LF: [$_{\text{SDP}}$ IOTA [$_{\text{PDP}}$ the [boy]]]
c. $[\![\text{the boys}]\!] = \iota x[[\![\text{boy}]\!](x)]$ if there is exactly one boy, undefined otherwise[15]

(47) a. a boy
b. LF: [$_{\text{SDP}}$ EX [$_{\text{PDP}}$ a [boy]]]
c. $[\![\text{a boys}]\!] = \lambda Q.\exists x[[\![\text{boy}]\!](x) \wedge Q(x)]$

A final issue that will be relevant for our purposes is the position of number within the DP structure articulated in (40). It has often been argued that the basic merge sight of cardinal numerals is either at or adjoined to the head Num hosting the syntactic and semantic realization of number within the DP (Danon 2011; Ionin and Matushansky 2006; Longobardi 2001; Nelson and Toivonen 2000; Ritter 1991). This leads us to identify PD with Num, a position that Zamparelli explicitly embraces in later work (Heycock and Zamparelli 2005: fn.17). I will therefore assume that number features enter the derivation in the PD-layer.[16]

3.2 No WLM for subjects of PNs

With the requisite syntactic background now in place, I offer in this section my explicit proposal that WLM is ruled out for the subject of PNs, deriving anti-distributivity.

To begin, consider the canonical anti-distributivity example in (48). In the absence of a distributive operator, we predict this utterance straightforwardly to be a presupposition failure: the singular PN is defined only over atoms, and the plural subject denotes a plurality.

(48) a. Several women are a doctor.
b. $[\![\text{a doctor}]\!] = \lambda x : \text{ATOM}(x).\ *\text{DOCTOR}(x)$
c. $[\![\text{several women}]\!] = \lambda Q.\ \exists X[|X| = \text{SEVERAL} \wedge *\text{WOMAN}(X) \wedge Q(X)]$
d. $\exists X[|X| = \text{SEVERAL} \wedge *\text{WOMAN}(X) \wedge (\lambda x : \text{ATOM}(x).*\text{DOCTOR}(x))(X)]$ (*undefined*)

Suppose now that we insert a distributive operator over the derived predicate created by moving the subject DP. The only way this structure will be interpretable is if we WLM the PDP comple-

PD combined with a covert SD (EX or IOTA), if the DP is headed by a cardinal numeral or other weak quantifier, the denotation of that quantifier will now be included in the presupposition introduced on the index by the indexed quantifier (see (i)). This will not affect any of the results of the ensuing section, but it is worth pointing out.

(i) a. [1 [[EX$_1$ [several pl boys]] laughed]]
b. $\lambda x : |x| = \text{SEVERAL} \wedge *\text{BOY}(x).*\text{LAUGH}(x)$

[15]There is a redundancy here, as the PD head *the* introduces uniqueness and existence presuppositions that are duplicated by the IOTA operator. For Coppock and Beaver (2015), this problem does not arise, as they asume that the definite PD head *the* introduces only a uniqueness, not an existence, presupposition. Our conclusions are compatible with their assumptions, but I have opted to use the more familiar (and redundant) definitions above for ease of presentation.

[16]Depending on our assumptions, this may necessitate recursive PD-projections, to host number in the presence of other overt elements of category PD. This is probably independently needed given that overt elements of PD can co-occur, as in (i).

(i) a. The two boys
b. Several hundred papers
c. Twenty five books
d. A few pictures

ment to SD in the higher copy, as in (49a). Focusing on the lower copy, the general structure of the PN is cases such as this is as in (49b). Setting aside the label of the top-most projection, this is the same structure associated with argument DPs, which have an SD head that selects for a PDP complement (see (49c)).

(49) a. LF: [$_{TP}$ [**EX** [**several pl woman**]] [Dist [1 [are [. . . [**EX**$_1$ [a sg doctor]]]]]]]

b. Lower copy + PN: [SD [$_{PDP}$ PD [Num NP]]]
c. Argument DP: [SD [$_{PDP}$ PD [Num NP]]]

Drawing on this similarity, my proposal is that (49b) is obligatorily parsed as a SDP, rather than a PDP, and that this leads to a crash at the PF-interface. WLM is therefore ruled out in these cases, blocking distributivity. To be slightly more explicit, I adopt the common assumption that heads can select for particular categories of arguments (Chomsky 1965; Grimshaw 1979; Pesetsky 1982; Svenonius 1994; Adger 2003; Collins 2003; Lechner 2004; Kobele 2006; Pesetsky and Torrego 2006; Müller 2010 among many others). In the present case, we can then say that SD selects for a PDP complement, which prevents a variety of semantically plausible but completely impossible Merge operations, e.g., combining a type $\langle et, \tau \rangle$ SD with an unsaturated *v*P, type $\langle et \rangle$, to form an SDP. My proposal concerning PNs then essentially amounts to the idea that language is endocentric (Chomsky 1957; Jackendoff 1977, Chomsky 1995; Collins 2003; Chomsky 2013), and is summarized in (50).

(50) **Endocentricity**:
Given a head X selecting for category Y, and a complex expression YP of category Y, the result of Merge(X,YP) is of category X, i.e., X projects.

In cases like (49b), (50) has the effect that Merge(SD,PDP) always results in SD projecting.[17]

Recall that the aim of applying WLM in these cases is to derive a structure with a moved plural DP that distributes over the predicate formed via movement. Bearing in mind the projection issue above, the LF for such a structure would presumably be as in (51): the relevant SD is merged first in the subject position of the PN and projects, then get re-merged in Spec(TP), at which point its NumP complement undergoes WLM.

(51) [$_{TP}$ [$_{\mathbf{SDP}}$ **SD** [$_{PDP}$ several women]] [. . . [$_{\mathbf{SDP}}$ **SD** [$_{PDP}$ a doctor]]]]

I will now show that (51) is ill-formed, so that WLM is blocked for the subject of PNs. To this end, even if we assume that the syntax can build such a structure, and there are a variety of potential syntactic violations, (51) crashes at the PF interface. This follows under the well-established facts (i) that (in English) exactly one copy per chain must be spelled out (Chomsky 1995; Nunes 2001; Fox and Pesetsky 2005; a.o.; Fox 2017) and (ii) that each terminal must be spelled out at least once (Chomsky 1995; Sportiche 2016; Fox 2017). For concreteness, I provide Fox's (2017) statements of these principles below.

(52) **Copy spell-out**:

[17] Note that this formulation does not lead us to expect that the bare SD should project in other cases of WLM. For example, suppose D is merged as the sister of V or P, both categories that presumably select for a SDP. Given (50), we correctly predict that V/P projects, just as if a full DP had been merged. Likewise, if SD is merged in the basic subject position, Spec(*v*P), whatever principles dictate that *v*P projects over an SDP can apply here without issue, as SD does not itself select for *v*P, so (50) does not apply.

Let [YP_2 [$_{XP}$ X^0[...YP_1 ...]]] be a phrase in which X^0 agrees with Y^0, the head that is shared by YP_1 and YP_2. For every terminal t dominated by YP_1, delete t at spell-out.
(Fox 2017: 31)

(53) **Full Interpretation** (Phonology):
At the point of linearization, every terminal must be associated with at least one position.
(Fox 2017: 33)

Given (52), *a doctor* in the lower copy of SDP is not spelled out in (51), violating (53). It follows that (51) crashes at the PF interface, so that WLM is ruled out for subjects of PNs.

(54) **No WLM with PNs**:
WLM is impossible for the subject of a PN.

Granting that WLM is a prerequisite for the distributive reading, this therefore derives the anti-distributivity effect.[18] Moreover, because nothing we have said hinges on the PN being indefinite, nor on the distributive operator being covert, the result here extends immediately to the other cases in the introduction: distributivity is also blocked over definite PNs and with overt distributors.

3.3 Predictions

The present analysis derives the anti-distributivity effect from the fact that number must be represented in the lower copy of a moved subject of a PN. This furnishes a number of novel predictions that I will now argue are borne out.

3.3.1 Agreement mismatch in partitive relatives

To begin, recall from the introduction that distributive readings are possible for the subject of PNs in the limited context of agreement-mismatched relative clauses. The agreement mismatch involved in such examples is limited to cases involving relative clauses (RCs) at the right edge of partitive DPs, so that Longenbaugh terms such cases *Agreement mismatch in partitive relatives* (AMPR). The availability of the distributive reading is straightforwardly predicted under the present analysis given Longenbaugh's (2017) treatment of AMPR, which I present briefly below.

(55) a. Sue is one of the five linguists who is/*are also a CEO.

[18] A natural question is how to rule out an LF along the lines below, where we have put a distributive operator within the PN. The resulting structure is interpretable and gives rise to intuitively correct truth conditions. Given that this reading is unavailable, something must be said to block this LF.

(i) a. *Several women are a doctor.
b. [[several [PL women]] [1 [... [[several$_1$ [PL women]] [Dist [a doctor]]]]]]
c. $[\![$[Dist [a doctor]]$]\!] = \lambda X.\, \forall x[(\text{ATOM}(x) \wedge x \sqsubseteq X) \rightarrow [\![$a doctor$]\!](x)]$

My proposal here is simply that *Dist*, as an adverb, cannot be applied within the DP. We can see this overtly with *each*, which I am taking to be roughly equivalent to *Dist* in meaning and distribution. Note, moreover, that blocking *Dist* from applying within NP is needed independently of the current discussion concerning predicate nominals: if *Dist* can apply within NP, it would be possible for singular DPs to end up with plural reference in general, as in (ii).

(ii) a. the student
b. [the [Dist [SG student]]]
c. $[\![$(iib)$]\!] = \text{MAX}\{X : \forall x[(\text{ATOM}(x) \wedge x \sqsubseteq X) \rightarrow *\text{STUDENT}(x)]\} = [\![$[the [PL students]]$]\!]$

b. Bill is one of the only professors here who is/*are a chef.

Before presenting the account, I briefly review some basic facts about the structure of partitives and RCs upon which the proposal builds. Concerning partitives, the crucial result is that they have a silent head noun that is optionally elided under identity with the post-prepositional NP (see Jackendoff 1977; Cardinaletti and Giusti 1992; Zamparelli 1995; Sauerland 2004; a.o.).

(56) a. [One [~~book~~ of [the books]]]
b. [Two [~~girls~~ of [the girls]]]

Concerning RCs, we will assume they are ambiguous between a matching structure, where a relative-clause internal DP headed by a null-operator is moved to Spec(CP) and the associated NP is elided under identity to a relative-clause external head noun, and a raising structure, where the head NP is pied-piped to the clause edge, then extracted to a relative-clause external position and composed with a determiner (Bhatt 2002; Hulsey and Sauerland 2006).

(57) a. $[_{DP}$ the [book $[_{CP}$ [which/Op ~~book~~]$_1$ [I read t_1]]]] (*Matching structure*)
b. $[_{DP}$ the $[_{CP}$ book$_2$ $[_{CP}$ [which/Op t_2]$_1$ [I read t_1]]]] (*Raising structure*)

Longenbaugh's proposal is then that agreement mismatch involves a matching relative clause where the relative-clause-internal head noun takes the (silent) head NP of the partitive as its antecedent, rather than the plural external head. The optionality of agreement reflects the choice between the structures in (58b,c).

(58) a. One of the books that was on the table …
b. [One [[sg book] [of [the [[pl books] $[_{RC}$ [Op [pl books]] [t were …]]]]]]] *Expected Agr*
Ellipsis
c. [One [[sg book] [of [the [[pl books] $[_{RC}$ [Op [sg book]] [t was …]]]]]] *AMPR*
Ellipsis

Agreement mismatched RCs therefore involve a syntactically and semantically singular RC-internal head.[19] We therefore predict that it should be possible to distribute a plural, relativized subject of a PN in AMPR environments just in case the agreement is mismatched. This is borne out, cap-

[19] In addition to the singular verbal agreement, the status of the head noun as syntactically singular can be diagnosed by the fact that it can bind singular pronouns, even if the tense/aspect of the clause makes it impossible to see singular verb agreement.

(i) a. John is one of the only students who Sue convinced to submit his paper.
b. Mary is one of the few professors who can still remember writing her dissertation.

The semantically singular status of the head noun can be independently confirmed by the fact that agreement mismatched subjects are incompatible with collective predicates (see (iii)), which require semantically plural subjects. Note collective predicates permit *syntactically* singular subjects (*team* in (ii)), so this prediction is independent of syntactic number.

(ii) They are/the team is/*Sue is gathering on the field.

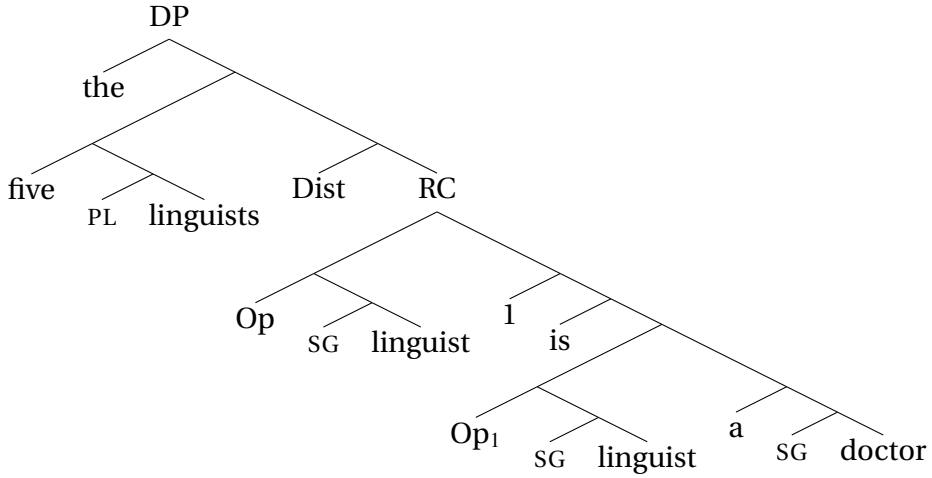
(iii) a. The police will arrest one of the two men who *is/are meeting in that café.
b. Sue is one of the many people who *is/are gathering at Boston Common today.

Moreover, the head NP in agreement mismatch contexts is unable to serve as the antecedent to a reciprocal pronoun (see (v)), which requires a semantically plural antecedent. Syntactically singular nouns can, in some cases, serve as the antecedent to a reciprocal (see (iv); e.g., Landau 2001: 49ff.), so this prediction is once again independent of syntactic number.

turing the core data. The structure assigned to such examples is as in (59), where the distributive operator applies at the top of the RC. The RC is then interpreted as in (59c),[20] so that the entire Dist+RC complex is assigned the denotation in (59d).[21] Combining this with the plural external head noun yields the distributive reading, as desired.[22]

(59) a. Sue is one of the five linguists who is also a CEO.

b.



c. $[\![\text{RC}]\!] = \lambda x : \text{ATOM}(x) \wedge *\text{LINGUIST}(x).\ *\text{DOCTOR}(x)$

d. $[\![\text{Dist RC}]\!] = \lambda X : *\text{LINGUIST}(X).\ \forall x[(\text{ATOM}(x) \wedge x \sqsubseteq X) \rightarrow *\text{DOCTOR}(x)]$

3.3.2 Bound plural pronouns

Consider now the second exception to the anti-distributivity property of PNs, the case of plural, bound-pronominal subjects.

(60) a. ?Several of the medical students think they're already a doctor.

b. These men all think they are they only one in the room. (Heim 2008)

These data too are predicted straightforwardly by the present account if we adopt the morph-syntactic transfer/deletion account of features on bound pronouns (Kratzer 1998, Schlenker 2003, von Stechow 2003, Heim 2008). I adopt the specific formulation of this account due to Heim, which holds that bound pronouns have the option of entering the derivation with unspecified φ-features (person, number, gender). These features are then valued at PF by copying the features on the associated binder. The φ-features on bound pronouns therefore need not be interpreted

(iv) The couple greeted ?each other/*themselves and sat down.

(v) a. Sally is one of the only linguists who *gets/get along with each other.

b. Smiths is one of the five grocery stores that *sells/sell each other's products.

[20] For simplicity of representation, I assume the RC-internal head reconstructs. Nothing hinges on this.

[21] I assume that universally quantifiers presuppose that every member of the restriction satisfies the presuppositions of the nuclear scope (Karttunen and Peters 1979; Heim 1983; Schlenker 2007).

[22] This analysis depends on it being possible to merge *Dist* at the root of a relative clause. While it is not unheard of for adverbial elements, specifically sentence adverbs, frequentatives, evaluatives, to attach high in the clausal domain, overt *each* cannot attach at the top of a RC.

(i) The five linguists (*each) who wrote a paper...

For now, we'll just have to assume that this is a difference between the overt and covert version of this operator.

at LF. This allows us to give a straightforward explanation for so-called "fake indexicality," where a 1st or 2nd person pronoun functions as a bound variable, e.g. *only I did my homework*. Crucially for present purposes, this theory also suggests a straightforward analysis of data like (61), where interpreting the PL feature on the bound pronoun would give rise to semantic anomaly.

(61) a. These candidates each think they are better than all their competitors.
(Heim 2008: 46)
b. These professors each believe they are smarter than the others.

Granting that the plural feature on a bound plural pronoun need not be interpreted, the data in (60) then follow straightforwardly on the present account: the underlying subject of the PN can be the number-neutral pronoun *they*, which is bound by the surface subject, as in (62).[23] The distributive operator can therefore be applied freely.

(62) a. These men each believe they're the only one in the room.
b. [[these [pl men]] [Dist [1 [[these$_1$] believe [... [[$_{\text{PDP}}$ they$_1$ [the only one in the room]]]]]]]]
c. $[\![\text{they}_1]\!]^g = g(1)$

3.3.3 Plural PNs

The present account also leads to the prediction that distributivity should be impossible for plural predicate nominals. The logic here is considerably simpler than in the singular case. In particular, consider the plural predicate nominal in (63), which on the present account has a structure as in (63b) and an interpretation as in (63c).

(63) a. Several women are doctors.
b. [$_{\text{PDP}}$ PL doctors]
c. $[\![[\text{PL doctors}]]\!] = \lambda X : \text{SUM}(X).*\text{DOCTOR}(X)$

Crucially, because we are assuming number features are presuppositional, *doctors* is undefined for atomic entities. It follows that even if it was possible to WLM in the subject of a plural predicate nominal, it would still be impossible to distribute, as the predicate is undefined for atoms.[24]

Unlike in the case of singular predicate nominals, where the anti-distributivity can be read off the surface unacceptability, we must work slightly harder to verify this prediction in the case of plural PNs. Here, no distributivity is needed to derive the most natural reading, as the plural subject can compose directly with the plural predicate.

With this in mind, one way to verify that distributivity is indeed blocked is to raise a plural subject over a singular indefinite, then checking for a reading where the subject distributes over the indefinite. This reading is not available.

(64) a. Several students are believed by a police inspector to be criminals.
b. These books seemed to a collector to be classics.

A second argument is that the overt distributive operator *each* is with plural PNs, again as ex-

[23] Note that we must still assume WLM applies in the A-movement step that displaces the surface subject to Spec(TP) and that is responsible for binding the lower pronoun.

[24] Independently, nothing we have said so far hinges on the number feature of the predicate nominal. It follows that WLM should be independently ruled out, blocking distributivity along the same lines as above. Thus, even if we don't assume number features are presuppositional, we are lead to the conclusion that singular and plural predicate nominals should behave similarly in rejecting distributivity.

pected [25]

(65) a. #These women are each doctors.
b. #These linguists are each semanticists.

3.3.4 Distributive determiners

The determiners *each, both, all* are well known to license distributive readings over their restriction.

(66) a. Each student here is working on a paper.
b. Both doctors ordered a prescription.
c. All applicants must submit an application by 5PM on July 31st.

Observe that DPs headed by *each* bot not by *both, all* are singular. It follows that the derived predicate created by moving the former type of DP but not the latter two types will contain a lower copy with a singular feature, so such DPs should make licit subjects of PNs. This is borne out:

(67) a. Each student here is also a professor.
b. LF: [[each [SG student]] [Dist [1 [... [[each$_1$ [SG student]] [PDP a [SG doctor]]]]]]]

(68) a. *Both/all students here are also a professor.
b. LF: *[[both [PL student]] [Dist [1 [... [[both$_1$ [PL student]] [PDP a [SG doctor]]]]]]]

The situation differs with plural PNs, where *all* and *both* DPs are licensed, counter to initial expectations. This is also true for the floated variants of these quantifiers (see fn. 25 above).

(69) a. Both/all students here are also doctors.
b. Both/all linguists here are syntacticians.

With *all*, this behavior is unsurprising once we grant that it licenses a non-distributive, collective reading (see (70); Brooks and Braine 1996; Kaup et al. 2002; Dotlačil 2010). Accepting this, the compatbility of *all* with plural PNs is expected, as no distributivity is required in such cases. Singular PNs, on the other hand, require an atomic argument and hence force distributivity, which is blocked, explaining the infelicity observed above.

(70) a. (All) the boys (all) lifted five pianos. (*collective reading possible*)
b. (All) the boys (all) gathered in the hallway.
(Dotlačil 2010)

With *both*, collective readings are considerably harder, although some speakers accept it (Dotlačil 2010).

(71) Context: exactly two men work in this office.
a. ?Both the men/the men both lifted five pianos. (*collective reading possible*)
b. ?Both the men/the men both gathered in the hallway.

[25] The quantifiers *both* and *all* are licensed here. I return to this issue below.

(i) a. These men are both doctors.
b. These linguists are both field-workers

I must leave to future research the question of why *both* is compatible with plural PNs, aside from pointing out that collective readings are somewhat possible.

3.4 A problem from Condition C

To conclude this section, I review some data involving Condition C that present a *prima facie* problem for the analysis. I propose a minimal elaboration on the core proposal that allows it to capture the problematic data while maintaining the essential results of the previous sections.

The problematic data is as follows. Recall from Section 2.1.2 that WLM was originally proposed to capture the bleeding effect of A-movement on Condition C. We are therefore lead to expect that A-movement of the subject of a PN should fail to bleed Condition C, since WLM is impossible here. This isn't borne out.

(72) a. *It seems to him_1 that that picture of John_1 is a fake.
b. That picture of John_1 seems to him_1 to be a fake.

In the context of DP-movement, we have so far treated *late merge* as an all-or-nothing procedure: the lower copy either contains a full SDP or a bare SD head. This conforms to the common position that *late merge* cannot be "embedded", that is, that that it must target the daughter of the root of the structure. The data in (72) therefore appear to contradict our result that *late merge* is impossible for the subject of a PN.

It turns out, however, that we can capture this data in a manner consistent with our proposal by minimally revising the conception of *late merge* we have adopted. In particular, Fox (2017) has recently argued that *late merge* must be allowed to target deeply embedded constituents, not just the daughter of the root. Allowing *embedded late merge* (ELM) of this sort suggests an immediate solution to the challenge posed by data like (72). In particular, ELM makes possible a derivation where the lower copy in a DP-movement chain contains both an SD head and the PD head it selects for, but lacks the NP complement to PD, which can be late merged in the higher copy. Because the lower copy is then the complex object [SD [PD ∅]], ELM allows us to avoid the projection problem with PNs while capturing the absence of Condition C effects. The example in (72b) would have an LF along the lines in (73), on this view.

(73) [[That [SD $\overbrace{\text{[picture of John}_1\text{]}}^{\textit{late merge}}$]] $\underbrace{\text{]}}_{\text{higher copy}}$ [1 [seems to him_1 [... to be $\overbrace{\text{[[}_{\text{PDP}}\ \underbrace{\text{[that}_1\ \text{[SD ∅]]}}_{\text{lower copy}}\ \text{[}_{\text{PDP}}\ \text{a fake]]}}^{\text{PN}}$]]]]]]

This makes it possible to reconcile the fact that WLM is impossible with PNs with the Principle A effects, taking WLM here to refer to the derivation where the lower copy contains only an SD^0 element.

Granting these assumptions, the challenge data like (72) pose to the present analysis shifts to ensuring that the new derviational machinery needed to capture these data does not undermine our analysis of the anti-distributivity effect. The remainder of this section is devoted to addressing this challenge.

Recall that distributivity is blocked over a derived predicate if the lower copy has an interpretable number feature. In the present context, we must therefore show that Num is obligatorily represented in the lower copy of a moved subject of a PN. To this end, recall from Section 3.1 that number is hosted at PD, and that PD heads may be recursively iterated when needed to introduce distinct elements. DPs can then have one of the following two structures: either Num is the only PD head in the structure, which plausibly covers bare plurals and PDP complements

to SDs like *every, most, all*, or there is at least one additional, iterated PD head, which covers all cases with an overt PD element.

(74) Num is the only PD head:

a. [$_{\text{SDP}}$ every/most/all [$_{\text{PDP}}$ PD.SG/PL student]] (*strong DPs*)

b. [$_{\text{SPD}}$ EX [$_{\text{PDP}}$ PD.PL students]] (*bare plurals*)

(75) There is at least one addition PD head:

a. [$_{\text{SDP}}$ IOTA [$_{\text{PDP}}$ PD.the [$_{\text{PDP}}$ PD.three [$_{\text{PDP}}$ PD.PL friends]]]] (*multiple overt PD^{0}s*)

b. [$_{\text{SDP}}$ EX [$_{\text{PDP}}$ PD.several [$_{\text{PDP}}$ PD.PL girls]]] (*weak DPs*)

In the former case, the minimal lower copy that avoids the projection problem will therefore be [$_{\text{SDP}}$ SD^0 [$_{\text{PDP}}$ Num $\emptyset$]], since Num is the only PD head. It follows that the lower copy must contain Num in such cases, ensuring the analysis carries through. In (76), I provide a general schematic for the lower-copy+PN in such instances.

(76) $\underbrace{[_{\text{PDP}}\ \overbrace{[_{\text{SDP}}\ \text{SD}_1\ \text{PD(=Num)}]}^{\text{lower copy}}\ [_{\text{PDP}}\ \text{PD NP}]]}_{\text{PN}}$

The latter case, in (75), presents more of a challenge. For the sake of the argument, let's assume that there is a PD head, PD′, that sits above Num in the structure.[26] Syntactically speaking, it is technically possible to avoid the projection problem by merging SD with PD′, producing a lower copy that does not contain a number feature. This would undermine the anti-distributivity result, so the present analysis depends on this structure being ruled out by extra-syntactic means.

My proposal to this effect is that while a structure like [SD [PD $\emptyset$]] maybe be syntactically well-formed, it is semantically uninterpretable unless a Num head is also present lower in the structure. In particular, recall that indexed determiners are interpreted either as type $\langle et, e\rangle$ or type e. Recall also that I am assuming that those PD heads that can co-occur with a plural number feature, e.g., *the, two, several, many, ...*, are all type $\langle et, et\rangle$. It follows that the structure [SD$_1$ PD′] is uninterpretable for both choices on interpretation for the indexed determiner:

(77) a. [SD$_{\langle et,e\rangle}$ PD$_{\langle et,et\rangle}$] (✗ *type mismatch*)

b. [SD$_e$ PD$_{\langle et,et\rangle}$] (✗ *type mismatch*)

Crucially, adding Num, which is type $\langle et\rangle$, to the structure in (77a) results in interpretability.

(78) [$_{\text{SDP:}\,e}$ SD$_{\langle et,e\rangle}$ [$_{\text{PDP:}\langle et\rangle}$ PD$_{\langle et,et\rangle}$ PD.Num$_{\langle et\rangle}$]] (✓ *interpretable*)

The present analysis can therefore be preserved as long as we assume that Num is type $\langle et\rangle$, while the elements *the, several*, and *n*, for $n \in \{1,2,3\cdots\}$, are all of type $\langle et, et\rangle$. I conclude this section by providing some independent justification for these types.

I begin by considering the types of the PD elements listed above. In each case, I will try to argue, in particular, that the meaning contributed by the relevant element cannot be captured with an entry of type et, which would allow it to combine felicitously with an indexed determiner. Consider first the case of *the*. Its contribution to the overall meaning is, informally speaking, to add uniqueness and existence presuppositions to its sister predicate. As we have seen, this

[26] It is difficult to independently determiner whether the PD head expressing number is merged above or below the PD heads bearing other overt PD elements, such as *the, a, several, etc.*. For the sake of the argument, I will therefore assume that Num is not the highest head, as this is the more challenging case.

meaning is easily encoded in the following type $\langle et, et \rangle$ denotation.

(79) $[\![\text{the}]\!] = \lambda P : \exists! x[P(x)].\ P$

The crucial question at present is therefore whether this basic contribution cab be captured with a denotation of type $\langle et \rangle$. In this case, the answer is a straightforward no: there is no way to contribute a uniqueness and existence presupposition to a predicate via predicate modification, which is how a type $\langle et \rangle$ *the* would be limited to combining with its sister. It follows that $\langle et, et \rangle$ is the lowest possible type that can successfully capture the intuitive contribution of *the*.

Moving on, the next PD heads I consider are the cardinal numerals. In the example denotations provided in Section 3.1, I assumed an extremely simple meaning for cardinal numerals whereby they simply add a cardinality assertion to the predicate they combine with, as in (80a). If this is indeed the intuitive meaning of such determiners, it is possible to accomplish the same overall meaning with a type $\langle et \rangle$ predicate, as in (80b) that combines via predicate modification with its sister.

(80) a. $[\![\text{three}]\!] = \lambda P.\lambda X.\ |X| = 3 \wedge P(x)$
b. $[\![\text{three}]\!] = \lambda X.\ |X| = 3$

However, as Ionin and Matushansky (2006) argue, it is likely that the denotations above are too simple to correctly capture the full range of readings possible with cardinal numerals. The argument concerns complex numerals, like *five hundred*. Assuming a simple syntax where numerals combine iteratively with their NP/NumP complement, as in (81b), we derive the contradictory interpretation for (81a) given in (81c).

(81) a. Five hundred books.
b. [Five [hundred books]]
c. $\lambda X.|X| = 5 \wedge |X| = 100 \wedge X \in *\text{BOOK}$

Ionin and Matushansky (2006) show that if we want to maintain the simple syntax in (81b), more sophisticated meanings are required for cardinal numerals. Their proposed lexical entries, which I repeat below, hinge on the following intuition: cardinal numerals do not measure cardinality in terms of the number of atomic individuals satisfying a given denotation, but rather in terms of how many discrete subparts a given set can be broken down into. They formally cash this out via the notion of partitions, as defined in (83).

(82) $[\![\text{n}]\!] = \lambda P_{et}.\lambda x_e.\exists S[\Pi(S)(x) \wedge |S| = n \wedge \forall s \in S[P(S)]]$

(83) a. $\Pi(S)(x) = 1$ iff:
(i) S is a *cover* of x and
(ii) $\forall z, y \in S[z \neq y \Rightarrow \nexists a[a \sqsubset z \wedge a \sqsubset y]]$ (cells of the partition do not overlap)

(84) A set C is a cover of a plural individual X iff X is the sum of all members of C.

As a demonstration, *five hundred books* is now interpreted as follows.[27]

[27] As Ionin & Matushansky discuss at length (Sec 3.1), this analysis depends on the complement to the lowest cardinal numeral being singular. If it were plural, the meaning would be compatible with a situation where the denotation of the noun consists of n distinct sub-groups, rather than the intended meaning where there are n distinct atomic subgroups. In our system, this means that a cardinal numeral must have two instances of number, the higher of which is expressed syntactically on the noun. Ionin and Matushanksy reach a similar conclusion.

(i) [PL [THREE [SG book]]]

(85) a. [Five [hundred books]]

b. $[\![\text{hundred book}]\!] = \lambda X_e.\ \exists S[\Pi(S)(X) \wedge |S| = 100 \wedge \forall s \in S[[\![\text{book}]\!](s)]]$
$= \lambda X_e.$ X is a plural individual divisible into 100 non-overlapping individuals p_i such that their sum is x and each p_i is a book

c. $[\![\text{five [hundred book]}]\!] = \lambda X_e.\ \exists S[\Pi(S)(X) \wedge |S| = 2 \wedge$
$\forall s \in S[\exists S'[\Pi(S')(s) \wedge |S'| = 100 \wedge \forall s' \in S'[[\![\text{book}]\!](s')]]]]$
$= \lambda X_e.$ X is a plural individual divisible into 2 non-overlapping individuals p_i such that their sum is X and each p_i is divisible into 100 non-overlapping individuals p_k such that their sum is p_i and each p_k is a book

(Ionin & Matushanksy 2006: 9)

Granting Ionin & Matushanksy's approach is on the right track, for our purposes it suffices to show that this more sophisticated meaning cannot be encoded in a meaning of type *et*. Ionin & Matushansky (Section 2.2) argue at length for exactly this conclusion, so I refer interested readers here for detailed discussion, but the gist of the argument is quite simple: predicate modification is an insufficient tool to compose cardinal numerals with the basic meaning above. For a brief demonstration, suppose we adopt the following type *et* meanings for cardinal numerals, which attempts to encode the meansings above in the lower type.

(86) a. $[\![\text{two}]\!] = \lambda X.\exists S[\Pi(S)(X) \wedge |S| = 2]$

b. $[\![\text{hundred}]\!] = \lambda X.\exists S[\Pi(S)(X) \wedge |S| = 100]$

c. $[\![\text{two hundred books}]\!] = \lambda X.\exists S[\Pi(S)(X) \wedge |S| = 2] \wedge \exists S'[\Pi(S')(X) \wedge |S'| = 100 \wedge [\![\text{books}]\!](x)]$

(Ionin & Matushansky 2006: 13)

Given (86), $[\![\text{two hundred books}]\!]$ is satisfied by any individual X that can be broken down into both 100 and 2 non-overlapping sub-groups, so any set of atomic books of cardinality 100 is in this denotation. This is obviously wrong. I will therefore follow Ionin & Matushansky in concluding that cardinal numerals must be of type $\langle et, et \rangle$

The remaining PD heads to be considered behave, by and large, like cardinal numerals in terms of how they compose with each other, suggesting a uniform analysis is warranted. I focus here on the specific cases of *several* and *many*, both of which combine with cardinal numerals in a manner similar to that documented above.

(87) a. Sue owns several thousand books

b. In reality, I lay many hundred miles away in an alien land. (COCA)

Given that *several thousand* in (87a) means something like >2,000 and *many hundred* in (87b) something like >200, by the same logic employed by Ionin & Matushansky, we need the more sophisticated relative-cardinality denotations to capture these meanings with a simple syntax. I therefore propose the following lexical entries with *several* and *many*, setting aside for now how the quantities SEVERAL and MANY are obtained.

(88) a. $[\![\text{several}]\!] = \lambda P_{et}.\lambda x_e.\exists S[\Pi(S)(x) \wedge |S| = \text{SEVERAL} \wedge \forall s \in S[P(S)]]$

b. $[\![\text{several}]\!] = \lambda P_{et}.\lambda x_e.\exists S[\Pi(S)(x) \wedge |S| = \text{MANY} \wedge \forall s \in S[P(S)]]$

The main point is that because these PD heads compose with cardinal numerals just like cardinal numerals, they too must be type $\langle et, et \rangle$.

To summarize, I have considered the main types of PD heads that can are licensed in a plural DP and argued that they must be typed as $\langle et, et \rangle$ in order to capture their intuitive contribution

to the meaning. It follows that [SD [PD ∅]] will be semantically uninterpretable for all choices of PD so far considered.

The final component of the argument is to argue that Num heads are type *et*. Here I appeal straightforwardly to simplicity: the intuitive meaning of number in the present system is to alter the definedness conditions of the sister predicate to the Num head, and this can be accomplished by positing a type ⟨*et*⟩ Num head that composes via predicate modification. As we have discussed, this is in contrast to the PD heads discussed above, which really seem to require higher-typed meanings. This sort of conceptual argument is admittedly hard to independently support empirically, but the general force of the proposal is intuitive and generalizes easily, i.e., always pick the lowest type that can capture the intuitive contribution of the element to the meaning.[28]

In summary, the framework that we have developed permits the following types of lower copies in DP movement chains, setting aside the possibility for ELM below the Num head.

(89) **Permissible lower-copy structures**:

a. [SD_1 ∅] (*wholesale late merge*)
b. [SD_1 [$_{PDP}$ … [Num ∅]]] (*embedded late merge*)
c. [SD_1 [$_{PDP}$ … [Num […NP]]]] (*no late merge*)

Given this state of affairs, we can explain the formerly problematic data introduced at the start of this section while still capturing the fact that plural subjects of PNs cannot distribute over them, as follows. Because of the projection problem pointed out in the previous section, WLM is blocked for the subjects of PNs. This means that the lower copy in such cases must have the form in (89b) of (89c). Option (89b) allows us to avoid Condition C violations that the moved DP would induce in its surface position. Because both (89b) and (89c) involve Num features, we also capture the anti-distributivity effect for such subjects.

4 A′-Movement and distributivity

Now that the core account of the PN facts is in place, I turn my attention to the related issue of how A′-movement interacts with distributivity. As discussed briefly in Section 2.1.2, T&H claim WLM is impossible in A′-chains. This leads to the prediction, following the results of Section 2.2, that A′-movement should never feed distributivity. In this Section, I show first that the data does not support T&H's claim unequivocally, and that WLM is possible in A′-chains in many instances. Adopting this more nuanced view, I then argue that the proposal from Section 2.2 sheds light on the mysterious fact that A′-movement, both overt and covert, is a relatively poor licensor of distributivity.

In developing their WLM proposal, T&H adopt the common claim that unlike A-movement, A′-movement does not bleed Condition C, in general (Freidin 1986; Lebeaux 1988; Fox 1999;

[28] Given the difficulty of empirically justifying the type of Num^0, it's worth considering briefly what happens if Num is a higher type. This does not, in and of itself, compromise the argument, since having a type ⟨*et*, *et*⟩ Num head simply means that some lower, type *et* head must be added to the structure to allow it to be interpretable. I illustrate below.

(i) a. [$SD_{\langle et,e\rangle}$ [several$_{\langle et,et\rangle}$ [Num$_{\langle et,et\rangle}$ ∅]]] (✗ *type mismatch*)
b. [$SD_{\langle e\rangle}$ [several$_{\langle et,et\rangle}$ [Num$_{\langle et,et\rangle}$ ∅]]] (✗ *type mismatch*)

It follows that there must be some lower *et* phrase present to allow compositionally. While I won't investigate this further, this suggests that even if we treat Num as higher-typed, the argument may still carry through if there is a lower NP head that can provide a type *et* base, allowing for ELM at some lower level.

a.o.).[29] In (90), I repeat some of the core examples that have been taken to support this claim.

(90) a. [??/*]Which argument that John$_1$ is a genius did he$_1$ believe?
(Fox 1999:164)
b. [*]Which corner of John$_1$'s room was he$_1$ sitting in?
(T&H 2009: 391)
c. [*]Which report that John$_1$ was incompetent did he$_1$ submit?
(Freidin 1986)
d. [*]The claim that John$_1$ was asleep, he$_1$ won't discuss.
(Chomsky & Lasnik 1993)

While the examples above are indeed degraded for some speakers, there is a considerable variation across both speakers and examples in how unacceptable structurally parallel examples are judged to be (Lasnik 1998, 1999; McCarthy 2004; Jacobson 2004; Kuno 2004). In fact, concerning (90)[d], Lasnik (1998) has himself rescinded his claim that the example is ungrammatical, and he likewise suggests that the difficulty of (90)[c] may be due to pragmatic factors. Many speakers I have consulted likewise do not find (90)[a,b] to be unequivocally impossible. These examples must moreover be compared against a variety of parallel cases in the literature which have been reported to be quite good.

(91) a. Which piece of evidence that John$_1$ was guilty did he$_1$ successfully refute?
b. How many arguments that John$_1$'s theory was correct did he$_1$ publish?
c. Which proof that Mary$_1$'s theory is superior to John's did she$_1$ present? (Lasnik 1998)
d. Whose allegation that John$_1$ was less than truthful did he$_1$ refute vehemently?
e. Whose claim that the senator$_1$ had violated the campaign finance regulations did he$_1$ dismiss as politically motivated? (Kuno 2004)
f. The claim that the director$_1$ was corrupt, he$_1$ was unwilling to discuss.
g. The widespread belief that John$_1$ is incompetent, he$_1$ deeply resents. (Postal 1997)

Given these data, it seems that we must acknowledge that with some examples and some speakers, A′-movement of an R-expression past a c-commanding co-indexed pronoun does result in a notable improvement. On the one hand, it therefore follows that the unequivocal claim that A′-movement does not bleed Condition C is untenable. On the other hand, the preponderance of data like (90) in the literature and of speakers who share, at least to some degree, the reported judgements, supports the existence of an A/A′-contrast in this domain: I am not aware of a single case where A-movement does not readily bleed Condition C, for all speakers. Moreover, even in the relatively acceptable examples in (91), there is in most cases still a contrast between the version of the example with a Condition C violation at the tail of the chain and the version without it.

(92) a. ?/??Which proof that Mary$_1$'s theory is superior to John$_1$'s did she$_1$ present?
b. Which proof that her$_1$ theory was superior to John$_1$'s did Mary$_1$ present?

Under the theory of movement adopted here, this state of affairs suggests that that WLM is indeed possible in A′-chains, subject to some as of yet unknown constraints. I will therefore adopt the following position concerning WLM.

[29] I set aside those cases where the offending R-expression is contained in a modifier to the A′-moved phrase, in which case it is widely acknowledged that A′-movement does bleed Condition C (Freidin 1986; Lebeaux 1988, 1998; Fox 1999; a.o.). None of the examples I discuss involve R-expressions contained in modifiers, so I set this complication aside here.

(93) **Late merge and A/A′-movement**:
Late merge is always possible in A-chains and marginally possible in A′-chains.

Taking (93) in conjunction with the results of Section 2.2, we predict that A′-movement should be only marginally capable of feeding distributivity, with acceptability of distributive readings corresponding roughly with the acceptability of A′-moving an R-expression out of a Condition-C-violating configuration.

(94) **A′-movement and distributivity**:
A′-movement feeds distributivity, with difficulty.

Consider first the case of covert A′-movement, and specifically QR. It is well documented in the literature that inverse-scope distributive readings are, in general, significantly harder than plain inverse scope readings (vanLehn 1978; Gil 1982; Tunstall 1998; Pylkkänen and McElree 2006; Dotlačil 2010). This result, which is otherwise very puzzling, is exactly as predicted: the inverse scope distributive reading requires both QR, itself already a costly operation (Wurmbrand 2015), *and* WLM in the higher copy of an A′-chain.[30]

(95) a. A linguist submitted five papers. (?? *distributive*)

b. [[[[EX $\overbrace{\textbf{five [pl papers]}}^{\text{WLM}}$] [Dist [1 [[a [SG linguist]] [2 [[a_2 [SG linguist]] [submitted [EX_1]]]]]]]]]

I provide several additional examples below.

(96) a. A student read several books. (?? *distributive*)
b. A boss fired these employees. (?? *distributive*)
c. A teacher recommended these students. (?? *distributive*)
d. A speech inspired the soldiers. (?* *distributive*)

Inverse-scope distributive readings are significantly easier for DPs headed by *every/each*, which is again predicted here: since the DPs are singular, the derivation need not involve WLM.

(97) a. A student came up with each of these ideas. (✓ *distributive*)
b. Someone dislikes every paper I write. (✓ *distributive*)

Similar results hold for overt A′-movement, where distributive readings for moved plural DPs are slightly more accessible but still generally marginal.

(98) a. Which books did a student write? (?? *distributive*)
b. I only bought the books that a student wrote. (? *distributive*)
c. Sally owns three paintings that a Spanish artist painted. (? *distributive*)
d. Which talks did a student give? (?? *distributive*)
e. Which stores did a famous singer endorse? (?? *distributive*)

[30] Question: how is WLM spelled out in this case, given that late-merged adjuncts are usually pronounced as disjoint from the rest of the quantifier, but we can never split D and NumP overtly:

(i) *I read several yesterday books

Also, does is inverse scope distributivity blocked if there is a WLM'd adjunct?

(ii) a. A linguist submitted five papers last year that I read.
b. A professor gave five students an A.

Finally, there is one context where inverse distributive readings are readily available, namely when the plural DP and the relevant singular indefinite are co-arguments within the VP.

(99) a. John gave a pumpkin pie to two girls. (✓ *distributive*)
(Roberts 1987)
b. Sue has written a report about several topics. (? *distributive*)

Strikingly, A′-movement readily bleeds condition C in parallel examples, suggesting that WLM is licensed here more easily than in the contexts above.

(100) a. *I told him$_1$ about two arguments that John$_1$'s claim is wrong.
b. Which two arguments that John$_1$'s claim is wrong did you tell him$_1$ about?

(101) a. *We asked him$_1$ about that picture of John$_1$.
b. Which picture of John$_1$ did you ask him about?

In conclusion, while I must leave for further work a precise determination of the factors controlling WLM in A′-chains, the results of this section conform to the core predictions of the present account, as well as provide an explanation for the puzzling fact that A′-movement does not readily license distributive readings in most contexts.

5 Remaining issues

5.1 Plurals

The simple theory of number I have embraced so far in the paper is well known to be inadequate, in unmodified form, to capture the full array of readings that plural DPs give rise to. To begin, the meaning of plural in 2.1.1 gives rise to truth conditions that are intuitively incorrect for bare plurals in the scope of negation, in polar interrogatives, and in the antecedent of conditionals.

(102) a. John doesn't have children.
≈ John doesn't have a single child.
b. Does John have children?
≈ Does John have one or more children?
c. If John has children, he shouldn't come.
≈ If John has one or more children, he shouldn't come.

Thus for example in (102a), the requirement that a plural DP denote a plurality means that (102a) is true in a situation where John has one child, since he therefore fails to satisfy the predicate *have children*. In these contexts, what is apparently needed is an "inclusive" interpretation for the plural that includes atomic individuals. This of course contrasts with the behavior of plurals in many other contexts, where the traditional meaning that excludes atoms seems to be active.

(103) a. John has children.
≈ John has two or more children.
b. Sue read books.
≈ Sue read two or more books.

Second, plural DPs containing a bound variable, as in (104), have been argued to have atomic entities in their denotation, so that, e.g., some boys can have only one friend.

(104) Every boy$_1$ should invite his$_1$ friends.

This is likewise not predicted on the present account, where $[\![\text{his}_1 \text{ friends}]\!]^g$ presupposes that $g(1)$ has at least two friends. In this section, I review the main approaches to these data, showing that none is fully compatible with our conclusions.

5.2 Alternatives

In this section, I discuss the alternative solutions of Sauerland et al. (2005), Spector (2007) & Zweig (2009), de Swart and Farkas (2010), and Grimm (2012). To begin, Sauerland et al. start from the hypothesis that plural DPs are semantically inclusive. The exclusive reading then arises via competition with the singular, which they assume introduces a presupposition that its argument be atomic. The pragmatic principle *Maximize Presupposition* (Heim) then dictates that singular must be used in all contexts where its presupposition is satisfied. The use of plural therefore implicates that the speaker was unwilling to commit to the presupposition associated with the singular, which gives rise to a non-atomicity inference, and hence an exclusive reading. A special assumption is then invoked to block this inference in downward entailing contexts, to capture the data in (102). For our purposes, the issue with this account is that the denotation of plural DPs always includes atoms. In our framework, we can model Sauerland's framework by assuming that PL is semantically vacuous. It follows that the lower copy in a movement chain will never introduce a presupposition that the associated derived predicate is defined only for atoms, which recall was the crucial step in blocking distributivity.

(105) a. Several women are a doctor.
b. LF: [[several [PL women]] [Dist [1 [... [[several$_1$ [PL women]] [a doctor]]]]]]
c. $[\![[\text{Dist } [1 \ldots]]]\!] = \lambda x : *\text{WOMAN}(x).\ *\text{DOCTOR}(x)$

In short, because non-atomicity arises in this framework as an utterance level inference, distributivity is not blocked over plural lower copies in a movement chain.

Spector's and Zweig's accounts run into a similar problem. Both authors likewise assume that plural DPs are semantically inclusive, then derive the exclusive interpretation as a second-order implicature: in upward entailing contexts, the exclusive reading is strengthened to the inclusive reading, while in downward entailing contexts, the exclusive reading is stronger and so is maintained. Once again, the problem here is that the relevant strengthening to the exclusive reading is an utterance level inference, as in (106).

(106) a. John read books.
b. Literal meaning: $\exists X[*\text{BOOK}(X) \wedge *\text{READ}(X)(J)]$
c. Strengthened meaning: $\exists X[*\text{BOOK}(X) \wedge *\text{READ}(X)(J)] \wedge \nexists X[|X| \geq 2 \wedge *\text{BOOK}(X) \wedge *\text{READ}(X)(J)$

It follows that the predicate formed by moving a plural DP will always be defined over atoms, since plural DPs are always semantically inclusive, blocking the account of anti-distributivity.

de Swart and Farkas (2010) adopt instead the position that the plural morpheme is semantically ambiguous between an exclusive and an inclusive interpretation. The choice among these alternatives is then made on the basis of a general pragmatic principle that prefers an utterance be assigned the strongest meaning possible.

(107) **Strongest meaning hypothesis for plurals**:
for a sentence involving a plural nominal, prefer that interpretation of [PL] which leads to the stronger overall interpretation for the sentence as a whole.

In upward entailing contexts, the exclusive meaning is thus preferred, and in downward entailing contexts, the inclusive meaning.

To see the challenge facing this approach in the context of our data, consider the paradigm case in (108), where there is a distributive operator in the sister to the moved DP *several women*.

(108) a. Several women are a doctor.
b. [[several [pl women]] [Dist [1 [. . . [[several$_1$ [pl women]] [a doctor]]]]]]

As we have seen, choosing the exclusive meaning for plural results in a presupposition failure in these cases. Presumably, this forces the choice of the inclusive meaning here, so that there is no need to apply (107). Granting our assumptions about the interpretation of movement chains, de Swart and Farkas (2010) therefore predict that sentences like (108) should be grammatical, and moreover that they should have an obligatory inclusive meaning, counter to fact.

It follows that none of the three approaches outlined above is straightforwardly compatible with the core results of this paper. That said, Grimm (2012) argues that there is reason to suspect that these approaches are not completely on the right track. The accounts share in common the fact that they specifically link inclusive readings to downward entailing contexts. This is transparent for the exhaustification based approaches and for de Swart and Farkas's (2010) approach; on the approach of Sauerland et al., the reasoning is more opaque but the conclusions are the same: bare plural indefinites in downward entailing contexts should be interpreted inclusively. Grimm (2012), however, observes that the choice between inclusive vs. exclusive readings is not strictly correlated with the downward or upward monotonicity of the context in which the relevant plural DP appears.[31] I repeat two of the key examples illustrating this below, and refer the reader to Grimm 2012 for additional cases and experimental evidence confirming the judgements. First, as first observed by Zweig (2009), some modal auxiliaries license inclusive readings in their scope (see (109)), although such contexts are not downward entailing, nor do they reverse scalar implicatures (Grimm 2012: 3).

(109) Sherlock Holmes should question local residents to find the thief.
$\not\approx$ Sherlock Holmes should question non-atomic collections of residents

Second, some transitive opaque verbs like *look for* license an inclusive reading for an indefinite plural object. Thus in the context below, (110b) indicates that finding even a single hiker would be acceptable.

(110) a. Context: I was out hiking and strayed from the trail. I'm not lost and worried.
b. I'm looking for hikers.
$\approx$ I'm looking for one or more hikers.

Following Zimmermann (2006), *look for* licenses upward monotonic inferences, so that the before above is especially surprising. Monotonicity alone, then, does not offer a sufficient characterization of the environments where inclusive and exclusive interpretations are available, pro-

[31] Inclusive readings are also not dependent on NPI-licensing environments, which is important since NPIs are licensed in some non-downward entailing environments.

(i) a. I am surpised that anything was there. (✓ *NPI*)
b. I am surprised that boxes were in the office (✓ *exclusive plural*)

(ii) a. #Both students who saw anybody reported to the police. (✗ *NPI*)
b. Both students who saw spies reported to the police. (✓ *inclusive plural*)

viding independent evidence against the three accounts presented above.

Grimm's alternative proposal is that the inclusive readings of bare plurals involves a special interpretive device, available only to bare plurals, whereby their number feature is ignored and they are treated as number-neutral predicates of kind-instantiation (following a related proposal by Krifka). This meaning then competes with the exclusive plural interpretation that we have been assuming thus far, with the choice between the two meanings determined by the type of information speakers are interested in at the point of utterance. Bare plural DPs are thus taken to be ambiguous along the lines below.

(111) a. $[\![\text{dogs}]\!] = \lambda x.\ x$ is an instantiation of the kind *dog*
b. $[\![\text{dogs}]\!] = \lambda X :\ \text{SUM}(X).\ *\text{DOG}(X)$

The choice among these readings is then determined by way of Roberts's (2012) theory of discourse, which models it as a winnowing process through which the participants try to determine which of many potential possible worlds they are in. The process proceeds via a question-and-answer procedure. Each point in the discourse is associated with a *Question Under Discussion*, which represents a partition on the currently entertained set of possible worlds (**?**). A given utterance is then felicitous iff it provides a (partial) answer to the QUD, thus helping to chose amongst cells in the partition and narrowing the entertained set of possible worlds. The choice among the inclusive and exclusive uses of the plural is then dictated by what the speaker takes to be the immediate QUD. The two relevant possibilities are what Grimm calls the existential and quantitative QUDs, which query the existence and the quantity of the DP in question. Taking the example *Ed saw dogs*, the corresponding existential and quantitative QUDs are then as in (112).

(112) a. Did Ed see any dogs? (*existential* QUD)
b. How many dogs did Ed see? (*quantitative* QUD)

The existential QUD represents a situation where it has not yet been determined whether or not Ed saw any dogs, while the quantitative QUD one where the speakers are trying to determine the number of dogs Ed saw. In a typical discourse, it follows that the quantitative QUD (QQUD) is only appropriate if the existential QUD (EQUD) has been resolved in the affirmative (see Grimm 2012: 7ff.): the quantity of dogs Ed saw is not normally at issue unless it has been determined that he saw one or more dogs in the first place.

The inclusive and exclusive readings are then licensed, respectively, when it is the EQUD and QQUD that is at stake. Consider the utterance *Ed didn't see dogs*, and suppose first that the EQUD is at stake. In this case, only the inclusive reading of the plural leads to an answer: on the exclusive reading, all that is being denied is that *Ed* saw a plurality of dogs, which doesn't resolve the EQUD. If it is instead the QQUD that is at stake, in most contexts we can assume the EQUD has been resolved in the affirmative (see above), which at present amounts to it being contextually entailed that Ed saw one or more dogs. In such a context, *Ed didn't see dogs* does not answer the QQUD on the inclusive reading, since the context only contains worlds where *Ed* saw one or more dogs. It follows that only the exclusive reading is licensed in such a context.

The fact that indefinite plurals in the scope of negation are typically inclusive then follows from the observation that in a typical context of use for such a sentence, the EQUD has not been affirmatively resolved: in most contexts where I utter *Ed didn't see dogs*, it is not true that Ed saw at least one dog. The QQUD, and correspondingly the exclusive reading, is therefore not ordinarily licensed in such cases. If, however, we select a context where an affirmative answer to the EQUD is plausibly entailed, exclusive interpretations become available, as predicted.

(113) I FAILED Algebra in High School – FAILED IT. You have to do a LOT of NOTHING to fail a class. I didn't turn in assignments, I didn't ask for help, and that's why I got an F.

World knowledge dictates that a class typically has many assignments, and presumably the speaker did at least one of them. The EQUD, *did the speaker turn in any assignments* is therefore affirmatively resolved, licensing the QQUD and hence the exclusive reading. A similar logic extends to the other core cases. I refer to reader to Grimm 2012 for discussion.

Assuming we limit the special, inclusive reading to bare plurals, this approach is compatible with the core data presented above. That said, it leads to a number of predictions that do not appear to be borne out, and moreover cannot capture the case of dependent plurals containing bound variables. The clearest prediction that fails to obtain is that bare plural DPs in, e.g., the scope of negation should license distributivity under A′-movement and should make acceptable subjects of singular predicate nominals: since these phrases are interpreted inclusively, there should be no problem distributing over their lower copy. While the distributive reading may be slightly more accessible in some cases (see (114)), it does not appear that bare plurals in the scope of negation can be the subject of singular predicate nominals, although adding *any* seems to make the reading easier.

(114) a. Sue didn't read (any) books that a student recommended last year. (? *distributive*)
b. John didn't solve (any) problems that a mathematician suggested. (? *distributive*)

(115) a. #No elephants are a reptile.
b. ?I doubt that #(any) doctors are also a good lawyer.

6 Conclusion

I began this paper by pointing out a puzzle concerning the impossibility of distributing a plural DP over a singular predicate nominal. I reviewed the main approach to this puzzle in the literature, due to Dotlačil (2011), and presented novel evidence showing that this approach faces empirical hurdles, in addition to the better known conceptual issues. In the remainder of the paper, I sketched a novel approach to this phenomenon that both overcomes the conceptual and empirical challenges with Dotlačil's approach. At the core of the proposal is the claim that if plural features contribute an exclusive interpretation, the *Dist* is barred from applying to the derived predicate formed by moving a plural DP, granting that the plural feature is represented in the lower copy. I then showed that Takahashi and Hulsey's (2009) mechanism of WLM provides a means of ensuring the plural feature is absent in the lower copy, licensing otherwise unavailable distributivity. It follows that movement can feed distributivutiy only if the plural feature has been WLM'd in the higher copy only. I then argued that the syntax of PNs is such that WLM is blocked for their subjects, deriving the anti-distributivity effect. Finally, I extended this approach to explain the difficulty of inverse-distributive readings under A′-movement. A pressing remaining issue that I have left for future research is how to reconcile my analysis with the fact that plural number appears to license inclusive readings.

References

Adger, D. (2003). *Core syntax: A minimalist approach*, volume 33. Oxford University Press Oxford.
Bhatt, R. (2002). The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics*, 10(1):43–90.

Bowers, J. (1993). The syntax of predication. *Linguistic inquiry*, 24(4):591–656.
Bowers, J. (2001). *Predication*. Blackwell Publishers Ltd.
Cardinaletti, A. and Giusti, G. (1992). Partitive "ne" and the QP-Hypothesis: a case study. In Fava, E., editor, *Proceedings of the XVII Meeting of Generative Grammar: Trieste, Feb. 22-24 1991*, pages 122–141. Turin: Rosenberg and Sellier.
Chierchia, G. (1998). Plurality of mass nouns and the notion of 'semantic parameter'. In Rothstein, S., editor, *Events and grammar*, volume 70, pages 53–103.
Chomsky, N. (1957). *Syntactic structures*. Mouton.
Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT press.
Chomsky, N. (1981). *Lectures on government and binding*. Foris Publications, Dordrecht, The Netherlands.
Chomsky, N. (1995). *The minimalist program*. MIT Press, Cambridge, Massachusetts.
Chomsky, N. (2000). Minimalist inquiries: The framework. In Martin, R. Michaels, D. and Uriagereka, J., editors, *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, pages 89–155. MIT Press.
Chomsky, N. (2013). Problems of projection. *Lingua*, 130:33–49.
Collins, C. (2003). Eliminating labels. In Epstein, S. and Seely, D., editors, *Derivation and explanation in the Minimalist Program*, pages 42–64. Oxford: Blackwell.
Coppock, E. and Beaver, D. (2015). Definiteness and determinacy. *Linguistics and philosophy*, 38(5):377–435.
Danon, G. (2011). Agreement and DP-Internal Feature Distribution. *Syntax*, 14(4):297–317.
de Swart, H. and Farkas, D. (2010). The semantics and pragmatics of plurals. *Semantics and pragmatics*, 3:6–1.
Dotlačil, J. (2010). *Anaphora and distributivity: A study of same, different, reciprocals and others*. PhD thesis, Utrecht University.
Dotlačil, J. (2011). Fastidious distributivity. In *Proceedings of SALT*, volume 21, pages 313–332.
Elbourne, P. D. (2005). *Situations and individuals*, volume 90. MIT Press Cambridge, MA.
Fara, D. G. (2001). Descriptions as predicates. *Philosophical Studies, 102*, pages 1–42.
Fox, D. (1999). Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry*, 30(2):157–196.
Fox, D. (2002). Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry*, 33(1):63–96.
Fox, D. (2017). Embedded Late Merge and the Theory of Displacement. *Ms*.
Fox, D. and Pesetsky, D. (2005). Cyclic linearization of syntactic structure. *Theoretical linguistics*, 31(1-2):1–45.
Freidin, R. (1986). Fundamental issues in the theory of binding. In *Studies in the Acquisition of Anaphora*, pages 151–188. Springer.
Gil, D. (1982). Quantifier scope, linguistic variation, and natural language semantics. *Linguistics and Philosophy*, 5(4):421–472.
Grimm, S. (2012). Plurality is distinct from number neutrality. In *Proceedings of NELS*, volume 41.
Grimshaw, J. (1979). Complement selection and the lexicon. *Linguistic inquiry*, 10(2):279–326.
Hackl, M. (2000). *Comparative quantifiers*. PhD thesis, Massachusetts Institute of Technology.
Heim, I. Artikel und definitheit. In von Stechow, A. e. a., editor, *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, pages 487–535. de Gruyter, Berlin.
Heim, I. (1983). On the projection problem for presuppositions. In *WCCFL 2*, pages 114–âĂŞ125. Barlow, M. et al.
Heim, I. (2008). Features on bound pronouns. In Harbour, D. Adger, D. and Béjar, S., editors, *Phi theory: Phi-features across modules and interfaces*, pages 35–56.

Heim, I. and Kratzer, A. (1998). *Semantics in generative grammar*. Blackwell.
Heycock, C. and Zamparelli, R. (2005). Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural language semantics*, 13(3):201–270.
Hulsey, S. and Sauerland, U. (2006). Sorting out relative clauses. *Natural language semantics*, 14(2):111–137.
Ionin, T. and Matushansky, O. (2006). The composition of complex cardinals. *Journal of semantics*, 23(4):315–360.
Jackendoff, R. (1977). *X' syntax: A study of phrase structure*. Number 2 in Linguistic Inquiry Monographs. MIT Press: Cambridge, Mass.
Jacobson, P. (2004). Direct compositionality: Is there any reason why not. In *workshop in linguistics and philosophy, University of Michigan*.
Karttunen, L. and Peters, S. (1979). Conventional implicatures in Montague Grammar. In Oh, C.-K. and Dineen, D., editors, *Syntax and semantics 11: Presupposition*, pages 1–56. New York: Academic Press.
Kayne, R. S. (1989). Null subjects and clitic climbing. In Jaeggli, O. and Safir, K. J., editors, *The Null Subject Parameter*, pages 239–262. Kluwer Academic Publishers, Dordrecht, Holland.
Kiparsky, P. and Tonhauser, J. (2012). Semantics of inflection. In Maienborn, C. Heusinger, K. and Portner, P., editors, *Semantics: An International Handbook of Natural Language Meaning*, pages 2070–2097.
Kobele, G. M. (2006). *Generating Copies: An investigation into structural identity in language and grammar*. PhD thesis, University of California, Los Angeles, Los Angeles, CA.
Kratzer, A. (1998). More structural analogies between pronouns and tenses. In *Semantics and linguistic theory*, volume 8, pages 92–110.
Kratzer, A. (2008). On the plurality of verbs. In DÃűlling, J. Heyde-Zybatow, T. and Schäfer, M., editors, *Event structures in linguistic form and interpretation*, volume 5, pages 269–300.
Krifka, M. Common Nouns: A Contrastive Analysis of English and Chinese. In Carlson, G. and Pelletier, F., editors, *The Generic Book*, pages 398–411. Chicago: Chicago University Press.
Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. *Lexical matters*, 2953.
Kuno, S. (2004). Empathy and direct discourse perspectives. In *The handbook of pragmatics*, pages 315–343. Blackwell Publishing Ltd.
Landau, I. (2001). *Elements of control: Structure and meaning in infinitival constructions*, volume 51. Dodrecht: Kluwer.
Landman, F. (1996). Plurality. In Lappin, S., editor, *The Handbook of Contemporary Semantic Theory*.
Lasnik, H. (1998). Some reconstruction riddles. In *Proceedings of the 22nd Annual Penn Linguistics Colloquium*, volume 5, pages 83–98.
Lasnik, H. (1999). *Minimalist Analysis*. Blackwell Publishers, Oxford.
Lebeaux, D. (1988). *Language acquisition and the form of the grammar*. PhD thesis, University of Massachusetts, Amherst.
Lebeaux, D. (1998). Where does binding theory apply? *Technical report*.
Lechner, W. (2004). Extending and Reducing the MLC. In Stepanov, A. Fanselow, G. and Vogel, R., editors, *Minimality Effects in Syntax*, pages 205–240. Mouton de Gruyter: Berlin.
Longenbaugh, N. (2017). Permissive distributivity. *Unpublished ms, MIT*.
Longobardi, G. (2001). *The structure of DPs: Some principles, parameters, and problems*. Citeseer.
Marti, L. (2017). Inclusive plurals and the theory of number. *Ms*.
Matushansky, O. (2006). Head movement in linguistic theory. *Linguistic inquiry*, 37(1):69–109.
Matushansky, O. (2015). Against PredP. In *Proceedings of IATL*, volume 30. MITWPL.

McCarthy, C. (2004). Reconstruction and Condition C: Some unexpected symmetries. *McGill working papers in linguistics*, 18(2):25–44.
Müller, G. (2010). On deriving CED effects from the PIC. *Linguistic Inquiry*, 41(1):35–82.
Nelson, D. and Toivonen, I. (2000). *Counting and the grammar: Case and numerals in Inari Saami*. na.
Nunes, J. (2001). Sideward movement. *Linguistic Inquiry*, 32(2):303–344.
Partee, B. (1987). Noun phrase interpretation and type-shifting principles.
Pesetsky, D. and Torrego, E. (2006). Probes, goals and syntactic categories. In *Proceedings of the 7th annual Tokyo Conference on Psycholinguistics*, pages 25–60.
Pesetsky, D. M. (1982). *Paths and categories*. PhD thesis, Massachusetts Institute of Technology.
Postal, P. (1997). Strong crossover violations and binding principles. *Paper presented at the Eastern States Conference on Linguistics*.
Pylkkänen, L. and McElree, B. (2006). The syntax-semantics interface: On-line composition of sentence meaning. *Handbook of psycholinguistics*, 2:537–577.
Rezac, M. (2013). Case and licensing: evidence from ECM+DOC. *Linguistic Inquiry*, 44(2):299–319.
Ritter, E. (1991). Two functional categories in modern Hebrew noun phrases. *Syntax and Semantics*, 25:37–60.
Roberts, C. (1987). *Modal subordination, anaphora, and distributivity*. PhD thesis, University of Massachusetts, Amherst.
Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5.
Romero, M. (1998). *Focus and reconstruction effects in wh-phrases*. PhD thesis, University of Massachusetts at Amherst.
Sauerland, U. (1998). *The meaning of chains*. PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.
Sauerland, U. (2004). The interpretation of traces. *Natural Language Semantics*, 12:63–127.
Sauerland, U., Anderssen, J., and Yatsushiro, K. (2005). The plural is semantically unmarked. In Kepser, S. and Reis, M., editors, *Linguistic evidence*, pages 413–434. Mouton de Gruyter Berlin.
Schlenker, P. (2003). Indexicality, logophoricity, and plural pronouns.
Schlenker, P. (2007). Anti-dynamics: Presupposition projection without dynamic semantics. *Journal of Logic, Language and Information*, 16(3):325–356.
Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Sauerland, U. and Stateva, P., editors, *Presupposition and implicature in compositional semantics*, pages 243–281. Springer.
Sportiche, D. (2005). Cyclic np structure and the interpretation of traces. In *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*. Berlin/New York: Mouton de Gruyter.
Sportiche, D. (2016). Neglect (or doing away with Late merger and Countercyclicity).
Stowell, T. (1981). *Origins of Phrase Structure*. PhD thesis, Massachusetts Institute of Technology.
Svenonius, P. (1994). C-selection as feature-checking. *Studia Linguistica*, 48(2):133–155.
Takahashi, S. and Hulsey, S. (2009). Wholesale late merger: Beyond the A/A′ distinction. *Linguistic Inquiry*, 40(3):387–426.
Tunstall, S. (1998). *The interpretation of quantifiers: semantics and processing*. PhD thesis, University of Massachusetts-Amherst.
van Urk, C. (2015). *A uniform syntax for phrasal movement: A case study of Dinka Bor*. PhD thesis, Massachusetts Institute of Technology.
vanLehn, K. A. (1978). Determining the scope of english quantifiers. *Technical Report*.
von Stechow, A. (2003). Feature deletion under semantic binding: Tense, person, and mood

under verbal quantifiers. In *Proceedings of NELS*, volume 33, pages 379–404.
Watanabe, A. (2006). Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language & Linguistic Theory*, 24(1):241–306.
Winter, Y. (2001). *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. MIT press Cambridge, MA.
Wurmbrand, S. (2015). The cost of raising quantifiers.
Zamparelli, R. (1995). *Layers in the determiner phrase*. PhD thesis, University of Rochester.
Zimmermann, T. E. (2006). Monotonicity in opaque verbs. *Linguistics and Philosophy*, 29(6):715–761.
Zweig, E. (2009). Number-neutral bare plurals and the multiplicity implicature. *Linguistics and philosophy*, 32(4):353–407.