

Towards an etiology of outer indices*

Amy Rose Deal
University of California, Berkeley

1 Constraints on reflexive interpretation

One of the core pieces of conventional wisdom about reflexive pronouns is the idea that they must be interpreted as bound variables, where the variable binder is (an abstraction operator just below) the local antecedent. It is not sufficient, according to this wisdom, for reflexives to be coreferential with their antecedents; they must be semantically bound by them. A classic demonstration of this requirement features a focus-sensitive element like *only* added to an otherwise referential subject.

(1) Only Sally evaluated herself.

Sentence (1) says that apart from Sally, no one had the property of being an x such that x evaluated x . That's what we get by interpreting the reflexive via variable binding. The sentence can't mean that apart from Sally, no one had the property of being an x such that x evaluated Sally – a meaning that would have been expected if we allowed mere coreference in this type of example.¹

In terms of Heim's (1993) inner and outer indices, a requirement for semantic binding amounts to a requirement that the inner index on a reflexive pronoun match the outer index on its antecedent.²

(2) [Sally]₁ evaluated [herself]₁.

How is this requirement enforced? In his paper *In search of phases*, Johnson (2007) proposes that reflexive pronouns simply carry an inner-to-outer matching requirement, (3), as part of their semantic contribution. (The rest — in particular, the size restriction on P — arises from the phase theory.)

* Kyle Johnson inspired me to insist on the idea that syntactic and semantic investigations can't be carried out in isolation from one another. I thank him for that, and I hope that this small paper lives up to the example he has set! Thanks, too, to Emily Clem and Peter Jenks for comments on the manuscript, and to the editors of this volume for their patience.

1 For purposes of coreference, the term coreferent with the reflexive could be *Sally*, which is only a part of the surface subject. For that reason, this type of example raises well-known questions about binding and c-command. See, e.g., Heim 1993, Safir 2004: Ch. 4, and Buring 2005b for discussion. I won't have anything to say about this here.

2 I suppress an inner index on *Sally* here.

(3) *Principle A* (Johnson 2007: 155)

Some P containing an anaphor, α , must be interpreted as $\lambda x_1.[p \dots \alpha_1 \dots]$.

The statement in (3) improves on traditional formulations of Principle A in that it clarifies the role of inner and outer indices in reflexive binding. When we say that reflexives must be syntactically bound — coindexed under c -command with an antecedent in an A-position — we mean, if (3) is correct, that their inner index must match the outer index of such an antecedent.

What I want to show in this short paper is that (3)— and, in turn, the conventional wisdom it represents — is in fact not correct. Reflexives as a class do not require that their inner index match a locally c -commanding outer index, i.e., semantic binding. Reflexives in ECM subject position reveal a weaker requirement: their inner index may only match a locally c -commanding *inner* index, leading to the possibilities of coreference and co-binding alongside semantic binding. However we account for a semantic binding requirement in cases like (1), then, it should not go simply by way of the lexical representation of reflexive pronouns. Whether semantic binding is indeed required is instead to be determined by the pronoun's syntactic context.

To get this view off the ground, we will need to address some lingering questions about where inner and outer indices come from. I suggest, following much recent work, that indices are features whose distribution is regulated in part by Agree. These features may originate both on DPs and on heads in the clausal spine. I will implicate the former as the source for inner indices and the latter as the source for outer indices.

2 Broadening the constraints: ECM subject reflexives

Kiparsky (2002) and Safir (2004) have pointed out that ECM subject positions fall into a gap between syntactic and semantic binding. Reflexives in this position are not exempt in the sense of Pollard & Sag 1992; they require syntactic binding, and so cannot be discourse-governed. Semantic binding, however, is not required. An illustrative example is (4):

(4) Only Kyle considers himself to not be photogenic.

(4) is ambiguous in a way that (1) is not. One reading is the semantic binding reading: apart from Kyle, no one had the property of being an x such that x considered x to not be photogenic. The second reading reports that no one besides Kyle considers Kyle to not be photogenic. That is the coreference reading that is missing in (1). We see that difference in the two LFs shown below, each prefaced with its representation in terms of inner and outer indices.

Towards an etiology of outer indices

- (5) a. [Only Kyle]₁ considers [himself]₁ to not be photogenic.
Only Kyle [$\lambda x_1. x_1$ considers x_1 to not be photogenic]
b. [Only Kyle]₁ considers [himself]₁ to not be photogenic.
Only Kyle₁ [$\lambda x_2. x_2$ considers x_1 to not be photogenic]

There is a third possibility as well for the interpretation of ECM subject reflexives, nicely shown by an example from [Safir 2004](#). Consider his sentence (6).

- (6) Every Democratic candidate insisted that only she expected herself to win.

Safir observes that the sentence is ambiguous. It has two LFs which may be represented in our terms as (7b) and (8b), again prefaced with Heimian double-indexed representations. Safir's paraphrases are given in (7c) and (8c).

- (7) a. [Every Democratic candidate]₁ insisted that [only she]₁ expected [herself]₁ to win
b. Every Democratic candidate [$\lambda x_1. x_1$ insisted that only x_1 [$\lambda x_1. x_1$ expected x_1 to win]]
c. Each candidate insisted that no other candidate expected to win.
(8) a. [Every Democratic candidate]₁ insisted that [only she]₂ expected [herself]₁ to win
b. Every Democratic candidate [$\lambda x_1. x_1$ insisted that only x_1 [$\lambda x_2. x_2$ expected x_1 to win]]
c. Each candidate insisted that everyone else thought she would lose.

In (7), the reflexive is semantically bound by its local subject; there is local inner-to-outer index matching. In (8), the reflexive is semantically bound by the higher subject *every Democratic candidate*. It is merely co-bound with its local subject; both the reflexive and its local subject are bound by the higher subject semantically.

ECM subject reflexives may be semantically bound by, coreferential with, or co-bound with their local antecedents. What is in common across these options? There is a simple generalization about the LFs above:

- (9) *Inner Index Matching*

If A syntactically binds B, A and B must match in inner index.

That is, what these examples suggest is that we can understand the interpretation of syntactic binding in terms of inner indices, the very indices that determine how each DP is itself interpreted. It does not matter how abstractors are indexed (which is a matter of outer indices), and therefore, it does not matter whether there are

relationships of semantic binding. Semantic binding is one outcome among several that are compatible with Inner Index Matching.

Let us see how far this will get us. We should consider the further LFs possible for Safir's example (6) where the intermediate subject is not semantically bound by the highest one. In this case the pronoun *she* acquires its referent contextually; it is syntactically and semantically free. Two readings are possible when this is so.

- (10) a. [Every Democratic candidate]₁ insisted that [only she₂]₂ expected [herself₂] to win
 b. Every Democratic candidate [$\lambda x_1. x_1$ insisted that only x_2 [$\lambda x_2. x_2$ expected x_2 to win]]
 c. Each candidate insisted that only $g(2)$ expected to win.
- (11) a. [Every Democratic candidate]₁ insisted that [only she₂]₃ expected [herself₂] to win
 b. Every Democratic candidate [$\lambda x_1. x_1$ insisted that only x_2 [$\lambda x_3. x_3$ expected x_2 to win]]
 c. Each candidate insisted that only $g(2)$ expected $g(2)$ to win.

These possibilities are consistent with Inner Index Matching. And of course the reading is missing where the reflexive is entirely semantically unbound, but contextually mapped to a referent:

- (12) *Missing reading 1*
 a. [Every Democratic candidate]₁ insisted that [only she₁]₂ expected [herself₃] to win
 b. Every Democratic candidate [$\lambda x_1. x_1$ insisted that only x_1 [$\lambda x_2. x_2$ expected x_3 to win]]
 c. Each candidate insisted that everyone else thought $g(3)$ would lose.

A further missing reading is one where the reflexive is bound semantically by the matrix subject but does not share the (inner or outer) index of its local subject.

- (13) *Missing reading 2*
 a. [Every Democratic candidate]₁ insisted that [only she₂]₃ expected [herself₁] to win
 Every Democratic candidate [$\lambda x_1. x_1$ insisted that only x_2 [$\lambda x_3. x_3$ expected x_1 to win]]
 b. Each candidate insisted that only $g(2)$ expected that candidate to win.

Where things become more interesting is in the case of LFs which violate Inner Index Matching but are alphabetic variants of the examples with semantic binding, (7) or (10), e.g.:

- (14) a. [Every Democratic candidate]₁ insisted that [only she₁]₂ expected [herself₂] to win
b. Every Democratic candidate [$\lambda x_1. x_1$ insisted that only x_1 [$\lambda x_2. x_2$ expected x_2 to win]]
c. Each candidate insisted that no other candidate expected to win.

Here the reflexive is semantically bound by the abstractor just below its local antecedent. Any index could be substituted for the three occurrences of the index 2 in (14b), with no effect on the meaning. This makes it hard to judge on the basis of meaning alone whether the reflexive and its antecedent must match in inner index. Faced with semantically equivalent LF pairs like (7)/(14), then, we could maintain the pure Inner Index Matching generalization (thereby maintaining that (7) is an admissible LF but (14) is not) – after all, the only examples where we can really *tell* if Inner Index Matching holds are those where the reflexive is not (locally) semantically bound. Alternatively, we could move to a more permissive generalization which admits both (7) and (14):

(15) *Either Index Matching*

If A syntactically binds B, A and B must match in either inner or outer index.

The question will have to be decided on theoretical grounds. In the next section, I will propose a theory that makes Either Index Matching the slightly more natural of the two generalizations.

3 The road to LF

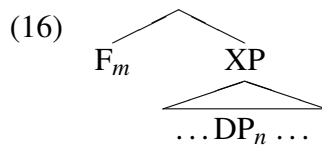
If Index Matching (Inner or Either) is the correct generalization, what kind of generalization is it? We could think about this type of pattern as an artifact of anaphor-antecedent agreement. This becomes natural if indices are represented in the syntax by index features, as a number of authors have recently concluded.³ Presumably, index features are features of D, like Case, which are not inherent in D elements, but which are obligatorily added to them as they are entered into a numeration.⁴ Syntactic binding certainly requires at least some feature matching between reflexive and

³ See, e.g., Rezac 2004, Hicks 2009, Kratzer 2009, and Moulton 2009.

⁴ Chomsky (1995: §4.2.2) calls these features ‘optional’, though that must be taken as a technical term. A D cannot be merged into a structure without a Case feature of some type (modulo default case in the sense of Schütze 2001); likewise, it cannot be merged without an index feature. It is merely the choice of one index feature over another that constitutes the option.

antecedent, and Index Matching can be thought of simply as a particular instance of this wider requirement. The question is now how we produce LFs that contain inner and outer indices for the antecedents of reflexives. Outer indices, in particular, must be scrutinized. If indices are ϕ -features (Rezac 2004: ch 3), indicating how a DP itself should be interpreted, how do outer indices fit in to the system?

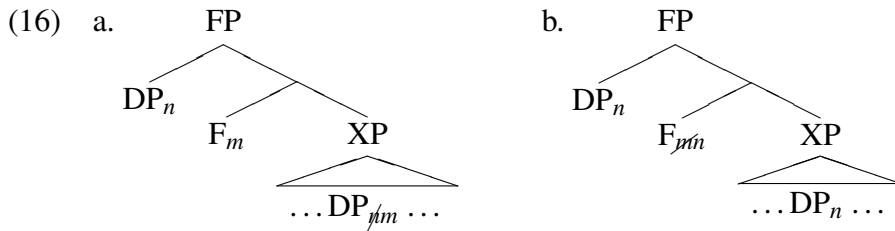
Let's assume with Heim & Kratzer (1998) that the lambda binders we have been positing arise in some sense from movement of the associated arguments. (Since all our examples involve subject binders, the movement in question can be ordinary A-movement to SpecIP.) Let's further suppose that movement is driven by functional heads which bear their own indices, added to them (as to Ds) as they enter a numeration. In the narrow syntax, then, (external) Merge builds structures like (16), where head F bears index feature m and DP bears possibly distinct index feature n .



The next step is Move, and following Chomsky (2000), that is a composite of two operations. One is Agree between F and the XP-internal instance of DP. The other is Merge of a new copy of DP in F's specifier. What, exactly, are the features that are subject to Agreement? Usually, in movement of subjects to SpecIP, it is ϕ -features that play the central role. If index features are ϕ -features, index feature agreement should take place in (16).

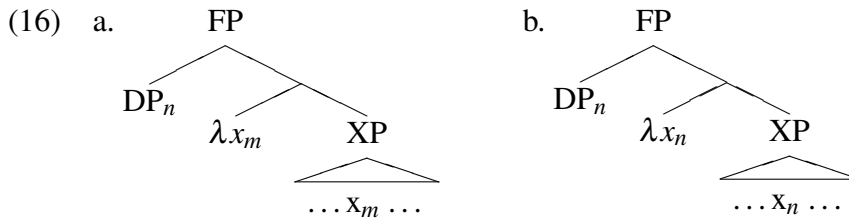
But wait. Both F and DP have index features that are valued. (If probes always have unvalued features, it must be other ϕ -features on F that trigger ϕ -Agree in this case.⁵) In terms of index features, starting from structure (16), the result of Agree and Merge could in principle be either (16a), where the index of F is shared with DP, or (16b), where the index of DP is shared with F. (A key assumption here is that sharing of an index feature overwrites the previous index feature associated to the relevant locus. For clarity, I represent the original indices with strikeouts below.)

⁵ That is to say that Agree for some ϕ feature [u G] brings back both [G] and an index feature [i] to the probe. This could be thought of as part of a larger generalization that Agree brings back (i.e., interacts with) the entire ϕ set whenever a particular ϕ feature forms the satisfaction condition for Agreement (Deal 2015).



I propose a simple outcome: the grammar makes both options available. Both structures serve as potential inputs to an algorithm for semantic interpretation.

The first step of this algorithm is to translate syntactic objects and their purely syntactic features into semantically interpretable objects. I'll use the name 'LF' for the result of this step. Given the role of functional heads in driving movement, it would be reasonable to suppose that the binders we have been positing are simply the LF translations of the functional heads involved in DP movement.⁶ The LFs corresponding to structures like (16) have translated the functional head as a binder and the lower copy of the moving DP as a variable (compare Fox 2002).



These are LF representations of the type we have been positing, simply in tree-structural form. By introducing functional heads and DPs with independent index features, and interpreting these as variables on DPs and as abstractors on functional heads, we have effectively produced the split between inner and outer indices. The proposed route to these representations shares with Buring 2005a the result that lambda binders are in principle indexed freely; they need not share the index of the DP in their specifier position. This is what makes it possible to tell apart semantic binding and co-binding in ECM examples like (6), and semantic binding and coreference in ECM examples like (4).

Overall, what this setup makes possible is a scenario where the in-situ copy of the local antecedent for an reflexive doesn't match the moved copy in index features. The source of the disjunctive requirement imposed by Either Index Matching is now apparent: the reflexive must agree in index features with its antecedent, but either the higher or lower copy thereof can be considered. Matching of the higher copy

⁶ Important antecedents here are Adger & Ramchand 2005 and Kratzer 2009, who argue that semantic binding may be established directly between bindees and binding functional heads, rather than DP "antecedents".

produces inner index matching. Matching of the lower copy produces inner-to-outer index matching. By contrast, a theory adopting purely Inner Index Matching would need to additionally require that only the higher (and less local) of the two copies be considered.

4 Constraining the system

If there has been progress thus far, it comes at the cost of a new problem of overgeneration. By replacing Johnson's (3) with Index Matching (whether Inner or Either), we have come to predict that a requirement for semantic binding should be absent not just for ECM subject reflexives, but also for reflexives in other syntactic positions. While this is correct for ECM structures like (4), it is incorrect for simple reflexive sentences like (1). What makes these two cases different?

What I will say here will be brief and will follow closely in the footsteps of previous work. What is needed is a principle along the lines of (16) (cp. Safir's (2004) Locally Reflexive Principle). This encodes both a preference for one type of LF over another, and a restriction on that preference's scope.

(16) *Preference for semantic binding*

Only semantic binding, and not coreference/cobinding, is possible between reflexive and antecedent, when the two are arguments of the same semantic predicate.

We should put both pieces of the proposal in (16) in perspective. The preference itself recalls a seminal idea from Reinhart 1983a,b which has informed much of the following work on Principle B (Grodzinsky & Reinhart 1993, Heim 1993, Büring 2005b, Roelofsen 2010). The central insight of that literature can be appreciated via the formulation from Grodzinsky & Reinhart 1993:

(17) *Rule I: Intrasentential Coreference*

NP A cannot corefer with NP B if replacing A with C, C a variable A-bound by B, yields an indistinguishable interpretation.

The preference for semantic binding posited in (16) is different from Rule I in two ways. First, it is different in lacking the caveat about indistinguishable interpretations. In (1), for instance, coreference and semantic binding yield distinguishable interpretations, but this does not improve the status of the LF with coreference.

- (1) Only Sally evaluated herself.
- a. ✓LF: Only Sally₁ [$\lambda x_1. x_1$ evaluated herself₁]
 - b. ✗LF: Only Sally₁ [$\lambda x_2. x_2$ evaluated herself₁]

Second, the preference for semantic binding (16) and Rule I are different in their domains of application. The preference for semantic binding applies only to bindees in ordinary object position, not ECM subject position. But Rule I considerations are visited equally on ordinary objects and on ECM subjects. The broad scope of Rule I makes sense in view of Reinhart's proposal that Rule I effects are in essence pragmatic and driven by norms of efficient conversation. Speakers and hearers in discourse consider entire sentences or sequences of sentences at a time, and draw conclusions from them. The narrow scope of the preference for binding in (16), by contrast, makes sense if this condition is in essence grammatical and driven by constraints on efficient computation. Chomsky (2001) conjectures that such constraints induce cyclicity. Grammatical mechanisms sequentially consider limited windows of structural representation. The generalization expressed in (16) accords with the intuition of Johnson (2007) that we might identify these windows of computation, or phases, by considering locality domains for the grammar of reflexives.

The ultimate result we arrive at should be distinguished from Johnson's, however, and from that of other authors working to understand locality in binding in terms of the phase theory (e.g., Lee-Schoenfeld 2008, Hicks 2009, Rooryck & Vanden Wyngaerd 2011, Despica 2015). We can now see that the syntax and semantics of reflexive binding are sensitive to subtly different domains of locality. Syntactic binding relations are established in domains which include both the matrix and embedded subjects in simple cases of ECM. It is this type of domain — a syntactic domain for syntactic binding — which is featured in going accounts of binding locality within a phase theory. Let's call this domain the *syntactic phase*. A different type of domain is involved in the grammaticized preference for binding. This preference winnows the range of LFs created by syntactic binding. The winnowing is internal to the semantic component and operates only on structures that are local in crucially semantic terms — the domain of a verb and its semantic arguments. We could call this the *semantic phase*. That name is meant to suggest a parallel between components of grammar in their use of limited windows of representation as a route to computational efficiency.

5 The end result

Are reflexive pronouns indeed always interpreted as semantically bound by their local antecedents, as the conventional wisdom would have it? No: that is what we learn from ECM subject reflexives. A theory that hopes to deal with these reflexives adequately must allow for the broader range of interpretive possibilities that comes about when the local antecedents of reflexive pronouns are associated with mismatching inner and outer binding indices. In this case, reflexives reveal the abil-

ity to be coreferential or co-bound with their local antecedents, rather than strictly semantically bound by them.

The tasks are then two-fold: to explain where the mismatching outer indices come from, and to explain why mismatches between inner and outer indices are not tolerated in *all* cases of reflexive binding. My suggestion on the first count goes by way of a crucially syntactic component for the binding theory. Index features are present in syntax and transferred by Agree between DPs and the functional heads that attract them. My suggestion on the second count puts together Reinhart's (1983a, 1983b) idea that semantic binding is preferred by linguistic systems with Johnson's (2007) view that we find locality domains in grammar by considering the domains that are relevant for the evaluation of interface constraints.

References

- Adger, David & Gillian Ramchand. 2005. Merge and move: *Wh*-dependencies revisited. *Linguistic Inquiry* 36(2). 161–193.
- Büring, Daniel. 2005a. *Binding theory*. Cambridge University Press.
- Büring, Daniel. 2005b. Bound to bind. *Linguistic Inquiry* 36(2). 259–274.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–156. Cambridge, Mass: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, Mass: MIT Press.
- Deal, Amy Rose. 2015. Interaction and satisfaction in ϕ -agreement. In Thuy Bui & Deniz Ozyildiz (eds.), *Proceedings of NELS 45*, 179–192. Amherst: GLSA.
- Despic, Miloje. 2015. Phases, reflexives and definiteness. *Syntax* 18(3). 201–234.
- Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33(1). 63–96.
- Grodzinsky, Yosef & Tanya Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24(1). 69–102.
- Heim, Irene. 1993. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. SFS-Report 7–93 Universität Tübingen Germany.
- Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.
- Hicks, Glyn. 2009. *The derivation of anaphoric relations*. John Benjamins.
- Johnson, Kyle. 2007. In search of phases. In S. Karimi, V. Samiiian & W. Wilkins (eds.), *Phrasal and clausal architecture*, 146–166. John Benjamins.
- Kiparsky, Paul. 2002. Disjoint reference and the typology of pronouns. In Ingrid Kaufmann & Barbara Stiebels (eds.), *More than words: a festschrift for Dieter*

- Wunderlich*, Akademie Verlag.
- Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2). 187–237.
- Lee-Schoenfeld, Vera. 2008. Binding, phases, and locality. *Syntax* 11(3). 281–298.
- Moulton, Keir. 2009. *Natural selection and the syntax of clausal complementation*: University of Massachusetts at Amherst dissertation.
- Pollard, Carl & Ivan A. Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23(2). 261–304.
- Reinhart, Tanya. 1983a. *Anaphora and semantic interpretation*. Chicago, Illinois: University of Chicago Press.
- Reinhart, Tanya. 1983b. Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy* 6. 47–88.
- Rezac, Milan. 2004. *Elements of cyclic syntax: Agree and Merge*: University of Toronto dissertation.
- Roelofsen, Floris. 2010. Condition B effects in two simple steps. *Natural Language Semantics* 18. 115–140.
- Rooryck, Johan & Guido Vanden Wyngaerd. 2011. *Dissolving binding theory*. Oxford University Press.
- Safir, Ken. 2004. *The syntax of anaphora*. Oxford University Press.
- Schütze, Carson T. 2001. On the nature of default case. *Syntax* 4(3). 205–238.