

Rule-I, intensional predicates, and children’s pronominal reference assignment*

Yosef Grodzinsky

*Edmond and Lily Safra Brain Research Center
The Hebrew University of Jerusalem, Israel
Institute for Neuroscience and Medicine (INM-1)
Forschungszentrum Jülich, Germany*

1 Introduction

The birthday of Kyle Johnson, my friend, colleague, and co-author of a (yet-to-be-published) paper on parsing, is an opportunity for reflection on thoughts past. During my long friendship with Kyle, the world of pronominal reference has seen debates and developments. Much of the little I understand of this world is due to Kyle’s patient and lucid explanations. It is thus with gratitude that I dedicate this squib to him.

My intention here is to briefly revisit the intra-sentential coreference rule, known as Rule I in its various incarnations (Reinhart 1983, Grodzinsky & Reinhart 1993, GR henceforth, Heim 1998, 2009, Fox 2000, Büring 2005). I will bring up data that are not typically considered in the context of syntactic binding, and interpret them using Heim’s perspective on Rule I. Then, I will show how this perspective can help to evaluate children’s notorious failure to assign the right referential value to pronouns in certain contexts at a certain age. I will use this perspective to argue that, contrary to GR’s account, children at age 4 are in full command of Rule I and its implementation, and are hampered by no processing deficiency in this respect. Children’s errors, I will argue, stem from an inability to recover from failed semantic composition.

* I’d like to thank Luka Crnić for sharp and constructive comments, and the students in my HUJI Spring 2017 seminar—Nitzan Kugler, Kedem Snir and Ella Tzuberi—for taking my false starts with stride. Thanks also go to the editors of this volume for their initiative and for inviting me to contribute. This research was made possible by grants from the Israel Science Foundation and from ELSC.

2 The current picture and how we came to have it

2.1 Counterexamples to Condition B

A syntactic binding theory in which locality is the divide between pronouns and reflexives (e.g., Chomsky 1981) is immediately confronted with counterexamples, most notably, instances in which a pronoun and a local antecedent have the same referential value:

- (1) *Local Coreference* — “*apparent violations of Condition B*”
- a. John and Mary have a lot in common. He loves her and *she* loves *her*.
(Evans 1980)
 - b. I dreamt I was Brigitte Bardot and *I* kissed *me*. (Lakoff 1972)
 - c. Q: Who is this woman?
A: She must be *Zelda*. *She* praises *her* to the skies. Only *Zelda* would do that. (Evans 1980)

A well-known account (Reinhart 1983, Grodzinsky & Reinhart 1993) proposes that when this shared referential value results in a special interpretation, it does not fall under the Binding Theory. Each of the above is said to meet this requirement.

2.2 Rule I

Reinhart builds on an idea of Bach & Partee (1980), who propose to handle strict/sloppy pronominal ambiguities in VP-ellipsis by distinct grammatical modules. A pronoun, on this view, may be either a bound variable (as in sloppy identity meanings of VP-ellipsis), or a referential pronoun whose assignment is invariant, and that has the same referential value as some antecedent (strict identity). Reinhart makes use of the distinction between bound and referential pronouns, and argues that phenomena like those in (1), henceforth *Local Coreference*, are due not to binding, but rather, to sameness of referential value of a pronoun and a local antecedent, which is licit in (1), but not elsewhere (2a):

- (2)
- a. **Mary* likes *her*
 - b. *Mary* likes *herself*

The need, then, is to distinguish coreference from variable binding on the one hand, and on the other hand, to regulate it so that cases like (2a) are ruled out. Thus a special rule is proposed, relying on the observation that pronominal coreference is blocked in environments where a pronoun that shares referential value with a local antecedent can be replaced by a bound one (2b). The rule expresses the intuition that

Local Coreference is possible only when it accomplishes something that binding cannot do, according to which:

- (3) *Rule I* (Grodzinsky & Reinhart 1993)
 NP A cannot corefer with NP B if replacing A with C, C a variable A-bound by B, yields an indistinguishable interpretation.

2.3 Heim’s refinement

The idea behind all versions of Rule I is that of alternative meaning: in GR’s version, if the coreferential meaning is “indistinguishable” from the bound one, then the former is blocked. However, the notion of distinguishability is vague. Heim (1998, 2009) makes this notion precise. The referential value that the “offending” pronoun is assigned is determined by the utterance context. This enables its value to be the same as that of a local antecedent, but in the absence of binding. Covaluation, the relation between two coreferential pronouns that are not bound, is defined thus:

- (4) *Covaluation* (Heim 2009)
 Let α and β be occurrences of DPs of type e in an LF ϕ , and let C be a subjective utterance context. Then β is covalued with α in ϕ and C iff for all $\langle w, g \rangle \in C$ and all $g' \supseteq g$, $\llbracket \phi \rrbracket^{w, g'} = \llbracket \phi^{\alpha/\beta} \rrbracket^{w, g'}$, where $\phi^{\alpha/\beta}$ is the result of replacing β by α in ϕ .

Heim assumes a version of Rule I that is has the same structure as GR’s, except that her system, which ties the referential value of the pronoun to the utterance context, allows for a meaning difference between the two alternatives — the meaning of an utterance with two covalued pronouns may be distinct from the meaning of that utterance with a bound alternative. Thus, by Heim’s Rule I, Local Coreference is possible if covaluation produces a meaning that is distinct from binding. But why is that so in (1)? Heim’s answer: the referential value of the pronoun in these cases is distinct from the bound one because the pronoun is an individual concept, and as such, covaluation with a local antecedent produces a meaning that is distinct from a bound one. In (1a), Mary loves the person that she is.

3 Further predictions

Heim’s claims can be put to the test (along lines suggested in Grodzinsky 2007 and Grodzinsky & Sharvit 2008). First, if a bound and a covalued meaning are distinct, then coordinating one with the negation of the other would not lead to a contradiction. Thus, given S, a sentence with a covalued pronoun, we construct S’ by replacing that pronoun with a bound one, and check whether S&¬S’ is contradic-

tory. Mary may love the person that she is in (1a), but that doesn't mean that she is in love with herself. By this logic, these elaborations of (1) should be acceptable:

- (5)
- a. John and Mary have a lot in common. [He adores her] and [_S she adores her], although [_{S'} she does not (necessarily) adore herself].
 - b. I dreamt I was Brigitte Bardot and [_S I kissed me], but [_{S'} I didn't kiss myself].
 - c. Who is this woman? She must be Zelda. [_S She praises her to the skies], but I wouldn't say [_{S'} She praises herself to the skies].

To convince a skeptical reader that (5) is acceptable, let's set up a context against which (5a) needs to be evaluated: modest John and modest Mary are watching a movie. They both adore the main actress, but being at a distance from the screen, they fail to realize that the actress is actually Mary. Thus modest Mary, who doesn't adore herself, *does* adore that woman on the distant screen, who, unbeknownst to her, happens to be she herself. This situation makes (5a) true. Similar contexts can be set up for the other cases in (5).

Descriptively, it is easy to see that the pronoun in S is has a *de re* reading, whereas the one in S' is *de se* (Lewis 1979), and that $\llbracket S_{de\ se} \rrbracket \neq \llbracket S'_{de\ re} \rrbracket$, and thus for (5), it holds that $\llbracket S_{de\ se} \rrbracket \& \neg \llbracket S'_{de\ re} \rrbracket \neq \perp$.

The *de re/de se* distinction is typically attributed to *believe*-type predicates. However, it can also be obtained with intensional predicates that combine with an individual, not a proposition, namely, with an individual concept (Moltmann 1997, Grodzinsky & Sharvit 2008). As it turns out, this is exactly the type of predicate in (1):

- (6)
- a. In May 2017, John admired the president of the United States, but he did not admire Trump.
 - b. John praised the president of the United States, but he did not praise Trump.
 - c. John dreamt about the president of the United States, but not about Trump.

The president who John admired at the date mentioned in (6a) could have been Obama (who John happens to falsely believe continued for a third term), and not Trump (who he despises in May 2017). Likewise, he may have praised the U.S. president, whoever this person may have been (6b), or dreamt about that personality (6c), but not about Trump. Yet as there is a single individual who is the U.S. president at any given time, John could not have met two different individuals who were president at that date. Replacing the intensional predicate with an actional one therefore results in a contradiction:

- (7) #In May 2017, John met the president of the United States, but he did not meet Trump.

These observations can be summarized thus: (a) Rule I only permits Local Coreference where covaluation produces a meaning that is different from binding; (b) this difference is contingent on there being an individual concept in object position. If these observations are valid, we expect actional predicates to block Local Coreference. With such predicates, the extension of the pronoun should be the same in all possible worlds, and covaluation should not produce a reading that is distinct from binding. This expectation is borne out, as noted in [Grodzinsky 2007](#). Replacing the predicates in (1) with actional ones results in blocked Local Coreference:

- (8) a. John and Mary do the same thing on Tuesdays. #He lends *her* money and *she* lends *her* money.
b. #I said I was Brigitte Bardot and *I* kissed *me*.
c. Q: Who is this woman? A: She must be Zelda. #*She* just gave *her* money.

Heim's account predicts the contrast between (1) and (8). Moreover, it opens the way for the construction of contexts that are contrived so that an individual concept object can sneak even into a sentence with an actional predicate:

- (9) a. Walking into the crowded stadium, John and Mary stumbled upon the same object. It was so crowded that they didn't even realize that he kicked her and *she* kicked *her*.
b. Dick and Jane each have Jane's home number on the top of their speed dials. When they came to the party, they both left their cell phones on the couch. Oddly enough, they then sat down with equal clumsiness: they both sat on their phones inadvertently, and hit the speed-dial in the same way. Thus, he phoned her and *she* phoned *her* (i.e., her home phone).

So Heim's construal of Rule I (4) is vindicated, with one important modification: covaluation is possible only if one of the covalued DPs is type $\langle s, e \rangle$ (rather than type e). With this conclusion in pocket, let me examine whether this view of the rule can help to elucidate the old acquisition puzzle.

4 Rule I and the acquisition puzzle

4.1 The experimental record

Four-year-old children have repeatedly evinced non-adult like performance in a Truth-Value Judgment Task (TVJT) where the assignment of a referential value to a pronoun is at issue (most notably Chien & Wexler 1990; see Hartman et al. 2013 for a recent review of the experimental record). The most stable performance pattern is summarized in (10). Each test sentence is typically preceded by a fairly plain context, in which the characters involved are introduced, and accompanied by an image ($A \circ B$ stands for an image with a reflexive action; $A \rightarrow B$ stand for an image with a transitive action). Children are asked to answer a question, or determine the truth-value of a declarative sentence:

(10)	Context and sentence	Image	
		<i>Match</i> % correct	<i>Mismatch</i> % correct
	<i>This is A. This is B...</i>		
a.	Is A touching herself?	~90 $A \circ B$	~90 $A \rightarrow B$
b.	Is A touching her?	~90 $A \rightarrow B$	~50 $A \circ B$
c.	Is every A touching herself?	~80 $A \circ B$	~80 $A \rightarrow B$
d.	Is every A touching her?	~80 $A \rightarrow B$	~80 $A \circ B$

At the relevant age, children master conditions A and B, as demonstrated by their high performance levels in (10a,c,d), and replicated multiple times. In addition, they correctly confirm the truth of the Match (M) in (10b) at high levels. But why do they fail to identify the falsehood of the MisMatch (MM) in (10b), and moreover, perform at chance-level in this condition? Whatever their deficiency is, it cannot stem from an absence of a constraint on Local Coreference, because such deficiency would imply that they lack the means to block Local Coreference, and therefore erroneously answer in the affirmative. But this answer mode would lead to below-chance level performance, rather than chance — the level observed in the failed condition.

4.2 A Rule I-based account

In GR, we proposed a processing account for this sticking data point. We observed that Rule I divides the experimental pie correctly: cells (10a,c,d) are only governed by the binding theory, as in all of them, the pronoun must be a bound variable (either by it being a reflexive, or by a quantificational antecedent), The M part of (10b) requires neither the binding theory no Rule I. Indeed, children are well above

chance in all of these conditions. The only condition leading to failure is the MM part of (10b) — the one in which both local binding and Local Coreference need to be rejected.

We thus noted that the distinctions within the children's behavior, then, are correctly made by Rule I. But why are they *at chance level* on the failed condition? If they don't master Rule I, they should accept (10b), and err systematically, namely be at 0% correct. To explain their chance behavior, we took Rule I to be an algorithm, and relegated the deficit to a lack of processing resources that its implementation requires: to rule out identity in referential value of the pronoun and the local antecedent A in (10b), this algorithm must construct an alternative representation of the sentence by replacing the pronoun with a bound one, and then compare this alternative to the original representation. Local Coreference would be deemed as licit depending on the results of the comparison. Yet for the Replace-and-Compare algorithm to be implemented, a processing resource is required, which we argued was lacking in children at the relevant age. Therefore, when asked to implement Rule I, their system collapses, and the result is guessing. The same reasoning holds for Heim's perspective.

4.3 Additional results

A recent study by [Hartman et al. \(2013\)](#) contrasted the failed condition (11a) with one in which the pronoun is phonologically reduced — an option available in English (11b):

- (11) a. I think Elmo painted him. (Full Pronoun)
 b. I think Elmo painted 'm. (Reduced Pronoun)

Hartman et al., whose stimuli were preceded by rather elaborate contexts (thanks to methodological suggestions by [Conroy et al. 2009](#)), successfully replicated [Chien & Wexler's \(1990\)](#) result — chance performance (52.8% correct) in the MM part of (11a). For (11b), in which the pronoun was phonologically reduced, they found elevated performance level — the children rejected Local Coreference 80% of the time. They concluded that children treat reduced pronouns in English as clitics, similar to the way pronominal clitics are treated by young speakers of many languages (e.g., [McKee 1992](#), [Escobar & Gavarró 1999](#)). For example, the Italian *Gianni lo asciuga* 'Gianni dries him' was used by McKee in an analogous experiment with a comparable age group in Italy, where children performed correctly in the MM condition, that is, gave high (correct) rejection rates. Thus pronoun-containing sentences with reduced English pronouns, as well as clitics in many languages, do not lead to failure. We can safely conclude that these sentences do not require the Replace-and-Compare algorithm, or else the processor would collapse, and lead to

chance performance, on a par with the documented failure on (11a). Thus Hartman et al. show that the deficit is more restricted than previously supposed.

5 A composition failure account

Looking now at the experimental record, we find two classes of pronoun-containing constructions that indicate that children successfully eschew Rule I when asked to evaluate Local Coreference: (a) They correctly reject it in sentences where the potential local antecedent (subject) is quantificational (10d); (b) They also correctly reject sentences in which the pronoun to be covalued with a local antecedent (object) is a clitic (11b). In both instances, a bound variable analysis is forced, and the pronoun cannot be a candidate for covaluation. Children's successful rejection of Local Coreference leads to a safe conclusion: in these instances, their system does not even attempt to implement Rule I's Replace-and-Compare algorithm. They only fail when a full pronoun (that may be covalued if analyzed as an individual concept) composes with an actional predicate (that does not combine with an individual concept), and the resulting meaning combines with a referential subject.

Whence this pattern? In the successful conditions, either the pronoun is a clitic, or the local antecedent is quantificational. If children use these facts to stop the Replace-and-Compare algorithm from applying, why does their system collapse when an actional predicate must compose with a full pronoun (and the resulting meaning combines with a referential subject)? Why can't they use the actional predicate as a clue that covaluation should not even be considered (especially in the vanilla context that Chien & Wexler used)? Here is my tentative answer, presented in four steps:

- i. children know Rule I and suffer no processing limitation with respect to it;
- ii. the Replace-and-Compare algorithm ignores the predicate. It only looks at categories that may be referential. If one of them must be a binder or a bindee, the algorithm is blocked from applying;
- iii. in the problematic (10b) and (11a) that lead to chance performance, the referential subject and the full pronoun object are good candidates for covaluation. The pronoun can be analyzed as an individual concept, type $\langle s, e \rangle$; this brings about a composition problem, as the actional predicate, being transitive (type $\langle e, \langle e, t \rangle \rangle$) fails to compose with this type;
- iv. failing to construct a coherent meaning, the children are reduced to guessing.

My conclusion: at age 4, Children are in full command of Rule I and its implementation, and are hampered by no processing deficiency in this respect. When a pronoun has a phonological shape that makes it suitable to covaluation, it gets the appropriate type, and that leads to a composition failure. Yet, while adults successfully recover from composition failures, children do not, at least in cases of the sort described above. Their failure to recover, for all I know, may not be linguistic, but may be due to their general lack of experience with test situations (cf. Crain et al. 1996).

This analysis carries several predictions. Some of these are too subtle to be elaborated on in this brief chapter, but the main point is this: once intensional predicates replace the actional ones, children's performance should soar. To test this prediction is not an easy task. We are now getting ready to do that, and with any luck, we'll be able to report our results by the time we are called to contribute to Kyle's next festschrift.

Happy birthday, Kyle!

References

- Bach, Emmon & Barbara H. Partee. 1980. Anaphora and Semantic Structure. In Jody Kreiman & Almerindo E. Ojeda (eds.), *Papers from the Parasession on Pronouns and Anaphora*, 1–28. Chicago Illinois: University of Chicago.
- Büring, Daniel. 2005. *Binding Theory*. Cambridge: Cambridge University Press.
- Chien, Yu-Chin & Kenneth Wexler. 1990. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language acquisition* 1(3). 225–295.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.
- Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz & Colin Phillips. 2009. Equal treatment for all antecedents: How children succeed with principle B. *Linguistic Inquiry* 40(3). 446–486.
- Crain, Stephen, Rosalind Thornton, Carole Boster, Laura Conway, Diane Lilo-Martin & Elaine Woodams. 1996. Quantification without qualification. *Language Acquisition* 5(2). 83–153.
- Escobar, Linda & Anna Gavarró. 1999. The acquisition of Catalan clitics and its implications for complex verb structure. *Report de Recerca*, Rep. No. GGT-99-3, Grup de Gramàtica Teòrica, Universitat Autònoma de Barcelona.
- Evans, Gareth. 1980. Pronouns. *Linguistic Inquiry* 11(2). 337–362.
- Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, Massachusetts: MIT Press.
- Grodzinsky, Yosef. 2007. Coreference and Self-ascription. Ms, McGill University.

- Grodzinsky, Yosef & Tanya Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24(1). 69–102.
- Grodzinsky, Yosef & Yael Sharvit. 2008. Coreference and Self-ascription. Ms, McGill University and University of Connecticut.
- Hartman, Jeremy, Yasutada Sudo & Ken Wexler. 2013. Principle B and Phonologically Reduced Pronouns in Child English. Ms. UMass, Amherst, Institut Nicod and MIT.
- Heim, Irene. 1998. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In Orin Percus & Uli Sauerland (eds.), *The Interpretive Tract (MITWPL 25)*, 205–246. Cambridge, Massachusetts: MIT.
- Heim, Irene. 2009. Forks in the Road to Rule I. In Muhammed Abdurrahman, Anisa Schardl & Martin Walkow (eds.), *Proceedings of Thirty-Eight Annual meeting of the North East Linguistic Society*, 339–358.
- Lakoff, George. 1972. Linguistics and Natural Logic. In Donald Davidson & Gilbert H. Harman (eds.), *Semantics of Natural Language*, 545–665. Dordrecht: Reidel.
- Lewis, David. 1979. Attitudes de dicto and de se. *Philosophical Review* 88. 513–543.
- McKee, Cecile. 1992. A comparison of pronouns and anaphors in Italian and English acquisition. *Language Acquisition* 2(1). 21–54.
- Moltmann, Friederike. 1997. Intensional verbs and quantifiers. *Natural Language Semantics* 5(1). 1–52.
- Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. Chicago, Illinois: University of Chicago Press.