THE REST OF THE BINDING THEORY
Karin Golde

YY Software Corporation

1. Introduction

It has been acknowledged for some time that what is frequently referred to as the "reflexive" pronoun in English (*himself*, *herself*, etc.) does not always serve a reflexive function. That is, its interpretation may depend on another argument of the same head, as in (1a), or it may not, as in (1b).

(1)　　a. John$_i$ likes himself$_i$.
　　　　b. The image of herself$_i$ in a new apartment drove Mary$_i$ to work harder.

The goal of this paper is to build a single multi-clausal constraint which will license these and other uses of these pronominal forms.[1] Therefore they will henceforth be referred to as *self-pronouns* (spros), to make clear that it is not just the reflexive function which is at issue. In addition, this paper will attempt to account for the restrictions on the reference of personal pronouns (ppros) like *him*, *her*, *it*.

Most previous binding theories accounting for pronominal reference tend to focus either on the syntactic, semantic, or discourse related factors involved, either reducing all effects to one of these three levels, or treating discourse effects as a separate phenomenon. However, I will present evidence for the necessity of ordered constraints on the distribution of spros and ppros at the syntactic, semantic, and discourse levels.

2. Syntactic constraints

As shown in (1a), repeated below, the spro may have a coargument at the syntactic level. However, if this is the case, then the spro must be coindexed with another argument of the same head; hence the unacceptability of (2b).

(2)　　a. John$_i$ likes himself$_i$.
　　　　b. *John$_i$ knows that Mary likes himself$_i$.

This pattern also holds for cases where the spro is an argument of a noun, as in the following examples.

(3)　　a. Mary's$_i$ image of herself$_i$ has no connection to reality.
　　　　b. *John$_i$ knows that Mary's image of himself$_i$ has no connection to reality.

In Pollard and Sag (1994), these facts about spros are captured by Principle A, a constraint which depends on the syntactic concept of obliqueness.

(4)　　Principle A: A locally o-commanded spro is locally o-bound.[2]

---

[1] Other functions include the adverbial emphatic (*the king did it himself*) and the adnominal emphatic (*the king himself did it*). See Golde (1999) for an analysis of these and their relation to the pronominal spro.
[2] Pollard and Sag use the more traditional term "anaphor" rather than "spro".

In (2) and (3) the spro is locally o-commanded, since it has a less oblique coargument, but only in the (a) examples is it locally o-bound, and thus licensed by Principle A.

Principle B accounts for constraints on the reference of ppros, which may not be coindexed with a less oblique coargument. This is formulated by Pollard and Sag as follows:

(5)   Principle B: A ppro is locally o-free.

This accounts for the fact that a ppro may not be coindexed with a less oblique coargument, as in the following example:

(6)   *John$_i$ likes him$_i$.

3. Semantic constraints

Pollard and Sag do not attempt to formally extend Principle A to account for those cases where the spro is locally o-free. This paper picks up where they leave off, adding two additional levels to Principle A to account for semantic and discourse related factors.

First, the semantic effects can be observed when the spro is the object of a relational noun, a noun which is like a verb in that it has arguments to which it assigns thematic roles. For example, the noun *note* must have a writer, and usually has an addressee, while *faith* has an experiencer and a theme.

(7)   Jill kept notes to herself on her bulletin board. [*Notes must be written by referent of spro*]
(8)   John believes that unwavering faith in himself will ensure his success. [*Referent of spro must be the one who has faith*]

As these examples show, when the object of a relational noun is an spro, then the higher unexpressed argument must be interpreted as coreferential.
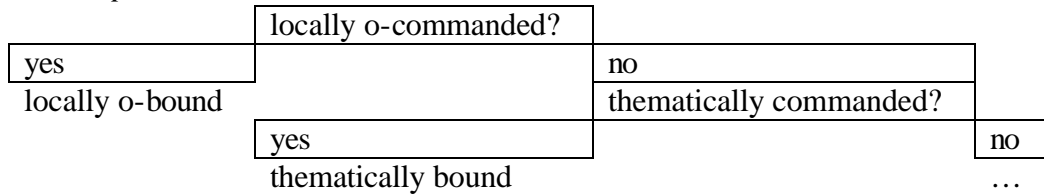
Williams (1994) takes such data as evidence that all anaphoric binding takes place at the thematic level. However, the following examples show that this cannot be right.

(9)   a. John$_i$ believes himself$_i$ to be a shoo-in this election.
        b. *John$_i$ believes him$_i$ to be a shoo-in this election.

With subject-to-object raising verbs like *believe*, the raised object is not assigned a thematic role by the higher verb, yet syntactically it acts as the verb's object. Unless spros can be bound at a syntactic level, this type of data is unaccounted for.

Thus I retain Pollard and Sag's Principle A, but add another clause to account for the data in (7) and (8).

(10)   Principle A (first revision)
       Is the spro...

	locally o-commanded?		
yes		no	
locally o-bound		thematically commanded?	
	yes		no
	thematically bound		…

Now even if the spro is locally o-free, Principle A checks to see if there is a "higher" theta role commanding it; if so, then it must be bound by that theta role.

Thematic command (th-command) is defined in terms of th-outranking, a relation among the roles in a given nucleus.

(11)   For θ_1 and θ_2 roles of a *psoa* σ, θ_1 TH-OUTRANKS θ_2 iff θ_1 has more Proto-Agent entailments than θ_2, and/or fewer Proto-Patient entailments than θ_2.

The order imposed on the thematic roles of a given head is based on the number of Proto-Agent and Proto-Patient entailments satisfied by each role, as put forth in Dowty (1991). The Proto-Agent entailments include properties typical of agents, such as volitional involvement in the event or state, sentience and/or perception, and causing an event or change of state in another participant. The Proto-Patient entailments include properties such as undergoing a change of state and being causally affected by another participant.

By defining the thematic hierarchy in terms of degree of agentivity and patienthood, we are freed from having to use the standard inventory of thematic roles, such as Agent, Patient, Goal, etc. As Dowty demonstrates, such a list is not adequate for capturing the fine-grained differences among roles. However, I will continue to use terms like "agent" informally where the intended meaning is clear.

We can then define th-command as follows:

(12)   For X a valent of a word *w* with CONT|NUCL σ, X is TH-COMMANDED iff there exist roles θ_1 and θ_2 of σ such that
       (i) θ_1 th-outranks θ_2, and
       (ii) the value of θ_2 is structure shared with X's index.

Finally, th-binding is defined in terms of th-command:

(13)   X is TH-BOUND iff there exist roles θ_1 and θ_2 of a soa σ such that
       (i) θ_1 th-commands θ_2, and
       (ii) the values of θ_1 and θ_2 are structure shared with X's index.

Returning to (7) and (8) above, we can see that the spro is th-commanded by the noun's higher ranking theta role. Principle A correctly predicts that it must be th-bound; that is, the noun's roles must be interpreted as being coindexed.

Note also that the two clauses of Principle A are crucially ordered with respect to each other. Consider the following contrast between the determiner and possessive pronoun:

(14)   Mary often wrote herself notes on things that needed to be done around the office. She passed some of them on to John, who posted the/*his notes to herself on a bulletin board.

Given the context, the notes are interpreted as being written by Mary to Mary. If we started by checking the second clause of Principle A, then both versions of the second sentence would be ruled in, since the spro is th-commanded and th-bound. However, as it stands Principle A rules out *his notes to herself*, because the spro is locally o-commanded, but not locally o-bound.

Finally, Principle B is also revised to account for semantic effects on ppros.

(15)   Principle B: (final version)
         A ppro is locally o-free and th-free.

Thus if the relational noun's object is a ppro rather than an spro, Principle B correctly predicts that the unexpressed agentive role cannot be coreferential with the ppro.

(16)   *Mary$_i$ put notes to her$_i$ in a drawer. [writer = addressee]

Mary may be interpreted as the writer of some of these notes but not all of them. If she were the sole author, then both the WRITER and ADDRESSEE roles would be assigned to Mary's INDEX value, and the ppro would be ruled out by Principle B for being th-bound.

4. Further evidence for the semantic component

The behavior of pronominal objects of locative PPs provides intriguing support for the current approach. At least since Jackendoff (1972) and Cantrall (1974), it has been observed that both an spro and a ppro are possible in this position when coindexed with the verb's subject.

(17)   a. John$_i$ put the candles around him$_i$/himself$_i$.
         b. Mary$_i$ spilled the juice on her$_i$/herself$_i$.

Kuno (1987) predicts that it is whether the pronoun's referent is a "target" of the action, in the sense of being physically and/or intentionally targeted, which determines whether an spro or ppro is used. This claim prompted a set of experiments detailed in Golde (1999), which suggest that the more the object's referent is physically involved in the action, the more likely speakers are to prefer to express the object as an spro rather than a ppro. Thus the ppro is preferred in an example like (17a), where the candles are not physically affecting John, while the spro is preferred in (17b), where Mary is altered to a degree by being soaked by the juice.

According to Dowty's set of Proto-Patient entailments, an argument which is causally affected and/or undergoes a change of state is more likely to be expressed as an object. In the present situation, it seems that the more patient-like an argument is, the more likely it is to be treated as a direct argument of the verb, where it is locally th-bound by the verb's

subject, and hence licensed as an spro. In a case like (17a), where the pronoun's referent is not physically affected, the pronoun is treated simply as an argument of the preposition, and hence is not one of the verb's thematic roles. Since it is th-free with respect to the subject, it is licensed as a ppro by Principle B.

While it is not clear at present how to formalize this insight, the phenomenon provides clues as to how thematic argument structure should be approached. Namely, it appears to be fluid to some degree, allowing for variation based on speakers' understanding of the real-world relationships among the participants in a given event.

5. Discourse constraints

Perhaps the best-known effect on the acceptability of non-reflexive spros is that of point of view, referred to almost synonymously as involving "logophoricity" or "empathy." For example, in the following sentences (based on ones constructed by Pollard and Sag), the spro is acceptable only when it refers to the person whose point of view is being represented.

(18)   a. $John_i$ was getting tired of Mary's dirty campaign tactics. He was especially upset when she had a doctored photo of him_i/$himself_i$ printed in the paper.
    b. Mary was going to get even with John. She was sure that she could regain the upper hand by getting a doctored photo of him_i/*$himself_i$ printed in the paper.

This issue is too complex to do justice here; instead, I will simply show that however it is characterized, this effect may be captured as the last clause of Principle A, as in the following formulation.

(19)   Principle A: (final version)
    Is the spro…

	locally o-commanded?		
yes		no	
locally o-bound		thematically commanded?	
	yes		no
	thematically bound		coreferential with the SC

I will use the term "subject of consciousness" (SC) as it is used by Zribi-Hertz (1989), to refer to the person whose point of view is being taken.

In (18), the pronoun is the object of a picture noun, one which heads an NP denoting an image or aspect of an entity. As such, it is distinguished from a relational noun, which heads an NP denoting a state or event, and generally has an agent-like thematic role. Being locally o-free and th-free, the spro in (18a) is required only to be coreferential with the SC.

However, if the spro is the object of a relational noun, then it is th-commanded, and must be th-bound.

(20)   a. $John_i$ was getting tired of Mary's dirty campaign tactics. He was especially upset by the doctored photo of him_i/$himself_i$ she had printed in the paper.

b. John$_i$ was getting tired of Mary's dirty campaign tactics. He was especially upset by the open letter to him$_i$/*himself$_i$ she had printed in the paper.

In (20b), *letter* is not a picture noun, and so the use of the spro implies that coreference is intended with the writer of the letter, Mary. In other words, the spro is th-commanded by the WRITER role without being th-bound, so it is ruled out by Principle A.

This shows that the last two clauses of Principle A are also crucially ordered. If coreference with the SC were checked before th-command, the spro in (20b) would be ruled in incorrectly. Thus we have evidence for constraints on spros at three levels (syntactic, semantic, and discourse), as well as evidence that these constraints are crucially ordered with respect to one another.

References

Cantrall, William. 1974. *Viewpoint, Reflexives, and the Nature of Noun Phrases*. The Hague: Mouton.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67:547-619.

Golde, Karin. 1999. Binding Theory and Beyond: An Investigation into the English Pronominal System. Doctoral dissertation, Ohio State University.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Kuno, Susumu. 1987. *Functional Syntax*. Chicago: University of Chicago Press.

Pollard, Carl, and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Williams, Edwin. 1994. *Thematic Structure in Syntax*. Cambridge: MIT Press.

Zribi-Hertz, Anne. 1989. A-type binding and narrative point of view. *Language* 65:695-727.