LINGUIST on-line conference on
Geometric and Thematic Structure in Binding
October, 1996.

Based on work in progress with Martin Everaert and Eric Reuland.

DRAVIDIAN ANAPHORA and implications for emphatic anaphors.

Tanya Reinhart

Dravidian languages (particularly Malayalam) were brought up in this conference as posing problems for the reflexivity approach of R&R (by Lidz and by Hamilton). In fact, anaphora in these languages provides direct evidence for that approach. It is also independently interesting, since it differs in some respects than the better studied anaphora systems, so I would like to survey these languages here in some detail. At the same time, one of the anaphora strategies found in these languages may shed some light on the problem of emphatic anaphors across languages.

Dravidian languages:
Kannada, (K) Malayalam, (M) Telugu, Tamil. SOV order.

1. inventory:

Here is first a quick survey of the basic anaphora inventory these languages all have in common. (I turn directly to where they differ.) I will use English words with Dravidian word order. Most of the original examples can be found in Jayaseelan's papers. Some are p.c.

-All have a pronoun (M: awan) and a SE anaphor (M: taan, similar form in the other languages). No SELF anaphor as such.

-The distribution of the SE anaphor in the local context, is the same as observed in R&R for Dutch: SE obeys condition B, as in (1a) (pointed out in great detail in Jayaseelan (e.g. 1995). The contrast observed in Dutch between object (1a) and ECM subject (1b) is found here as wall.

1) a) *Ram taan sees (Ram sees SE)
   b) Ram [taan fool] considers (Ram considers SE fool)
   c) *Ram [awan fool] considers. (Ram considers him fool)

Pronouns obey the chain condition, hence the contrast between (1b) and (1c). The facts of (1) are confirmed in all Dravidian languages, though they differ in the reflexivization system.

-The distribution of SE in the long distance context is similar to Chinese (which differs from Scandinavian, on points b, c):

-a. Subject orientation and C-command obligatory.
-b. Occurs freely also in subordinate clauses.
-c. Blocking effects (matching features), like in Chinese.

-Full names: No condition C-effects with two full names. Yes: condition B-effects. (This is similar to what Lasnik found in Vietnamese and Thai. I believe that the same is true for English as well. This is an issue of rule I, of Grodzinsky and Reinhart, and irrelevant for the present discussion).

Further notes on SE in the local context:
The findings in (1) are in conflict with what Hamilton says about Malayalam (quoting Mohanan 1982)). But there is ample evidence that the relative judgment in (1), repeated below, is found not only in all Dravidian languages, but also in Hindi. According to my p.c. checking of Korean, the same is found also there.

1) a) *Ram taan sees (Ram sees SE)
a') *Ram awan sees (Ram sees him)
b) Ram [taan fool] considers (Ram considers SE fool)
c) *Ram [awan fool] considers. (Ram considers him fool)

In evaluating conflicting reports of the judgments, the crucial point to consider is that also in Scandinavian, condition B violations are always weak, relative to the chain condition violation with pronouns. So if we compare a sentence with SE, like (1a) to the same sentence with a pronoun, as in (1a'), (1a) is always much better. In R&R's system this difference follows, first, since the pronoun in these contexts violates both condition B and Chain, and, more importantly, because the chain condition is viewed as a stronger -core syntax- condition, compared to the weaker - semantic - condition B. The crucial way to evaluate the Reflexivity predictions is to compare (1a) with the judgment on the ECM structure (1b), where the weak condition B is observed. In all languages I mentioned above, there is a clear contrast here - (1b) is perfect, while (1a)- marginal. Probably, the same is true in many other languages.

One of the reasons this has not been widely noted is that the traditional BT cannot distinguish between objects and ECM subjects. Hence the SE case (1a) was always only compared to the pronoun case, (1a'), and the more relevant contrast (1a)-(1b) was not checked in the literature reporting the availability of local binding of SE. Another way to put it: condition B effects with SE are observable only given the reflexivity condition B. The traditional BT-condition B fails to distinguish objects and ECM subjects, since it was formulated to capture the distribution of pronouns. Pronouns, indeed, behave alike in these two contexts. For R&R this follows, since their distribution is governed also by the chain condition, by which objects and ECM subjects equally form a chain with the matrix subject. This, in fact, was one of the major arguments in R&R for separating the reflexivity and the chain conditions.

Hamilton suggests that condition B locality effects with SE are the marked, or 'language-specific' case. However, the fact that it is found in completely unrelated families of languages is significant. I would say that it is the other way around: The reported lack of this locality effects in Chinese and Japanese is the mystery to be explained.

2. The dravidian anaphora system.

There is no SELF marking on arguments (No SELF anaphors) in any of the Dravidian languages. Still, conditions A, B ('Reflexivity') are clearly and strictly obeyed. There are two systems of dealing with reflexivization:

a. Marking on V or agr: Morpheme 'koL' is the reflexive marker. So everything is as in Dutch, except that the marking is on the verb.

b. Duplication, or 'protection from binding': The coindexed element is embedded inside a more complex NP, thus avoiding a reflexivization environment, and all issues of condition B and chain.

-Kannada, is purely of system (a). (This is shown in Lidz' LI squib).

-Malayalam is purely of system (b) (no koL or any reflexive marker). The whole issue of reflexivization is avoided, by never allowing coindexation of co-arguments. (This is how Jayaseelan explains it. He believes the same is true universally. SELF anaphors have the function of protecting the NP from reflexivization (avant la lettre).

-Telugu combines both. The duplication system is different than Malayalam, and there is also an inflectional reflexive marker koL. While Kannada and Malayalam are easy to explain, Telugu is difficult, and appears more 'unsystematic'. I will therefore not discuss it here. (But I can send some summary of the facts and the conceptual directions open to account for them, upon request.)

-Tamil falls under the duplication method. But I don't have the data re whether it is like Malayalam or like Telugu.

3. Inflectional marking: kannada

(For some actual sentences, and more detailed analysis, see Lidz (1995) squib)

3.1. Reflexive marking and reflexivity condition A:

In R&R it was argued that reflexivization can, in principle, be marked morphologically either on the verb or on the argument (SELF) (R&R, p. 662). However, the only instance we had for marking on the verb was intrinsic reflexivization. It turns out that Kannada is the full fledged example of the later option: The only way to reflexive mark the predicate is by marking the verb. This applies freely to any verb, as in (2a).

2a) Rama tann-annu hode-du-koND-a
Rama SE-acc hit-pp-**koL**.pst-3 (Rama SE self-hit)
Rama hit himself

b) *Rama Mary hit-koL-s (=Rama self-hit Mary)

c) *Rama thinks-koL that he is a fool.

The way reflexivity condition A is active is illustrated in (2b): Since a reflexive marked predicate is not reflexive, the sentence is out. (2c) shows that koL is indeed a reflexive marker, and not some general anaphora marker.

3.2. Conditions B and chain:

The following chart summarizes how the facts are derived in the Reflexivity framework:

			B	Chain
3a)	*Rama loves SE		*	ok
b)	*Rama loves him		*	*
c)	Rama koL-love SE		ok	ok
b)	*Rama koL-love him[1]	ok	*	
4a)	Rama believes [SE fool]		ok	ok
b)	Rama koL-believe [SE fool]		ok	ok

(4b) deserves a comment: It poses no problem to Reflexivity condition B, but the question is why it is allowed by condition A, given that the matrix predicate is reflexive-marked and still, no two of its direct arguments are coindexed. Recall that for independent other problems, condition A applies in R&R to syntactic rather than semantic predicates. The definition of syntactic predicates incorporates ECM subjects. Rama and SE are, therefore, coarguments of a syntactic predicate, even though not of a semantic one. So this syntactic predicate is appropriately reflexive.

3.3. Some implications.

The morphological reflexive-marking of the verb in Kannada completes the similarity between reflexivization and agreement systems: Just as agreement can be morphologically-marked either on the argument or on the verb, so can reflexive marking.

As far as I know, Kannada provides also a confirmation of another prediction in R&R. A question raised by R&R (p. 662) is whether one reflexive marking is sufficient to license the coindexation of more than two arguments, e.g. in the case of a three-place predicate. Based on reflexive-marking of the argument, we concluded that this is the case, and defined reflexive-marking accordingly. Again, we did not have any example involving a morphological marking of the verb, rather than the argument. But, according to R. Amritavalli (p.c), This is indeed true of Kannada.

5 a) Rama pustakavannu tanage koTTukoNDa
Rama book-acc SE-dat give-pp-**koL**.pst-3
Rama gave a book to himself

[1]My notes say that there are dialect differences on this option (with pronouns.)

b) *Rama pustakavannu tanage koTT-a[2]
Rama book-acc SE-dat give-pst-3
Rama gave a book to himself

c) *Rama tannannu tanage koTT-a
Rama SE-acc SE-dat give-pst-3.
Rama gave a book to himself

d) Rama tannannu tanage koTTukoNDa
Rama SE-acc SE- dat give-**koL**.pst.3
Rama gave himself to himself

Note, first, that the reflexive morpheme koL can license coindexation of the dative argument with the subject, as in (5a). Without koL, anaphora is impossible here, as in (5b). Similarly, the dative and the accusative arguments cannot be coindexed in the absence of reflexive marking with koL, as in (5c). But when koL is present it licenses coindexation of all three arguments, as in (5d). (The equivalent paradigm with Dutch zelf is exemplified in detail with R&R's (21)).

3.4. Unsolved problems:
We have no solution for the NP issue discussed in Lidz' squib. More generally, not enough work has been done, in the Reflexivity framework, on SE anaphora inside NPs.

4. The protective way: malayalam.
(Detailed survey and examples can be found in Jayaseelan's (1995) and previous papers.)

4.1. The basic facts.
-All NPs (Pron, SE, Proper name, no data given on Definite descriptions) can occur with the suffixed tanne:

6) -SE: taan -- taan-tanne
-pron: awan --awan-tanne
-P names: Rama -- Rama-tanne

-With locally coindexed NPs (which would otherwise be co-arguments of the same predicate), occurrence of tanne is obligatory - None of the left members of (6) can occur alone in this context.

-All three right members of (6) can occur locally. I.e. the suffixed tanne exempts from both condition B and Chain.

-The tanne does not induce locality. All (right and left) members of (6) can occur in the long-distance context. The distribution of the complex (right) unit is determined by the original

[2]I am not sure I got the full verbal morphology right here from my notes. The relevant point, however is that there is no koL morpheme.

anaphoric element:

-taan-tanne, like the single SE taan, is subject oriented, requires a c-commanding antecedent, and shows blocking effects.

-awan-tanne, like awan, is not subject oriented and does not require c-command.

4.2. Jayaseelan's account:

-tanne is some kind of a focused element.

-It exempts from condition B (and chain), because the relevant DP is no longer coindexed locally with its antecedent (just as in John likes his mother, John and his are not locally coindexed.) The internal structure of the DP is (7), with tanne as D. The specifier DP is not coindexed with the top DP (just as 'his father' is not coindexed his). In (8), the top DP is a coargument with the subject, but not coindexed with it, hence all is fine.

7 DP

$DP_{(i)}$ D'

taan/awan D

tanne

8) $Rama_{(i)}$ [[taan]$_{(i)}$-tanne] loves.

COMMENT: It is not fully clear that the idea of tanne as D can be maintained. Haripasad (dissertation) noted that in Telugu (which is independently different) the suffix can occur after PP, as in (9) (which is [kind of] Telugu, hence the koL])

9) Jaanu [[PP tana-too] tannu] spoke-koL.
with
Jaanu spoke with SE(self)

Jayaseelan (pc) acknowledged that the same is true for Malayalam. Alternatively, the structure could be some sort of an adjunction of the tanne to DP, PP, etc. In any case, it should have, for Malayalam, the property of blocking co-arguments coindexation and chains. (Telugu is more complex, as I mentioned.)

4.3.Condition A.

So far, then, there is just no reflexivization in Malayalam. Note that if the inflectional marker koL existed in the language, we would have gotten a violation of condition A with the complex anaphors, as in (10).

10) $Raama_{(i)}$ [[$taan_{(i)}$]-tanne] loves-koL

(Since koL reflexive-marks the predicate, its arguments should be coindexed. But there is no

coindexation of co-arguments in (10).) However, there is no koL in Malayalam. So the language just chose to avoid the issue of reflexivization altogether (a point discussed also in Lidz' article to this conference). Reflexivity Condition A is appropriately met, since there are no reflexive-marked predicates in this language.

5. Implications for emphatic anaphors.

What we saw is that there are two systems of dealing with the reflexivity requirements: One is to reflexive mark the predicate, either on the verb, or on the argument. The other, is to embed the anaphoric argument, thus avoiding a reflexive predicate altogether (Malayalam). Jayaseelan argues that in fact, there is universally, only one system, namely, the second (Malayalam) type. The SELF in English and other languages is just a focus-emphatic marker, exempting pronouns from condition B. If this is true, then condition A could be dispensed with, and we only need reflexivity condition B. I think that although, possibly, this line could be further developed, it may be far-fetched, at the present stage of our understanding of the problems. If true, it would be very difficult to explain why it is so hard to get this focus marking in, say, English, when the anaphor is an argument. Specifically, why is it so much easier to find the anaphor in the standard local context (11) than in (12a)?. (In Malayalam, no such contrast is found.) Further, why do we find the familiar contrast in (12a,b)?. (If anything, it should be easier for the anaphor to be focus in (12a) than in (12b). Still it is very difficult to get (12a), even as focus.) In this respect, English is different from Malayalam, which allows complex anaphors in equal ease in all syntactic contexts[3].

11) Max praised himself.
12 a) *Max wants us to praise himself
b) Max wants us to praise Lucie and himself.

But the general finding of Jayaseelan, namely that there is another form and function of complex anaphors is very important. Possibly, there are, universally, two ways anaphors may satisfy the reflexivity requirements, and accordingly, even in languages with real SELF anaphors, like English, there may be two syntactic functions of SELF -as reflexive-marker and as a 'protector' of an index. When both are available, the second is used for 'emphatic' purposes. I mentioned in my previous comments on Lidz that in R&R we argued that all instances of emphatic SELF anaphors in an argument position, are foci, in the standard sense, and we were forced to assume they undergo focus movement at LF, which then exempts them from condition A. But independently of this problem, I don't really believe that foci must move at LF. So, it would be nice if an alternative account could be found.

---

[3]The desire to eliminate condition A from the reflexivity framework has very good reason: Condition A looks more syntactic than condition B. First, it has to be stated in terms of syntactic, rather than semantic, predicates, and next, it is stronger than condition B -yielding much worse derivations when violated. Nevertheless, I don't think the line could be assuming the Malayalam-type protection system universally. Rather, the alternative seems to be reducing condition A to the chain condition, which applies at very similar environment. This line was proposed and fully executed by Fox (1993).

Baker (1994) proposes a distinction between anaphoric and "non anaphoric intensive use"[4]. The latter is generated as he himself, in the nominative, but is realized as himself in the accusative (Hence the confusion, since it is invisible in the accusative.) The idea that some accusative anaphors are just another form of the he himself emphatic form was proposed also by Bickerton (1987), who argued, if I remember correctly, that the pronoun or pro is deleted in the accusative.

Under this analysis, then, emphatic English forms may be viewed as having the same structural properties as those of Malayalam. Note, however, that SELF, unlike tanne in Malayalam, does have a reflexive-marking function. Hence, the contribution of the syntactic embedding in the complex (emphatic) anaphor is not just to 'protect' the pronoun form conditoin B and chain effects, but also to 'protect' the SELF from reflexive-marking. In a structure like 'he-himself' the SELF is too deeply embedded to reflexive mark the verb of which it is an argument. For this reason, such anaphors can occur also in a context like (12a), without reflexive-marking the verb. Nevertheless, there should be clear contextual reason to prefer this more complex way to express anaphora. As I said in my previous reply to Lidz, I think this reason is that it is easier to stress a complex anaphor than a pronoun. When the reflexivity conditions allow the use of a normal anaphor, as in (12b), this normal anaphor can be used for stress, perspective, or other reasons. But when it cannot occur, only the emphatic form can be used to enable the anaphor to carry stress, without, at the same time, reflexive-marking the verb incorrectly.

But along with protecting the SELF from reflexive-marking, the complex embedded form still has the same effect, noted by Jayaseelan, of protecting also the embedded pronoun ('he' in 'he-himself') from condition B (and Chain). Hence, in the local context, it can be used to obtain coreference without creating a reflexive predicate. This was the case in Lidz example of ellipsis (Max defended himself better than Bill did), which I discussed in the previous round.

We should keep in mind, however, that introducing ambiguity of SELF anaphors in English is a dangerous move. Unless we have very clear evidence and tests for identifying the two forms, we run into the danger of unfalsifiability. -We may end up being always right, but in the wrong way. To my recollection, Baker does not substantiate his syntactic analysis very deeply, so it would be necessary to find more evidence for a syntactic ambiguity in English, along the lines he proposes. Specifically, more work is needed on how the analysis still entails the contrasts in (11)-(12). Nevertheless, the idea is worth pursuing.

I don't have much more to say on how to argue for such a structural distinction, but at first glance, it seems that Hamilton's finding on Icelandic may be relevant here. In Icelandic, there are two forms of the SELF anaphor: 'sjalfur sig' and 'sig sjalfur'. Hamilton notes that they do not have the same distribution. Since there are two forms here, there is room for checking whether they could each belong to a different type of those outlined by Baker.

HAMILTON SAYS:

---

[4]Baker argues that the distribution of 'logophors' is freer in British than American English. He notes also that emphatic anaphors occur in an argument position of a verb in British English more frequently than assumed by R&R. (He gives examples from Jane Austin, in his (14). In fact, similar facts can be found in the Zribi-Hertz collection. Though, statistically, they are rare in her collection.)

"Unlike _sjalfur sig_ (which in all relevant respects behaves as predicted by R&R's system), _sig sjalfur_ may not be bound by a coargument (_sig sjalfur_ is possible in (13) only on an emphatic reading). (Note 4) This is so even though _sig sjalfur_ and _sjalfur sig_ both contain the putative SELF morpheme _sjalfur_: (Note 5)

(13) Jon(i) talar oft vid sig sjalfan(*i)/sjalfan sig(i)
John talks often to himself

However, like _sjalfur sig_, _sig sjalfur_ may be nonemphatically locally bound when it is the object of an adjunct _about_-PP (see below on the adjunct status of _about_-PPs):

(14) Maria(i) taladi vid Jon um sig sjalfa(i)/sjalfa sig(i)
Mary talked to John about herself

Neither _sig sjalfur_ nor _sjalfur sig_, however, may be (nonemphatically) locally bound when the object of some other adjunct PP:

(15) Jon(i) sa snak nalaegt ser sjalfum(*i)/sjalfum ser(*i)
John saw snake near himself "

COMMENT:
Given this data-base, it seems that only the SELF-SE form (_sjalfur sig_) is the reflexive-marking SELF anaphor in Icelandic. Hence, it can occur in the reflexivization environment (13). Given the logic of Jayaseelan and Baker, the syntax of the second form (SE-SELF) is different, so the index of the whole NP is not the same as the index of 'sig'. When this second form appears in the local context (13), all SELF does is to protect the index of sig from being co-bound. Since icelandic, unlike Malayalam, also allows actual coindexation in this context, there should be a reason why it is avoided, in favor of the SE_SELF form. The most easily accessible interpretation is that the emphatic reading is desired, as observed by Hamilton. But perhaps it could also be used to obtain coreference rather than variable binding, in contexts like those discussed by Lidz. In any case, no reflexive predicate is formed in this case. In non -reflexive contexts, as (14), the reflexivity conditions say nothing re which of the two should occur, since neither violates any condition. Hence, it is up to discourse to decide which of these is relevant, and whether the use is emphatic or not, cannot be determined in isolation.

Note, since this was brought up, that the problem in (15) is independent. It indicates a certain shortcoming of the treatment of PP anaphora in R&R. In fact, (15) is not so good (non-emphatically) also in English. An excellent alternative analysis of PP's in the reflexivity

framework is offered by Beit Arie (1994), which correctly rules (15) out (unless it is emphatic).

==========

Baker, C.L. (1994) "Locally free reflexives, contrast and discourse prominence in British English" Presented at the Annual Meeting of the Linguistic Society of America, Boston. (ms. University of Texas, Austin.)

Beit-Arie, Oren (1994), "PP structure and anaphora", ms. Tel Aviv University. Obtainable by e-mail: beitarie@aleph.co.il

Bickerton, D. "He himself - Anaphor, pronoun, or ...?" Linguistic Inquiry 18, 345-348, 1987.

Fox, Danny (1993) "Chain and binding: A modification of Reinhart and Reuland's "Reflexivity". Ms. MIT. (available at request: fox@mit.edu)

Jayaseelan, K.A. (1995) "Anaphors as Pronouns," ms. CIEFL: Hyderabad. (Also earlier manuscripts).

Lidz, J. (1995) "Morphological Reflexive Marking: Evidence from Kannada," Linguistic Inquiry 26:705-710.