

## **REFERENCE-TRACKING AND COMPETING CONSTRAINTS**

**Bernard Comrie**

*University of Southern California*

**Abstract:** The general phenomenon of reference-tracking includes the interaction of a number of competing principles. Two principles are discussed in this paper: (a) the use of more marked forms, including in particular reflexive pronouns, to indicate coreference in the most restricted domain(s) (the DOMAIN LOCALITY PRINCIPLE), and (b) the tendency for reflexive pronouns to require a subject antecedent (the ANTECEDENT PRINCIPLE). Under certain circumstances these two principles can conflict. Examples are examined from various languages to illustrate resolution of this conflict, including strictly grammaticalized solutions and more ad hoc possibilities.

**Keywords:** Reference-tracking; Reflexive pronouns; Rules; Constraints; Hierarchies; Typology

### **1. REFLEXIVES AND THE RESTRICTED DOMAIN**

The work reported on in this article is part of a more general study of reference-tracking from a crosslinguistic, typological viewpoint. A particular focus of this research is the establishment of crosslinguistically valid hierarchies to subsume particular observed patterns of crosslinguistic distribution. One of the patterns that will be of prime interest in this article is the strong tendency for marked indicators of coreference, such as reflexive pronouns, to be used in more local domains rather than in more extended domains, where "domain" can be understood, at least for present purposes, as "syntactic domain", as defined by the hierarchical structure of the sentence. A simple example can be seen by comparing English and Russian.

In English, reflexive pronouns are obligatory as indicators of coreference when, roughly speaking, the antecedent and the remention are both arguments of the same predicate, as in (1)–(2):<sup>1</sup>

- (1) John<sub>i</sub> saw himself<sub>i</sub> in the mirror.
- (2) John<sub>i</sub> wrote a letter to himself<sub>i</sub>.

For more extended domains, for instance if the remention is an adjunct, or a possessor noun phrase, or in a distinct (nonfinite or finite) clause, then English uses the ordinary pronoun, as in (3)–(6):

- (3) John<sub>i</sub> heard steps behind him<sub>i</sub>.
- (4) John<sub>i</sub> saw his<sub>i</sub> book.
- (5) John<sub>i</sub> asked Mary to make him<sub>i</sub> some tea.
- (6) John<sub>i</sub> said that Mary loves him<sub>i</sub>.

Thus, English requires a reflexive pronoun in the most local domain (arguments of the same predicate), and requires ordinary pronouns elsewhere.

By contrast, Russian, as illustrated in (7)–(12), requires a reflexive pronoun for arguments of the same predicate, for adjuncts, and for possessors; it optionally allows a reflexive pronoun, alongside an ordinary pronoun, across a nonfinite, infinitival clause boundary; and requires an obligatory ordinary pronoun across a finite clause boundary:

- (7) Volodja<sub>i</sub> uvidel sebja<sub>i</sub> (REFL) v zerkale.
- (8) Volodja<sub>i</sub> napisal sebe<sub>i</sub> (REFL) pis'mo.
- (9) Volodja<sub>i</sub> uslyšal za soboj<sub>i</sub> (REFL) šagi.
- (10) Volodja<sub>i</sub> uvidel svoju<sub>i</sub> (REFL) knigu.
- (11) Volodja<sub>i</sub> poprosil Tanju vskipjatit' sebe<sub>i</sub> (REFL)/emu<sub>i</sub> čaj.
- (12) Volodja<sub>i</sub> skazal, čto Tanja ego<sub>i</sub> ljubit.

(The Russian examples are literal translations of the English examples and in the same order, with substitution of Volodya and Tanya for John and Mary. *Emu* in (11) and *ego* in (12) can, of course, also indicate reference to some third party other than Volodya or Tanya.) Thus, in Russian use of the reflexive characterizes the most local domain, but extends a fair distance beyond this.

Note that the overall hypothesis is that it is possible to establish a cross-linguistically valid hierarchy (perhaps not entirely linear) ranging from most to least local domain. A language having reflexive pronouns distinct from nonreflexive pronouns will start with distinct reflexive pronouns in the most local domain, and may extend them beyond this. For each such language, there will be a cutoff point, although this must be understood to allow for a possible intermediate area where distinct reflexives are possible, but not obligatory, as with infinitival constructions like (11) in Russian.

Even more interesting cross-linguistic data are provided by languages with a threeway opposition, with ordinary pronouns in the least local domain, ordinary reflexives in an intermediate domain, and emphatic reflexives—an even more marked form of reflexive—in the most local domain. In the recent generative literature this pattern has been documented most

extensively for the Mainland Scandinavian languages. The examples below are from Danish.<sup>2</sup> The relevant forms are: emphatic reflexive *sig selv*, ordinary reflexive *sig* (possessive *sin*), ordinary pronoun masculine singular *ham* (possessive *hans*), feminine singular *hende* (possessive *hendes*).

In Danish, the emphatic reflexive is required when, roughly speaking, the antecedent and the remention are both arguments of the same predicate, as in (13)–(14):<sup>3</sup>

- (13) Peter kritiserer sig selv.  
       ‘Peter<sub>i</sub> criticizes himself<sub>i</sub>.’  
 (14) Sofie sendte brevet til sig selv.  
       ‘Sofie<sub>i</sub> sent a letter to herself<sub>i</sub>.’

In (13)–(14), the ordinary reflexive *sig* is not possible.<sup>4</sup> The ordinary reflexive is used obligatorily for adjuncts and possessors, as in (15)–(16):

- (15) Sofie lagde bøgerne bag sig.  
       ‘Sofie<sub>i</sub> put the books behind her<sub>i</sub>.’  
 (16) Peter tog sin frakke.  
       ‘Peter<sub>i</sub> took his<sub>i</sub> coat.’

(In (15) *sig selv* would be possible, but would give an explicitly emphatic reading.) In (15)–(16), nonreflexive *hende* and *hans*, respectively, would not be possible, with the interpretations indicated. The ordinary reflexive is possible, alongside the ordinary pronoun, across an infinitival clause boundary, as in (17):

- (17) Marie hørte Eva kritisere sig/hende.  
       ‘Marie<sub>i</sub> heard Eva criticize her<sub>i</sub>.’

(In (17), *hende* can also refer to some third party other than Marie or Eva. For coreference with Eva, the understood subject of the infinitival clause, the emphatic reflexive *sig selv* would, of course, be required.) If there is an intervening finite clause boundary, only the ordinary pronoun is possible, as in (18):

- (18) Marie siger at Eva kritiserer hende.  
       ‘Marie<sub>i</sub> says that Eva is criticizing her<sub>i</sub>.’

(*Hende* in (18) can, of course, also refer to some third party other than Marie or Eva. The emphatic reflexive *sig selv* would refer back to Eva. The ordinary reflexive *sig* would have no possible interpretation.) In terms of the hierarchy of degrees of locality of domain, Danish requires the emphatic reflexive for arguments of the same predicate, requires the ordinary reflexive for adjuncts and possessors, allows either the ordinary reflexive or the ordinary pronoun across an infinitival boundary, and requires the ordinary pronoun across a finite clause boundary.

Japanese presents data similar to those from Danish, though with somewhat different cutoff points; similar patterns are also attested for Korean and Chinese and for a number of South Asian languages, including both Indo-Aryan and Dravidian languages. In Japanese, the emphatic reflexive is *zibun zisin*, the ordinary reflexive is *zibun*, and the ordinary pronoun is *kare* (although zero anaphora is usually more natural). For arguments of the same predicate,

Japanese uses the emphatic reflexive, although at least some speakers also allow the ordinary reflexive, as in (19):

- (19) Taroo wa zibun zisin/zibun o but-ta.  
 Taro TOP EMPH:REFL/REFL ACC hit-PST  
 'Taro<sub>i</sub> hit himself<sub>i</sub>.'

For adjuncts and possessors, the ordinary reflexive is used, as in (20)–(21):

- (20) Taroo wa zibun ni is-satu hon o kat-ta.  
 Taro TOP REFL for one-CLF book ACC buy-PST  
 'Taro<sub>i</sub> bought a book for himself<sub>i</sub>.'
- (21) Taroo wa zibun no kao o arat-ta.<sup>5</sup>  
 Taro TOP REFL GEN face ACC wash-PST  
 'Taro<sub>i</sub> washed his face<sub>i</sub>.'

Japanese does not have an obvious analog of the distinction between infinitival and finite clauses in the European languages discussed above, and the ordinary reflexive is possible, alongside the ordinary pronoun, even across clause boundaries that correspond to finite clause boundaries in European languages; one thus finds the same pattern in Japanese examples (23) and (24):

- (22) Hiroshi wa Mitiko ni zibun no heya de  
 Hiroshi TOP Michiko DAT REFL GEN room in  
 benkyoo su-ru yoo ni it-ta.  
 study do-PRS way to say-PST  
 'Hiroshi<sub>i</sub> told Michiko<sub>j</sub> to study in his<sub>i</sub>/her<sub>j</sub> room.'
- (23) Hiroshi wa Mitiko ga zibun no heya de  
 Hiroshi TOP Michiko NOM REFL GEN room in  
 benkyoo si-te i-ru tte it-ta.  
 study do-CVB be-PRS that say-PST  
 'Hiroshi<sub>i</sub> said that Michiko<sub>j</sub> is studying in his<sub>i</sub>/her<sub>j</sub> room.'

(In (22)–(23), *zibun* can also refer to Michiko, as subject of the subordinate clause.)

These data are summarized in Table 1.

In closing this section, it should be noted that the emphatic reflexive, although in all examples known to me involving a morpheme or morphemes in addition to the ordinary reflexive, does not necessarily have this morpheme attached to or adjacent to the form of the ordinary reflexive. In Tamil, for instance, the emphatic reflexive is marked by a verbal morpheme *koL-* (past tense *koN-T-*), literally 'seize, hold', which is attached to the converb (gerund) form of the main verb. Compare (24), with the ordinary reflexive, and allowing either the subject of its own clause or the subject of the higher clause as antecedent, with (25), with the emphatic reflexive, allowing only its own subject as antecedent:

- (24) kumaar raajaa tann-aip parrip peec-in-aan en-ru ninai-tt-aan.  
 Kumar Raja REFL-ACC about talk-PST-3SM that think-PST-3SM  
 'Kumar<sub>i</sub> thought that Raja<sub>j</sub> was talking about him<sub>i</sub>/himself<sub>j</sub>.'



Table 1: Indication of coreference in selected languages

	English	Russian	Danish	Japanese
argument of same predicate	reflexive	reflexive	strong reflexive	strong reflexive/ ordinary reflexive
adjunct	ordinary pronoun	reflexive	ordinary reflexive	ordinary reflexive
possessor	ordinary pronoun	reflexive	ordinary reflexive	ordinary reflexive possible
across infinitival clause boundary	ordinary pronoun	reflexive/ ordinary pronoun	ordinary reflexive/ ordinary pronoun	ordinary reflexive possible
across finite clause boundary	ordinary pronoun	ordinary pronoun	ordinary pronoun	

- (25) kumaar raajaa tann-aip parrip peec-ik koN-T-aan en-ru ninai-tt-aan.  
 Kumar Raja REFL-ACC about talk-CVB hold-PST-3SM that think-PST-3SM  
 ‘Kumar<sub>i</sub> thought that Raja<sub>j</sub> was talking about himself<sub>j</sub>.’

## 2. ANTECEDENTS OF REFLEXIVES

In section 1, we examined the status of the remention, whether a reflexive pronoun (ordinary or emphatic) or an ordinary pronoun, in relation to the rest of the sentence. In this section, we are concerned with the status of the antecedent. In many languages, the antecedent of a reflexive pronoun must be a subject, or an actor.<sup>6</sup> Of the languages represented in Table 1, for instance, this is, more or less strictly, the case for Russian, Danish, and Japanese. In Russian, for example, albeit with a certain amount of speaker variation,<sup>7</sup> reflexives can normally only refer to the subject of the sentence—in (26) to *Volodja*, and not to the indirect object, *Tane*, even though the reflexive is not marked for the gender of its referent:

- (26) Volodja dal Tane svoju knigu.  
 ‘Volodya<sub>i</sub> gave Tanya<sub>j</sub> his<sub>i</sub>/\*her<sub>j</sub> book.’

The few exceptions that are cited in the literature and that are readily acceptable to all or nearly all speakers appear to be fixed phrases, not productive syntax, as for instance in *položít’ čto-to na svoe mesto* ‘to put something<sub>i</sub> in its<sub>i</sub> place’, i.e. ‘to put something back in its proper place’, *zastat’ kogo-to u sebja/v svoem dome* ‘to find someone<sub>i</sub> at home (lit. at self<sub>i</sub>)/in their<sub>j</sub> house’. In Danish example (27), the reflexive (in fact, given the configuration, the emphatic reflexive) *sig selv* can only refer back to Sofie:

- (27) Sofie fortalte Hanna om sig selv.  
 ‘Sofie<sub>i</sub> told Hanna<sub>j</sub> about herself<sub>i</sub>/\*j.’

(In (30), *ee* could also refer to some third participant other than Volodya or Tanya.) In other words, although both antecedent and remention are inside the VP, the fact that the remention is itself within a noun phrase inside the VP creates a greater domain distance. This does still mean, however, that the principle on the antecedent takes preference over the principle on domain distance.

More problematic is the case where we are dealing literally with two objects of the same predicate as antecedent and as remention. Danish has a distinct, grammaticalized solution to this problem. Neither the ordinary reflexive nor the emphatic reflexive is possible, since both require a subject antecedent. The ordinary pronoun is not possible, since it cannot indicate coreference in domains as local or more local than the clause. The solution is a distinct combination, using the ordinary pronoun together with the emphatic element *selv*, which has already been seen in the formation of the emphatic reflexive. This is illustrated in (31):

- (31) Sofie fortalte Hanna om hende selv.  
'Sofie told Hanna<sub>i</sub> about herself<sub>i</sub>.'

Although grammaticalized, this is still something of an ad hoc solution within the system as a whole, since there is nothing about this combination that would lead one to believe that it should be appropriate for the most local domain. Indeed, the same combination can be used to indicate coreference with a subject across a finite clause boundary ("logophoric environment"), as in (32):

- (32) Komponisten sagde at orkestret kun måtte spille symfonien med ham selv som dirigent.  
'The composer<sub>i</sub> said that the orchestra could only play the symphony with himself<sub>i</sub> as conductor.'

In Russian, there seems to be no strictly grammaticalized way of dealing with the situation, and different speakers and writers seem to adopt ad hoc strategies on the (admittedly very rare) occasions on which the construction crucially occurs. One case is (33), brought to my attention by Alan Timberlake:

- (33) Cvetaeva protivopostavljaet Kazanovu ne tol'ko ego ničtožnomu okruženiju v zamke, no i ego — samomu sebe.  
'Cvetaeva not only opposes Casanova<sub>i</sub> to his<sub>i</sub> insignificant surroundings in the castle, but also (opposes) him<sub>i</sub> to himself<sub>i</sub>.'

In the first clause, the antecedent is direct object and the remention is a possessor, so the ordinary pronoun is used for the remention. In the second, elliptical clause the antecedent is direct object and the remention is indirect object. The author's solution is to have recourse to an emphatic reflexive, combining the reflexive pronoun *sebjä* (dative *sebe*) with the emphatic pronoun *sam* (masculine singular dative *samomu*). Although the emphatic reflexive is not a grammaticalized form in Russian as it is in Danish or Japanese, its occurrence here is nonetheless entirely fitting from the viewpoint of the domain locality principle: objects of a single predicate form the most restricted domain, and it is precisely here that the author has recourse to the emphatic reflexive. But note that the usual antecedent condition is relaxed.

#### 4. CONCLUSION

In this paper, I have examined the interaction of two principles relating to reference-tracking, namely the principle on domain locality and the principle on the antecedent. Conflict between the two principles is possible in the admittedly rare instances where both antecedent and remention are within the VP. In some languages, such as English, the particular instantiation of the second principle means that the two principles do not come into conflict. (This could also be interpreted as consistent ranking of the domain locality principle over the antecedent principle.) In other languages, however, a stricter version of the antecedent principle, requiring a subject antecedent, leads to conflict. In some instances, the strict antecedent principle wins out, leading to use of an ordinary pronoun even in a highly local domain. In other cases, other solutions, grammaticalized or ad hoc, are resorted to.

#### NOTES

1. There is a certain amount of leakage across the argument/adjunct boundary, though the details are not relevant to our present concerns. Reflexives are required in standard English for benefactives, for instance, although these are not arguments, as in (i)–(ii):

- (i) John<sub>i</sub> bought a cake for himself<sub>i</sub>.
- (ii) John<sub>i</sub> bought himself<sub>i</sub> a cake.

2. I am grateful to Uffe Bergeton Larsen for providing me with the Danish examples. Many are taken or adapted from Jakubowicz (1992) and Vikner (1985), the latter not directly available to me.

3. As noted for English in note 1, the precise characterization of the dividing line requires further specification. Note in particular the different treatment in Danish of semantic benefactives encoded with a preposition (requiring *sig selv*) as in (i) and without a preposition (requiring *sig*) as in (ii):

- (i) Sofie købte en bog til sig selv.  
'Sofie<sub>i</sub> bought a book for herself<sub>i</sub>.'
- (ii) Sofie købte sig en bog.  
'Sofie<sub>i</sub> bought herself<sub>i</sub> a book.'

4. As has been observed in the generative literature, some verbs, called affected verbs, do allow the ordinary reflexive even in the most local domain, as in *Peter vasker sig* 'Peter washes (himself)'. The variant *Peter vasker sig selv* gets an emphatic reading on the reflexive object. These are in general verbs that describe actions typically done by an agent to him/herself, and in English often translate as intransitives. I assume that this is another principle that interacts with those discussed in the text, but I have not investigated the phenomenon in sufficient detail to propose a more specific account of the interaction.

5. In (21) *Taroo wa kao o aratta* would also be possible, with no specification of the possessor, it being inferred pragmatically that Taro probably washed his own face. This is probably not an instance of zero anaphora, since the overt ordinary pronoun in *Taroo wa kare no kao o aratta* prefers an interpretation of noncoreferentiality between *kare* and *Taroo*.

#### 4. CONCLUSION

In this paper, I have examined the interaction of two principles relating to reference-tracking, namely the principle on domain locality and the principle on the antecedent. Conflict between the two principles is possible in the admittedly rare instances where both antecedent and remention are within the VP. In some languages, such as English, the particular instantiation of the second principle means that the two principles do not come into conflict. (This could also be interpreted as consistent ranking of the domain locality principle over the antecedent principle.) In other languages, however, a stricter version of the antecedent principle, requiring a subject antecedent, leads to conflict. In some instances, the strict antecedent principle wins out, leading to use of an ordinary pronoun even in a highly local domain. In other cases, other solutions, grammaticalized or ad hoc, are resorted to.

#### NOTES

1. There is a certain amount of leakage across the argument/adjunct boundary, though the details are not relevant to our present concerns. Reflexives are required in standard English for benefactives, for instance, although these are not arguments, as in (i)–(ii):

- (i) John<sub>i</sub> bought a cake for himself<sub>i</sub>.
- (ii) John<sub>i</sub> bought himself<sub>i</sub> a cake.

2. I am grateful to Uffe Bergeton Larsen for providing me with the Danish examples. Many are taken or adapted from Jakubowicz (1992) and Vikner (1985), the latter not directly available to me.

3. As noted for English in note 1, the precise characterization of the dividing line requires further specification. Note in particular the different treatment in Danish of semantic benefactives encoded with a preposition (requiring *sig selv*) as in (i) and without a preposition (requiring *sig*) as in (ii):

- (i) Sofie købte en bog til sig selv.  
'Sofie<sub>i</sub> bought a book for herself<sub>i</sub>.'
- (ii) Sofie købte sig en bog.  
'Sofie<sub>i</sub> bought herself<sub>i</sub> a book.'

4. As has been observed in the generative literature, some verbs, called affected verbs, do allow the ordinary reflexive even in the most local domain, as in *Peter vasker sig* 'Peter washes (himself)'. The variant *Peter vasker sig selv* gets an emphatic reading on the reflexive object. These are in general verbs that describe actions typically done by an agent to him/herself, and in English often translate as intransitives. I assume that this is another principle that interacts with those discussed in the text, but I have not investigated the phenomenon in sufficient detail to propose a more specific account of the interaction.

5. In (21) *Taroo wa kao o aratta* would also be possible, with no specification of the possessor, it being inferred pragmatically that Taro probably washed his own face. This is probably not an instance of zero anaphora, since the overt ordinary pronoun in *Taroo wa kare no kao o aratta* prefers an interpretation of noncoreferentiality between *kare* and *Taroo*.

Judgments on *kare* 'he', *kanozyo* 'she', are notoriously fluid in contemporary Japanese, and I have not included their systematic examination in this paper.

6. The extent to which the grammatical relation or the semantic (thematic) role takes precedence varies from language to language. While this variation is an interesting area for investigation, the following discussion sticks to simple examples where subject and actor coincide. Other factors that play an important role in many languages, sometimes in interaction with grammatical relations and/or semantic roles, are topic-comment structure (with topics as preferred antecedents) and linear order (in particular with a requirement or preference for the antecedent to precede the reflexive).

7. Thus, some speakers of Russian will accept the direct object as antecedent of a reflexive in (i), albeit as a much less preferred alternative to the interpretation where the antecedent is the subject:

- (i) Volodja otvez Tanju k svojim roditeljam.  
'Volodyaj took Tanyaj away to hisj/herj parents.'

Such speakers have a system more similar to that described below for English. Other speakers will just as emphatically reject the interpretation with the direct object as antecedent and insist, in this interpretation, on the nonreflexive *ee* 'her' (which is acceptable in this interpretation to all speakers). The situation is no doubt complicated by the insistence of many normative grammars that (perhaps barring lexicalized exceptions) only subjects are permitted as antecedents of reflexives in Russian.

#### REFERENCES

- Jackendoff, Ray (1990). *Semantic structures*. MIT Press, Cambridge, MA.
- Jakubowicz, Celia (1994). *Sig en danois: syntaxe et acquisition*. In Hans-Georg Obenauer and Anne Zribi-Hertz (eds), *Structure de la phrase et théorie du liage* (Hans-Georg Obenauer, Anne Zribi-Hertz (Eds.)). Presses Universitaires de Vincennes, Saint-Denis.
- Maling, Joan (1990). Clause-bounded reflexives in modern Icelandic. In Joan Maling and Annie Zaenen (eds.), *Modern Icelandic syntax* (Joan Maling, Annie Zaenen (Eds.)), 277–287. Academic Press, San Diego.
- Vikner, Sten (1985). Parameters of binder and binding category in Danish. *Working Papers in Scandinavian Syntax* 23, 1–58.