

Long distance reflexives in Singapore Malay:

An apparent typological anomaly

Peter Cole and Gabriella Hermon

To appear in *Linguistic Typology* 2,1 (Spring 1998)

Abstract

There are a variety of properties which are believed to be associated with long distance reflexives. These include the requirement that the reflexive form be monomorphemic, subject orientation, and blocking by subjects intervening between the reflexive and its antecedent (in appropriate languages). Malay long distance reflexives have been claimed to contradict the expected typological correlations. We argue that Malay does not constitute a genuine counter example because, on closer examination, what appear to be long distance reflexives are in fact instances of a form which is indeterminate between a local reflexive and a pronoun.

Keywords: agreement, anaphora, binding domain, c-command, complex sentences, locality, long distance reflexive, Malay, pronoun, reflexive, syntax

1. Long distance reflexives and the parameterizing of the Binding Theory

An important topic in the recent generative literature on bound anaphora has been the examination and analysis of differences among languages in locality restrictions on reflexives. While this topic has clear typological implications, to the best of our knowledge, it has not been discussed in the typological literature. To take a well known case, reflexives in English require local (roughly, clausemate) antecedents while in Mandarin Chinese the antecedent can be non-local, any number of clauses from the anaphor:

(1) **Poirot_i thinks that Mary hurt himself_i.*

(2) *John_i believes that Bill_j said that Tom_k shaved himself*_{i/*j/k}.*

(3) *Zhangsan_i renwei Lisi_j zhidao Wangwu_k xihuan ziji_{i/j/k}.*

Zhangsan thinks Lisi knows Wangwu likes self

'Zhangsan thinks Lisi knows Wangwu likes him/himself.'

We shall refer to reflexives with non-local antecedents like those in (3) as long distance (LD) reflexives.

In this paper we shall review the typological properties believed to be associated with LD reflexives and will show briefly how these properties are predicted by current analyses of the construction. We will then turn to the main topic of the paper, an apparent instance of LD reflexives in Malay that seems to contradict the expected

correlations. We will show that on closer examination what appear to be LD reflexives in Malay are in fact a form which is indeterminate between a local reflexive and a pronoun.

We would like to begin by reviewing briefly the major generative analyses of LD reflexives in order to elucidate the typological predictions of these analyses. In early generative work on LD reflexives (e.g. Manzini & Wexler 1987 and Wexler & Manzini 1987), the cross linguistic variability seen in (1)-(3) was taken to be a primitive of grammatical theory: It was proposed that the locality conditions imposed by the Binding Theory are not universal, and that languages can differ parametrically in terms of the locality condition imposed by the grammar.

According to the Binding Theory,¹ a bound anaphor (i.e. reflexive or reciprocal) must have a c-commanding antecedent within a highly restricted domain, often referred to as the "Governing Category". (With few exceptions, the Governing Category is coextensive with the clause in which the anaphor occurs.) A pronoun must not have a c-commanding antecedent within that domain, but it can have an antecedent which does not c-command it or which is not within that domain. In contrast, a referring expression (e.g. a proper name or a relative clause) can never have a c-commanding antecedent, even one external to the Governing Category.

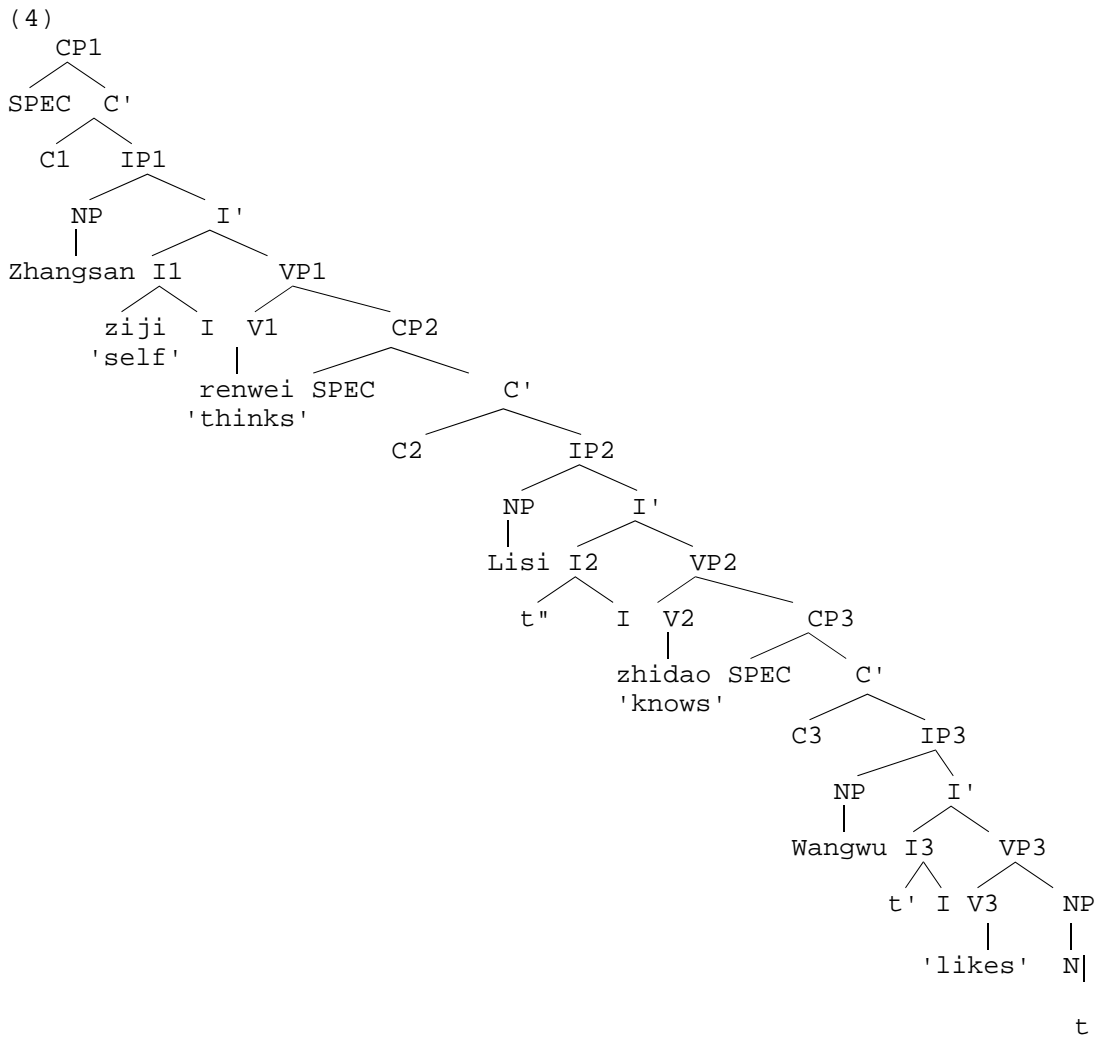
Accounts like that of Manzini and Wexler accommodate differences among languages by allowing the grammars of specific languages to contain statements extending or

restricting the binding domain. According to this view, pronouns and reflexives differ in a fundamental way. Reflexives require c-commanding antecedents, but pronouns do not. Thus, LD reflexives cannot refer to non-c-commanding antecedents. In contrast, pronouns need not have c-commanding antecedents and cannot refer to c-commanding antecedents within their local domains. These differences will be used later in the paper as tests for whether certain forms are pronominal or reflexive.

The parameterization hypothesis claims implicitly that the existence of LD reflexives is independent of other phenomena associated with the construction such as subject orientation (the requirement that the antecedent for the reflexive be a subject) and the blocking effect (the blocking of a reflexive-antecedent relation by intervening subjects). Thus, it fails to predict that other phenomena will correlate with the existence of LD reflexives in a language. In order to capture generalizations regarding such correlations, in much of the more recent generative literature, an alternative approach has been proposed. This approach makes much stronger typological predictions than does the parameterization of the locality domain. We shall first sketch the more recent approach and shall then review its typological implications.

It is claimed in much current work that, despite appearances to the contrary, reflexives must universally have local antecedents. In this literature, the appearance of non-locality seen in (3) is treated as due to Logical Form (LF) movement of the reflexive (or some analogous

device). For instance, according to the head movement analysis in Cole, Hermon, & Sung (1990), Cole & Sung (1994), Cole & Wang (1996) among others, the reflexive in (3) moves in LF to the INFL (or AGR, in Cole & Wang 1996) of its own clause, and from there to the INFL (or AGR) of the matrix clause.² When Binding Theory is applied, at LF, the reflexive stands in a local relationship to its antecedent (*Zhangsan*), as is seen in (4).³ Since the movement illustrated in (4) is head movement and not phrasal movement, it is predicted that LD reflexives must be monomorphemic since, with few exceptions, heads are monomorphemic. Thus, the fact that *himself* is a phrasal rather than an X⁰ reflexive explains the ungrammaticality of (2) on the LD interpretation.⁴



We would like to emphasize that the head movement analysis of LD reflexives not only eliminates the necessity of parameterizing the Binding Theory, but, as was first observed by Pica (1987), it also predicts correctly a number of otherwise mysterious properties of the construction:⁵

- (5) a. LD reflexives are monomorphemic (because heads are monomorphemic while phrases are not).
- b. LD reflexives are typically subject oriented (because, in most cases, only the subject c-

commands INFL, the site to which the reflexive moves).⁶

- c. In languages without subject-verb agreement, LD reflexives manifest the blocking effect, the blocking of a reflexive-antecedent relation due to the presence of an intervening subject with person features different from those of the local subject.⁷

These predictions will be illustrated further below.

According to the head movement analysis, these properties will hold of LD reflexives in all languages, and, hence, are typological properties of the LD reflexive construction.

While it is generally assumed in the generative literature that the properties outlined in (5) hold of LD reflexives across languages, there has in fact been no typological study in which this claim has been examined systematically. In the absence of such a study, it may be useful to provide some selected references to recent works on specific languages which largely conform to the regularities in (5):⁸

Czech: Toman (1991)

Russian: Rappaport (1986)

Serbo-Croatian: Progovac (1993)

Dutch: Koster (1985); Wilt (1985)

Norwegian: Vikner (1985); Hestvik (1990, 1991, 1992);

Hellan (1991);

Icelandic: Thrainsson (1991); Sigurjonsdottir (1993)
 Italian: Napoli (1979); Giorgi (1984, 1991); Manzini
 (1991)
 Kannada: Amritavali (1991)
 Malayalam: Jayaseelan (1997, 1998)
 Japanese: Kuno (1972, 1986, 1987); Katada (1989)
 Korean: Yang (1983, 1988); Kang (1988); Park (1985)
 Chinese: Y. H. Huang (1984); Battistella (1989); Tang
 (1989); Huang & Tang (1989); Cole, Hermon, &
 Sung (1990); Sung (1990); Cole & Sung (1994)

2. The exceptionality of long distance reflexives in Malay

On initial examination, long distance reflexives in Malay appear similar to those in Mandarin Chinese and other East Asian languages.⁹ As is shown in (6), the reflexive pronoun can take either a local or a long distance antecedent:

(6) *Ahmad_i tahu Salmah_j akan membeli baju untuk diri-nya_{i/j}.*¹⁰

Ahmad know Salmah will buy clothes for self-3SG
 'Ahmad knows Salmah will buy clothes for him/herself.'

In (6) *diri-nya* can refer either to the subject of its own clause, *Salmah*, or to the subject of the matrix clause, *Ahmad*, a pattern similar to that seen in (3) in Mandarin Chinese.

Despite their apparent similarity to Mandarin Chinese and other languages studied previously, as was pointed out by Yan Huang (1996), Malay reflexives present a number of anomalies with respect to the typological pattern observed in (5). First, in languages studied previously, LD reflexives are monomorphemic, while the seeming LD reflexive in (6) is bimorphemic, *diri-nya* self-3SG, like English *him-self*. According to the head movement analysis, it is the monomorphemic nature of LD reflexives that allows them to undergo head movement in (4).¹¹ In Mandarin Chinese the polymorphemic reflexive PRONOUN+*ziji* is strictly local:¹²

(7) Zhangsan_i renwei Lisi_j zhidao Wangwu_k xihuan ta-
*ziji*_{*i/*j/k}.¹³

Zhangsan think Lisi know Wangwu like he-self
 'Zhangsan thinks Lisi knows Wangwu likes him/himself.'

Thus, the use of the polymorphemic *diri-nya* as an LD reflexive in Malay is problematic, both in terms of the predictions of the head movement analysis and typologically.

Second, while LD reflexives in other languages are subject oriented, those in Malay are not. Examples (8) and (9) show that in Mandarin Chinese, while the local reflexive *ta-ziji* can refer to either a subject or a non-subject, the LD reflexive *ziji* is restricted to subject antecedents.

(8) Wangwu_i shuo Zhangsan_j zengsong gei Lisi_k yipian

guanyu

Wangwu say Zhangsan give to Lisi one.CLASS about
ta ziji_{i/j/k} de wenzhang.

he self POSS article

'Wangwu says that Zhangsan gave an article about
him/himself to Lisi.'

(9) Wangwu_i shuo Zhangsan_j zengsong gei Lisi_k yipian

guanyu

Wangwu say Zhangsan give to Lisi one.CLASS
about

ziji_{i/j/*k} de wenzhang.

self POSS article

'Wangwu says that Zhangsan gave an article about
him/himself to Lisi.'

In contrast, in Malay *dirinya* can take either a subject or
a non-subject antecedent, even when used long distance:

(10) Ali_i memberitahu Fatimah_j yang kamu menyukai

*dirinya*_{i,j}.

Ali tell Fatimah that you like self.3SG

'Ali told Fatimah that you like him/her.'

(11) Siti_i mengingatkan Mohamed_j yang saya tahu *dirinya*_{i,j}.

Siti remind Mohamed that I know self.3SG

seorang penjenayah.

one criminal

'Siti reminded Mohamed that I know he/she is a
criminal.'

Third, Malay is similar to Mandarin Chinese in failing to manifest subject-verb agreement. However, unlike other languages lacking subject-verb agreement (e.g. Mandarin Chinese and Malayalam), Malay fails to manifest the blocking effect: In Mandarin Chinese, when an immediately higher subject differs in person from a lower subject, LD reflexives are blocked, as is seen in (12)-(13).

(12) *Zhangsan_i renwei wo_j zhidao Wangwu_k xihuan
ziji_{*i/*j/k}.*

Zhangsan think I know Wangwu like self

'Zhangsan thinks that I know that Wangwu likes

himself.'

(13) *Ni_i renwei wo_j zhidao ziji_{*i/j} de taitai shi yige da hao
ren.*

you think I know self POSS wife is one big good
man

'You thought that I knew that my own wife was a very good
person.'

In (12) the subject of the clause in which *ziji* occurs is third person (*Wangwu*). The next subject up is the first person pronoun *wo*. As a result of the presence of *wo*, *ziji* cannot refer to the matrix subject, *Zhangsan*. This is despite the fact that *Zhangsan*, like *Wangwu*, is third person. The generalization, first noted by Y. H. Huang (1984) and formalized in Tang (1989), is that an intervening subject that differs in person from the local

subject will block the formation of an antecedent-reflexive relationship between an LD reflexive and a higher nominal that would otherwise be a possible antecedent. Example (13) shows the same effect. The presence of first person *wo* prevents *ziji* from referring to second person *ni*. These examples should be compared with (3) and (9), in which no blocker occurs and in which an LD relationship between reflexive and antecedent is well-formed.

As was discussed in Cole, Hermon, & Sung (1993), Cole & Sung (1994), and Cole & Wang (1996), the blocking effect appears to be limited to languages that fail to manifest subject-verb agreement. For instance, it occurs in Malayalam (Jayaseelan 1998), which has no verb agreement, but not in Kannada (Amritavali 1991), which has. The Blocking Effect also fails to occur in Italian, Icelandic etc., which have verb agreement. Thus, the blocking effect appears to be a typological property of languages which have LD reflexives and which fail to have verb agreement.

In contrast to Mandarin Chinese and Malayalam, the blocking effect fails to hold in Malay:

(14) *Aminah_i tahu saya memberi Siti_j buku tentang dirinya_{i/j/k}.*

Aminah know I give Siti book about self.3SG
 'Aminah knew I gave Siti a book about her/herself.'

(15) [*Ali dan Siti*]_i memberitahu saya yang [*Mohamed dan Ali and Siti* tell me that Mohamed and Fatimah]_j menulis rencana tentang diri mereka_{i/j/k}. Fatimah write article about self they 'Ali and Fatimah told me that Mohamed and Fatimah wrote an article about them/themselves.'

Examples (14) and (15) show that despite the intervening first person subject, *dirinya* in (14) and *diri mereka* in (15) can refer to the matrix subject.¹⁴

To summarize, Malay LD reflexives fail to obey the typological constraints which have been observed to hold of LD reflexives in other languages: They are bimorphemic while those in other languages are monomorphemic, they lack subject orientation and they do not exhibit the blocking effect.

3. Domain parameterization

We have shown that the Malay facts just summarized constitute a serious challenge both to contemporary theories of LD reflexives and to typological claims about their distribution. On the face of it, these facts, from a typological point of view, might seem to suggest that there are two distinct types of LD reflexives, one exhibiting the distributional pattern observed for Mandarin Chinese and the second exempt from the restrictions of that pattern. However, the apparent rarity of the pattern observed in

Malay suggests that this may not be a typologically sound claim.

Viewed from a generative perspective, the facts of Malay might seem at first to suggest the correctness of the general approach instantiated in the early generative work on LD reflexives. As was mentioned in Section 1, the line taken in those works was that the Binding Theory must be parameterized with respect to the locality domains found in a language. The parameterization proposal could be formulated as in (16):

(16) DOMAIN PARAMETERIZATION HYPOTHESIS (DPH)

The grammar of specific languages includes a statement of the domain of locality (Governing Category) required by the Binding Theory for that language; for Malay this is the root clause.¹⁵

Thus, Malay would differ from other languages (including Mandarin Chinese) in that the local domain is the root clause rather than the local clause. However, the parameterization of grammatical principles proposed in (16) is suspect since it runs directly counter to current trends in Generative Grammar such as the Minimalist Program, according to which all differences among languages are lexical, and grammatical principles are universal and invariant.

Putting theoretical issues aside, is (16) empirically the best way to capture how Malay differs from other languages? We shall argue that it is not. The DPH makes

the claim that in Malay in general the local domain for a nominal element will be the root clause. It can be shown, however, that (16) is *prima facie* false for Malay in general, and would have to be restricted to the specific form *diri*+PRONOUN. In addition to *diri*+PRONOUN, the lexicon of Malay includes personal pronouns (*dia* 'third person pronoun', *-nya* 'clitic form of *dia*') and an additional reflexive form, *diri*+PRONOUN *sendiri*.¹⁶ What is the local domain for each of these elements, the root clause, as predicted by the DPH, or the immediate clause?

Let us first consider the locality domain of the pronoun *dia*. As was mentioned earlier, pronouns, unlike reflexives, are restricted AGAINST occurring with c-commanding antecedents within their locality domain (while reflexives are REQUIRED to occur within that domain). Thus, (17) shows that the antecedent for *dia* cannot be the subject of the immediate clause; in (18) it is shown that the subject of a higher clause or an element in the discourse can be the antecedent for *dia*; finally, (19) shows that a non-c-commanding NP in subject position of the immediate clause is a possible antecedent for *dia*.

(17) *Ali_i nampak dia*_{i,j}*.¹⁷

Ali see he/she¹⁸

'Ali saw him/her.'

(18) *Fatimah_i tahu Ali_j mengajar dia_{i,*j,k}*.

Fatimah know Ali teach he/she

'Fatimah knows Ali saw him/her.'

(19) [*Bapa Ali*]_i nampak *dia**_{i,j,k}.

father Ali see he/she

'Ali's father saw him/her.'

The distribution of *dia* in (17)-(19) is just the same as that of the personal pronoun in English, which suggests that the locality domain for pronouns is like that of English rather than like (16). According to (16), the locality domain for Malay is the root clause. Thus, (16) would predict that the pronoun *dia* could not have a c-commanding antecedent anywhere in the root clause, that is, anywhere in the sentence. Thus, according to the DPH the reading of (18) in which *Fatimah* is the antecedent for *dia* should be ruled out, an incorrect prediction. The DPH clearly makes the wrong predictions with respect to Malay pronouns.

Turning to *diri*+PRONOUN *sendiri*, the restrictions on this form are similar to those on *himself/herself* in English. The antecedent for *diri*+PRONOUN *sendiri* must be a c-commanding NP in (roughly) the same clause as the reflexive:

(20) *Budak lelaki*_i itu nampak *dirinya sendiri*_{i,*j}.

child boy that see self.3SG alone

'That boy saw himself.'

(21) *Fatimah*_i tahu *Ali*_j mengajar *dirinya sendiri**_{i,j}.

Fatimah know Ali is.teaching self.3SG alone

'Fatimah knows that Ali is teaching himself.'

(22) [*Ibu Ali*]_i telah menyiram air di *dirinya sendiri*_i,
**j*. mother Ali already splash water on self.3sg alone
'Ali's mother splashed water onto herself.'

Example (20) shows that *dirinya sendiri* requires an antecedent within the sentence in which it occurs; (21) shows that the antecedent must be in the same clause as *dirinya sendiri*; and (22) shows that the antecedent must c-command the reflexive. This is the distribution predicted if *dirinya sendiri* has the same locality domain as do reflexives in English. If, however, the Governing Category for *dirinya sendiri* were the root clause, *Fatimah* in (21) would be a possible antecedent.

To summarize, examples (17)-(22) show that (16) cannot be a correct characterization of the locality restriction on binding in Malay in general, and would have to be restricted to the specific form *diri*+PRONOUN.

Before moving on to our own analysis, it may be useful to recapitulate briefly the facts which require explanation: (i) how *diri*+PRONOUN can have a local reflexive interpretation; (ii) how *diri*+PRONOUN can have a long distance interpretation despite the fact that it is not monomorphemic; (iii) why *diri*+PRONOUN is not subject oriented; and (iv) why Malay does not exhibit the blocking effect despite the fact that the language does not manifest verb agreement. We shall next show how these peculiarities of Malay are predicted correctly by an analysis which is fully compatible with head movement and which does not

contradict the overall typology of LD reflexives (as does the parameterization of the binding domain).

4. Feature conflation

We shall now turn to another, and, we hope, more adequate account of the distribution of *diri*+PRONOUN. We shall propose that, contrary to first appearances, *diri*+PRONOUN is not, in fact, a LD reflexive, but rather is a single form which is indeterminate between a local reflexive and a pronoun.

(23) FEATURE CONFLATION HYPOTHESIS (FCH)

The lexical item *diri*+PRONOUN is indeterminate between a pronoun and a reflexive.

This is to say, *diri*+PRONOUN is unspecified in the lexicon for the features [anaphor] or [pronoun]. For the sake of concreteness, we shall assume that these features are filled in when *diri*+PRONOUN is inserted in a sentence, and that the feature content restricts the interpretation of the sentence. We make this assumption because it facilitates a presentation in accord with the Binding Theory. On this view, a given instance of *diri*+PRONOUN can be either a pronoun or an anaphor, but no single instance of *diri*+PRONOUN can be both a pronoun and a reflexive.¹⁹

According to (23), *diri*+PRONOUN can be inserted into a sentence with the feature content of a reflexive or that of a pronoun. When *diri*+PRONOUN is inserted as a reflexive, it

must be bound in its Governing Category, which is taken to be roughly the clause, the same as that for *himself* in English, *ta ziji* in Mandarin Chinese etc. When, however, it is inserted as a pronominal, it must be free in its Governing Category.

Let us consider the predictions of the FCH with regard to the distribution of *diri*+PRONOUN. When *diri*+PRONOUN is inserted as [+anaphor, -pro], it is predicted by (23) that it must be bound in its local domain, roughly the immediate clause. Thus, the feature configuration [+anaphor, -pro] would predict correctly the local, reflexive use of *diri*+PRONOUN.

In contrast, when *diri*+PRONOUN is inserted as [-anaphor, +pro], it is predicted that as a pronoun it must take an antecedent outside its Governing Category, thereby explaining how a form which is not monomorphemic can receive a long distance interpretation. The fact that *diri*+PRONOUN is not subject oriented would be explained by the fact that pronouns do not display subject orientation.

Finally, the absence of the blocking effect would be explained similarly to the absence of subject orientation: because "long distance" *diri*+PRONOUN has the feature content of a pronoun, no blocking effect would be expected.

Thus, according to the FCH (23), the distribution of *diri*+PRONOUN follows directly from the fact that it is indeterminate between a local reflexive and a pronoun. As such, it is well-formed in the union of the environments in which local reflexives and pronouns are well-formed. Since it is not an LD reflexive, it would not be subject to the

typological restrictions specific to LD reflexives (monomorphemicity, subject orientation, and blocking effect).

5. Predictions regarding possible antecedents

We have shown that the Domain Parameterization Hypothesis (16) and the Feature Conflation Hypothesis (23) both predict the distributional facts about *diri*+PRONOUN which we have seen so far. The FCH has the advantage of restricting the irregularity to the indeterminacy of feature content of a particular lexical item. In contrast, the DPH must allow languages and lexical items to differ with respect to which grammatical principle is employed for the specification of the locality restrictions on reflexives: The domain can vary from highly restricted (the immediate clause) to unrestricted (the sentence).

While, all things being equal, Feature Conflation would be preferred on theoretical grounds, it is clearly desirable to determine empirically which hypothesis is more adequate. There is, in fact, at least one empirical difference in the predictions of the two hypotheses, one which, we believe, strongly favors Feature Conflation. It will be remembered that according to the DPH, all instances of *diri*+PRONOUN are reflexives, while according to the FCH *diri*+PRONOUN is indeterminate with regard to being a reflexive (and taking local antecedents) or a pronominal (hence taking LD antecedents). On the basis of both theory and typological evidence available from other languages, it

would be expected that reflexives would require a c-commanding antecedent within the same sentence, while pronouns would not require either a c-commanding antecedent or that the antecedent be contained in the same sentence as the pronoun. Thus, it should be possible to distinguish empirically whether "long distance" *diri*+PRONOUN behaves like a pronoun or a reflexive.

Turning to the facts of Malay, we saw previously that the local reflexive *diri*+PRONOUN *sendiri* requires a c-commanding antecedent in the same sentence, but the pronoun *dia* does not. Further examples are provided in (24)-(25):

(24) [_{NP} *Bapa Siti_i*]_j *tidak suka dirinya sendiri**_{i/j/*k}.

father Siti not like self.3SG alone

'Siti's father does not like himself.'

(25) [_{NP} *Ibu Ali_i*]_j *telah mengenalkan dirinya sendiri**_{i/j/*k}

mother Ali already introduce self.3SG alone
kepada kawan-kawan saya.

to friends my

'Ali's mother has introduced herself to my friends.'

These examples show that only the c-commanding NP (the father in (24) and the mother in (25)) can serve as antecedent for the reflexive *dirinya sendiri*. Furthermore, antecedents outside the sentence are not possible.

In contrast, the examples of (26)-(27) show that c-command is not required for the pronoun *dia* and that antecedents outside the sentence are possible:

(26) [*Bapa Ali*]_i nampak *dia**_{i/j/k}.

father Ali see he

'Ali's father saw him.'

(27) [*Ibu Ali*]_i telah mengenalkan *dia**_{i/j/k}

mother Ali already introduce he

kepada kawan-kawan saya.

to friends my

'Ali's mother has introduced him to my friends.'

In (26) and (27) *dia* can refer not only to the non-c-commanding possessor of the subject but also to an extra-sentential antecedent. Thus, in Malay, as well as in English and other familiar languages, a c-commanding antecedent within the sentence is required for reflexives but not for pronouns, and is impossible for pronouns within the local domain.

Examining now the predictions of Domain Parameterization and Feature Conflation regarding whether *diri*+PRONOUN requires a c-commanding antecedent, we see in (28) that *dirinya*, unlike *dirinya sendiri*, does not require a c-commanding antecedent:

(28) [_{NP} *Bapa Siti*]_i tidak suka *dirinya**_{i/j/k}.

father Siti not like self.3sg

'Siti's father does not like her/himself/him.'

(29) [_{NP} Ibu Ali]_i telah mengenalkan dirinya_{i/j/k}
 mother Ali already introduce self.3SG
 kepada kawan-kawan saya.
 to friends my
 'Ali's mother has introduced him/herself/her to my
 friends.'

Examples (28) and (29) show that unlike *dirinya sendiri*, *dirinya* can refer to a non-c-commanding antecedent (*Siti* in (28) and *Ali* in (29)). Furthermore, like the pronoun *dia*, but unlike the reflexive *dirinya sendiri*, *dirinya* can refer to an individual not specified in the sentence. Thus, *dirinya* in (28) might refer to some individual other than *Siti* or her father, e.g. *Ali*. Similarly, in (29) *dirinya* might refer to someone other than *Ali* or his mother, e.g. to *Siti*. Thus, the antecedents possible for *dirinya* are the union of the possible antecedents for the local reflexive *dirinya sendiri* and the pronoun *dia*. These facts are predicted by the FCH but not by the DPH.

6. Typological considerations

We have argued that the distribution of *diri*+PRONOUN in Malay is compatible with the hypothesis that *diri*+PRONOUN is indeterminate between reflexive and pronominal interpretations but incompatible with the hypothesis that *diri*+PRONOUN is a LD reflexive like Mandarin Chinese *ziji*. In fact, a state of affairs similar to that which we

propose for Malay is attested in other languages, and, thus, is not surprising from a typological perspective. For example, in Spanish the morphological distinction between pronouns and reflexives is restricted to third person; in all other persons, pronouns and reflexives have identical form.²⁰

(30)	PERSONAL PRONOUNS	REFLEXIVES
1SG	<i>me</i>	<i>me</i>
2SG	<i>te</i>	<i>te</i>
3SG.MASC	<i>lo</i>	<i>se</i>
3SG.FEM	<i>la</i>	<i>se</i>
1PL	<i>nos</i>	<i>nos</i>
2PL	<i>os</i>	<i>os</i>
3PL.MASC	<i>los</i>	<i>se</i>
3PL.FEM	<i>las</i>	<i>se</i>

Sentences (31) and (32) illustrate the use of the first person reflexive and non-reflexive pronouns, respectively.

(31) *(Yo) me veo en el espejo.*

I me see in the mirror
'I see myself in the mirror.'

(32) *María me vio en la tienda.*

María me saw in the shop
'María saw me in the shop.'

Thus, in the first and second persons, the analysis proposed for Malay *diri*+PRONOUN holds for Spanish as well.

In fact, this pattern is found not only in Spanish, but also in other Romance languages and in many Germanic languages. For example, in German *mich* 'me, myself' and *dich* 'you, yourself' are used both as reflexives and as pronouns (accusative). Only in third person are pronouns (*ihn* 'him', *sie* 'her', *es* 'it') distinguished from reflexives (*sich* 'himself').²¹

7. Conclusions

We have argued that, contrary to initial appearance, LD reflexives in Malay do not constitute counter examples to the typological generalizations which have been drawn on the basis of studies of a number of languages. Rather, the facts of Malay provide evidence that, in the Malay lexicon, pronouns and bound anaphors have been conflated, in a fashion similar to, but more extreme than, the conflation of pronouns and bound anaphors in Romance and Germanic.

It may be useful to summarize the types and properties of anaphoric elements, the existence of which have been argued for in this paper.

- (i) LOCAL (SHORT DISTANCE) REFLEXIVE
 - c-commanding antecedent;
 - antecedent in same local domain (Governing Category);
 - not subject oriented;
 - may be polymorphemic (XP)
- (ii) LONG DISTANCE REFLEXIVE

c-commanding antecedent;
antecedent external to local domain;
subject oriented;
blocking effect in appropriate languages;
monomorphemic (X^0)

(iii) PRONOUN

c-commanding antecedent not required;
cannot be c-commanded by antecedent within local
domain; not subject oriented;
no blocking effect;
XP or X^0

(iv) INDETERMINATE ANAPHORS

not subject to either c-command requirement of local
reflexives or to local non-c-command requirement of
pronouns;
not subject oriented;
no blocking effect;
XP or X^0

We have argued that Malay is unusual only in that the lexicon of the language contains indeterminate anaphors as well as the more familiar pronouns, and reflexives (local and LD). LD reflexives and indeterminate anaphors like Malay *diri*+PRONOUN, while superficially similar, can be distinguished in that the former require c-commanding antecedents (as in Mandarin Chinese) while the latter permit both non-c-commanding and extra sentential antecedents. While indeterminate forms are not widely attested, claims similar to those made here for *diri*+PRONOUN

have been made independently for analogous forms in Turkish (Kornfilt 1997).

We would like to conclude on a cautionary note. This paper shows the importance of the detailed study of apparent counter examples to typological generalizations. The status of *diri*+PRONOUN in Malay only becomes clear when it is considered in relation to other anaphoric forms in the language. Thus, general typological surveys must be complemented by detailed studies of languages that appear to be typologically anomalous.

Received: 30 December 1996

University of

Delaware

Revised: 12 September 1997

References

Amritavali, R. (1991). Lexical anaphora in Kannada.

Unpublished manuscript, Hyderabad: Central Institute of English as a Foreign Language.

Battistella, Edwin (1989). Chinese reflexivization: A

Movement to Infl approach. *Linguistics* 27: 987-1012.

Chomsky, Noam (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

Cole, Peter, Gabriella Hermon, & Li-May Sung (1990).

Principles and parameters of long-distance reflexives. *Linguistic Inquiry* 21: 1-22.

— (1993). Feature percolation. *Journal of East Asian Linguistics* 2: 91-118.

- Cole, Peter & Li-May Sung (1994). Head movement and long distance reflexives. *Linguistic Inquiry* 25: 355-406.
- Cole, Peter & Chengchi Wang (1996). Antecedents and blockers of long distance reflexives. *Linguistic Inquiry* 27: 357-390.
- Faltz, Leonard M. (1985). *Reflexivization: A Study in Universal Syntax*. New York: Garland.
- Giorgi, Alessandra (1984). Toward a theory of long distance anaphora: A GB approach. *The Linguistic Review* 3: 307-359.
- (1991). Prepositions, binding and q-marking. In Koster & Reuland (1991), 185-208.
- Hellan, Lars (1991). Containment and connectedness anaphors. In Koster & Reuland (1991), 27-48.
- Hermon, Gabriella (1994). Long-distance reflexives in universal grammar: Theoretical approaches and predictions for acquisition. In Barbara Lust et al. (eds.), *Syntactic Theory and First Language Acquisition: Cross-linguistic Perspectives. Vol. 2: Binding, Dependencies, and Learnability*, 91-112. Hillsdale, N.J.: Lawrence Erlbaum.
- Hestvik, Arild (1990). LF-movement of pronouns and the computation of binding domains. Doctoral dissertation, Brandeis University.
- (1991). Subjectless binding domains. *Natural Language and Linguistic Theory* 9: 455-496.
- (1992). LF movement of pronouns and antisubject orientation. *Linguistic Inquiry* 23: 557-594.

- Huang, C. T. James & Jane Tang (1989). The local nature of long-distance reflexives in Chinese. In *Proceedings of NELS 19*, 191-206. Amherst: Graduate Linguistic Student Association, University of Massachusetts.
- Huang, Yan (1996). A note on the head-movement analysis of long-distance reflexives. *Linguistics* 34: 833-840.
- Huang, Y.-H. (1984). Reflexives in Chinese. *Studies in Literature and Linguistics* 10.
- Jayaseelan, K. A. (1997). Anaphors as pronouns. *Studia Linguistica* 51.
- (1998). Blocking effects and the syntax of *Taan*. *Yearbook of South Asian Languages and Linguistics*.
- Kang, Beom-mo (1988). Unbounded reflexives. *Linguistics and Philosophy* 11: 415-456.
- Katada, Fusa (1989). What can long-distance anaphora say about operator systems in syntax? In *Proceedings of NELS 19*, 249-263. Amherst: Graduate Linguistic Student Association, University of Massachusetts.
- Kornfilt, Jaklin (1997). Long distance reflexives in Turkish. To appear in the *Proceedings of the Workshop on Long Distance Reflexives*, LSA Linguistic Institute, Cornell University, July 6, 1997.
- Koster, Jan (1985). Reflexives in Dutch. In Jacqueline Gueron, Hans Georg Obenauer, & Jean Yves Pollock (eds.), *Grammatical Representation*, 141-167. Dordrecht: Foris.
- Koster, Jan & Eric Reuland (eds.) (1991). *Long-Distance Anaphora*. Cambridge, Mass.: MIT Press.

- Kuno, Susumo (1972). Pronominalization, reflexivization and direct discourse. *Linguistic Inquiry* 3: 161-195.
- (1986). Anaphora in Japanese. In Sige-Yuki Kuroda (ed.), *Working Papers from the First SDF Workshop in Japanese Syntax*, 11-70. San Diego: Department of Linguistics, University of California.
- (1987). *Functional Syntax: Anaphora, Discourse, and Empathy*. Chicago: University of Chicago Press.
- Manzini, Rita (1991). Locality, parameters and some issues in Italian syntax. In Koster & Reuland (1991), 209-229.
- Manzini, Rita & Kenneth Wexler (1987). Parameters, Binding Theory, and learnability. *Linguistic Inquiry* 18: 413-444.
- Napoli, Donna Jo (1979). Reflexivization across clause boundaries in Italian. *Journal of Linguistics* 15: 1-27.
- Norhaida, Aman (1997). On the unity or non-unity of *ber*. To appear in the *Proceedings of the First Symposium on Malay and Indonesian Linguistics*, presented January 15, 1997, Penang, Malaysia.
- Park, Sung-Hyuk (1985). Pronominal and anaphoric elements in Korean. Doctoral dissertation, University of Texas at Austin.
- Pica, Pierre (1987). On the nature of the reflexivization cycle. In *Proceedings of NELS 17*, vol.2, 483-499. Amherst: Graduate Linguistic Student Association, Department of Linguistic, University of Massachusetts.

- Progovac, Ljiljana (1993). Long-distance reflexives: Movement-to-Infl versus relativized SUBJECT. *Linguistic Inquiry* 24: 755-72.
- Rappaport, G. (1986). On anaphor binding in Russian. *Natural Language and Linguistic Theory* 4: 97-120.
- Sigurjonsdottir, Sigríður (1993). Binding in Icelandic: Evidence from language acquisition. Doctoral dissertation, University of California at Los Angeles.
- Sung, Li-May (1990). Universals of reflexives. Doctoral dissertation, University of Illinois at Urbana-Champaign.
- Tang, C. T. Jane (1989). Chinese reflexives. *Natural Language and Linguistic Theory* 7: 93-121.
- Thráinsson, Huskulður (1991). Long-distance reflexives and the typology of NP's. In Koster & Reuland (1991), 49-75.
- Toman, Jindrich (1991). Anaphors in binary trees: An analysis of Czech reflexives. In Koster & Reuland (1991), 151-170.
- Vikner, Sten (1985). Parameters of binder and of binding category in Danish. *Working Papers in Scandinavian Syntax* 23.
- Wexler, Kenneth & Rita Manzini (1987). Parameter and learnability in Binding Theory. In Thomas Roeper & Edwin Williams (eds.), *Parameter Setting*, 41-89. Dordrecht: Reidel.
- Wilt, Koos van der (1985). Long distance anaphora in Dutch and Icelandic: A remark on learnability. *Linguistic Analysis* 15: 177-186.

- Yang, Dong-Whee (1983). The Extended Binding Theory of anaphors. *Language Research* 19: 169-192.
- (1988). On anaphor movement. In *Proceedings of NELS 18*. Amherst: Graduate Linguistic Student Association, University of Massachusetts.

Notes

The research reported on in this paper is part of a larger study of long distance reflexives across languages supported by the National Science Foundation (SBR-9122267 and INT-9423291).

The research discussed here was presented at project meetings at the National University of Singapore and the Universiti Sains Malaysia in Penang, Malaysia. We would like to thank those present at the meetings for their very helpful comments. Special thanks are due to Mashudi Kader and Norhaida Aman. We would also like to acknowledge the exceptionally helpful comments provided by two *Linguistic Typology* reviewers.

Correspondence addresses: Department of Linguistics (Cole, Hermon), Department of Educational Studies (Hermon), University of Delaware, Newark, DE 19716, USA; e-mail: pcole@udel.edu, gaby@udel.edu

¹ For the sake of concreteness, we are assuming the version of the Binding Theory in Chomsky (1986), but the details of the version of the Binding Theory are not important for the issues discussed in this paper.

² For earlier works in which an LF movement approach was employed, see Pica (1987), Battistella (1989), Tang (1989), Huang & Tang (1989), Sung (1990), Cole, Hermon, & Sung (1993), Hermon (1994), among others.

For other generative approaches to LD reflexives, see Napoli (1979), Manzini & Wexler (1987), Wexler & Manzini (1987), Kang (1988), Katada (1989), Progovac (1993), and the papers in Reuland & Koster (1991). A number of the papers in Koster & Reuland (1991) deal with logophoric and other discourse conditions on LD reflexives, a topic which we will ignore in the present paper.

³ In order to simplify the presentation, tree (4) does not show movement to C, which would be required by the Head Movement Constraint.

⁴ Cole, Hermon, & Sung (1990), Cole & Sung (1994), and Sung (1990) provide analyses which show (i) why long distance movement of phrasal reflexives is not expected and (ii) why long distance movement of the reflexive head out of the phrasal reflexive should not be possible. Our emphasis here is on the typological predictions of the approach, so we will not review the technical details.

⁵ See Cole & Sung (1994) and Cole & Wang (1996) for a detailed demonstration of how head movement makes these predictions. These predictions are illustrated in greater detail in Section 2.

Note that it is not necessary to parameterize whether a language employs head movement with reflexives. Head movement is predicted to occur whenever the reflexive is an X^0 element (e.g. Mandarin Chinese *ziji*) rather than a phrasal element (e.g. English *himself*, Mandarin Chinese *ta ziji* 'himself'). Of course, it is not always trivial to determine whether a given reflexive form is an X^0 or an XP.

⁶ It is shown in Cole & Wang (1996) that, at least with respect to Mandarin Chinese, subject orientation can be reduced to c-command, and that non-subject antecedents for reflexives are possible just when the noun phrase in question c-commands AGR.

⁷ This prediction was first made in Cole, Hermon, & Sung (1993), which provides a detailed discussion of the relationship of the Blocking Effect and subject-verb agreement.

⁸ See also Faltz (1985) and the papers in Koster & Reuland (1991). We would like to emphasize that the authors mentioned represent a diverse range of views. Thus, we make no claims with respect to the degree to which they would agree with the opinions expressed in this paper.

⁹ Our analysis of Malay is based on the judgments of educated Malay speakers in Singapore. Our factual claims pertain to this dialect, and we make no claims with respect to the distribution of reflexives in formal Bahasa Malaysia or in other dialects of Malay. Preliminary investigation suggests that there may be considerable dialectal variation along both social and geographical lines, a topic which we hope to investigate later.

¹⁰ The paradigm for Malay reflexives employing *diri* 'self' + PRONOUN in the dialect under discussion is as follows:

1SG *diri saya*
diri aku (less formal)

2SG	<i>diri kamu</i>	(other forms of 'you' may be substituted)
3SG	<i>diri-nya</i>	(the cliticized form is much more frequent)
	<i>diri dia</i>	
1PL	<i>diri kami</i>	(no distinction between inclusive and exclusive in this dialect)
	<i>diri kita</i>	
2PL	<i>diri kamu</i>	(other forms of 'you' may be substituted)
3PL	<i>diri mereka</i>	

Our informants were familiar with but did not feel comfortable using cliticized forms other than the third person singular *dirinya*.

All of the above forms can be used with local or non-local antecedents. For our informants, *dirinya* 'self.3SG' is restricted to third person singular antecedents. Thus, when used with reflexives, the morpheme *-nya* '3SG possessor' does not seem to have undergone the semantic "bleaching" which allows *buku-nya* (book-3SG) to mean either 'his/her book' or 'the book'. If *-nya* in *dirinya* were similarly "bleached" of person, *dirinya* should be able to refer to non-third-person antecedents. But sentences like (i) were rejected by our informants:

(i) **Saya_i tahu Salmah_j akan membeli baju untuk diri-nya_i.*

I know Salmah will buy clothes for self-3SG
'I know Salmah will buy clothes for him/herself.'

¹¹ As was seen above, this is a slight oversimplification. In order to undergo head movement, an element must be a head, and, hence X^0 . A compound should in principle be able to undergo head movement as long as it is dominated by an X^0 node.

¹² Locality appears to be due to polymorphicity rather than to the presence of person features: In Italian the reflexive *se* is limited to third person but it can take a LD antecedent. What distinguishes *se* from English *himself* and Malay *dirinya* is that *se* is a single morpheme while *himself* and *dirinya* are bimorphemic. It is precisely this distinction which is predicted by the head movement analysis, which claims that LD reflexives must be X^0 (in order to undergo head movement), but which makes no claims regarding the feature content of the reflexive morpheme.

¹³ Spoken Mandarin Chinese *ta* is not marked for gender or case. Thus, *ta ziji* can mean either 'himself' or 'herself'. For the sake of simplicity, we shall gloss *ta* as 'he'.

¹⁴ Example (15) also illustrates the fact that LD reflexives are possible with forms other than the third person clitic *-nya*: The LD reflexive in (15) is *diri mereka*, a non-cliticized form.

¹⁵ The DPH is a strawman and is not identical with the position taken in Manzini & Wexler (1987). Manzini and Wexler were aware of the inadequacies of approaches like (16), and proposed that each pronoun or anaphor would carry

its own specification of binding domain. See Sung (1990), Hermon (1994), and Cole & Sung (1994) for arguments showing that Manzini and Wexler, nonetheless, fail to capture the fact that in languages like Mandarin Chinese the properties in (5) go together.

¹⁶ Malay exhibits a number of additional reflexive forms, among them *diri*, *diri sendiri*, *sendiri*, and the verbal prefix *ber-*. We discuss the properties of *diri*+PRONOUN *sendiri* below. Of the remaining forms, *sendiri* alone is not used as a reflexive by our informants:

(i) **Ali nampak sendiri*.

Ali see alone

'Ali sees himself.'

It may be useful to say a few words about the remaining forms (despite the fact that none of them could be considered LD reflexives). The forms *diri* and *diri sendiri* are restricted to specific lexical items like *bunuh* 'kill':

(ii) *Dia cuba membunuh diri (sendiri)*.

he try TRANSITIVE.kill self (alone)

'He tried to kill himself.'

(iii) **Dia nampak diri (sendiri)*.

he see self (alone)

'He saw himself.'

Finally, *ber-* has a variety of uses, among them some translatable as reflexives:

(iv) Saya bercukur.

I INTRANSITIVE.shave

'I shaved myself.'

While *ber-* might appear on initial examination to be a reflexive morpheme, there is reason to believe that it is not and that such *ber-* forms are special cases of a more general non-reflexive use of *ber-*. See Norhaida (1997).

¹⁷ The indices *j* in (17) and *k* in (18) and (19) indicate that the reference of the pronoun *dia* can be an individual not represented in the sentence but, rather, determined by the discourse. The same convention will be followed below.

¹⁸ Like Mandarin Chinese *ta*, *dia* is indeterminate for gender. Additionally, Malay pronouns (and other nominals) do not vary according to case. Thus, we gloss *dia* as 'he/she'.

¹⁹ To these best of our knowledge, this proposal is empirically indistinguishable from the proposal that *diri*+PRONOUN is inserted without any specification for the features [anaphor] and [pronoun]. On the assumption that an element unspecified for these features would be exempt from the Binding Theory, it would, therefore, be well-formed in the union of the environments possible for pronouns and reflexives. It should be noted, however, that if *diri*+PRONOUN is taken to be unspecified for the features

[anaphor] and [pronoun], referring expressions will have to be specified as [-pronoun] and [-anaphor] in order to distinguish them from *diri*+PRONOUN. If referring expressions were also unspecified for [anaphor] and [pronoun], they would be expected to have the same anaphoric properties as *diri*+PRONOUN. Since referring expressions, unlike *diri*+PRONOUN, do not permit antecedents (in accord with the Binding Theory), this would be an incorrect expectation.

²⁰ The forms given in (30) are clitics, accusative in the case of the personal pronouns.

²¹ It might be argued that non-third person pronouns are irrelevant to the issues under discussion because they are independently referential, and, therefore, do not have the same anaphoric properties as do third person pronouns. This argument, however, would not be correct. Note that first and second person pronouns share the same distribution as third person pronouns with regard to their ability to tolerate coreferential c-commanding antecedents. Consider the wellformedness of (i), in which a first person pronoun is employed, while in similar sentences employing proper nouns, such as (ii), only a non-coreferential interpretation is possible; the grammaticality of (i) would be expected if the second instance of *I* has the grammatical properties of the pronoun *he* in (iii).

(i) *I_i said that I_i would go.*

(ii) *John_i said that John_j/*_i would go.*

(iii) *John_i said that he_{i/j} would go.*

If, however, first person pronouns had the grammatical properties of proper nouns, it would be expected that (i) is ungrammatical.