

DECOMPOSING MALAY ANAPHORIC EXPRESSIONS*

Hiroki Nomoto
Tokyo University of Foreign Studies
nomoto@tufs.ac.jp

Previous studies on anaphoric expressions in Malay centre on two forms, i.e. ‘*diri* + pronoun’ and ‘*diri* + pronoun + *sendiri*’, and analyse them non-compositionally. This paper shows that a compositional analysis of Malay anaphoric expressions is not only possible but is empirically more desirable than a non-compositional analysis, as it can account for a wider range of anaphoric expressions in a systematic manner.

1. Introduction

Previous studies on anaphoric expressions in Standard Malaysian/Singapore Malay (e.g. Cole and Hermon 1998, 2005; Nomoto 2011) centre on two forms: (i) ‘*diri* + pronoun’ and (ii) ‘*diri* + pronoun + *sendiri*’. They are non-compositional in that they treat these multimorphemic forms as single lexical items/words on a par with English reflexives such as *himself*. Consequently, the anaphoric properties of the relevant forms are ascribed to the whole expression.

This study proposes an alternative analysis. The proposed analysis is compositional, and hence the anaphoric properties of a multimorphemic form are ascribed to its constituent parts. It is demonstrated that the proposed analysis can deal with a wider range of anaphoric expressions in Malay, which include (i) ‘*diri* + pronoun’, (ii) ‘*diri* + pronoun + *sendiri*’, (iii) ‘*diri* + non-pronoun’, (iv) ‘*diri* + non-pronoun + *sendiri*’, (v) *diri*, (vi) *diri sendiri*, (vii) ‘(non-)pronoun + *sendiri*’, and (viii) *sendiri*. The proposed analysis offers more insight into the study of anaphoric expressions in related dialects/languages (e.g. Gil 2001; Paul 2004; Davies 2008; Kartono 2013), as these dialects/languages have been reported to employ similar multimorphemic anaphoric expressions.

This paper is organized as follows. Section 2 demonstrates that Malay indeed possesses the various anaphoric expressions listed above by presenting examples

* The research reported here was supported in part by the JSPS Grant-in-Aid for Young Scientists (B) (#26770135). I thank my Malay consultants Faridah Mohamed and Kartini Abd. Wahab as well as audiences at Tokyo University of Foreign Studies, the 21st Annual Meeting of the Austronesian Formal Linguistics Association (AFLA) and the 18th International Symposium on Malay/Indonesian Linguistics (ISMIL). I am grateful to Vincent Homer, who generously shared with me the L^AT_EX source file of his AFLA 16 proceedings paper (Homer 2009).

of them. Section 3 reviews Cole and Hermon's (2005) analysis of patterns (i) and (ii) as a representative of non-compositional analyses. After pointing out problems of their analysis, I propose an alternative compositional analysis in section 4 and show how it accounts for the anaphoric properties of the various types of anaphoric expressions in Malay in section 5. Section 6 concludes the paper.

2. The Inventory of Malay Anaphoric Expressions

Besides the two types that have attracted previous researchers' attention ((i) and (ii)), Malay has many other patterns of anaphoric expressions ((iii)–(viii)). All eight patterns are easily found in naturally occurring texts. They are all made up of one or more elements from *diri*, *sendiri* and noun phrases. This simple fact already suggests the possibility of a compositional treatment. I will show in section 4 that it is possible to reduce the eight patterns to just three. Examples of the eight patterns are given below.

(i) 'diri + pronoun' It is this pattern that has intrigued previous researchers the most, as it shows a hybrid property of reflexives and pronominals. At first, one may regard it as reflexive because, in elicitation sessions, native Malay speakers will normally use this pattern to translate English expressions with reflexives such as *himself* and *herself*. In support of this initial hypothesis, the pattern allows a local c-commanding antecedent, as indicated by index *j* in (1)–(2) below. However, a closer inspection reveals that the pattern also exhibits pronominal behaviours. Thus, *dirinya* also allows non-local and non-c-commanding antecedents, as indicated by index *i* in (1) and (2) respectively. The referent indicated by index *k* is a discourse referent that is salient in the current discourse but is not referred to by any DP in the sentence.

- (1) Ali_{*i*} kata [Siti_{*j*} mengambil gambar *diri-nya*_{*i/j/k*}].
Ali say Siti take picture DIRI-3
'Ali said Siti took a picture of *him/herself/her*.'
- (2) [Bapak Siti_{*i*}]_{*j*} tidak suka *diri-nya*_{*i/j/k*}.
father Siti not like DIRI-3
'Siti's father does not like *her/himself/him*.' (Cole and Hermon 2005:631)

(ii) 'diri + pronoun + sendiri' Unlike 'diri + pronoun', this pattern shows regular reflexive behaviours. It is hence subject to Condition A of the canonical binding theory; *dirinya sendiri* must be locally bound in (3) below.¹

¹ Note that examples like (i) involve a combination of pattern (i) 'diri + pronoun' and the adverbial *sendiri* meaning 'alone, by oneself', and should not be confused with pattern (ii).

- (3) a. Ali_i kata Siti_j mengambil gambar *diri-nya sendiri*_{*i/j/*k}.
 Ali say Siti take picture DIRI-3 own
 ‘Ali said Siti took a picture of *herself*.’
- b. [Ibu Siti_i]_j mengambil gambar *diri-nya sendiri*_{*i/j/*k}.
 mother Siti take picture DIRI-3 own
 ‘Siti’s mother took a picture of *herself*.’

For patterns (iii)–(viii), I only present examples (all taken from the DBP (Dewan Bahasa dan Pustaka) Corpus).² The anaphoric properties of these patterns will be described in section 5, where I show that they can be accounted for by the compositional analysis to be proposed in section 4.

(4) **(iii) ‘*diri* + non-pronoun’**³

Segala kejadian biologikal dan fizikal ke atas *diri remaja itu*
 all incident biological and physical to top DIRI adolescent that
 [...].
 ‘All the biological and physical incidents that happen to *the adolescent*
 [...].’

(5) **(iv) ‘*diri* + non-pronoun + *sendiri*’**

Apabila tersalah memilih teman, ia akan memberi kesan yang besar
 when mistake choose friend it will give effect REL big
 terhadap *diri remaja itu sendiri*.
 towards DIRI adolescent that own
 ‘If friends are chosen wrongly, that will have a big influence on *the adolescent*.’

(6) **(v) *diri***

Itu penting, sebab dalam hidup ini keyakinan *diri* adalah 50%
 that important because in life this confidence DIRI be 50%
 daripada kemenangan.
 from victory
 ‘That is important because in life *self*-confidence is 50% of the success.’

(i) John fikir (yang) diri-nya sendiri akan pergi ke KL besok.
 John think that DIRI-3 alone will go to KL tomorrow
 ‘John thinks that he himself will go to KL (= Kuala Lumpur) tomorrow.’
 (Cole and Hermon 2005:634)

² See my AFLA handout (http://www.tufts.ac.jp/ts/personal/nomoto/handout_afla21.pdf) for more examples.

³ Note that non-pronouns are not used as pronoun substitutes in this pattern, unlike *adik* ‘younger sibling’ in (7), which is used as a substitute for the second person pronoun.

(7) (vi) *diri sendiri*

Sampai bila adik harus membohongi *diri sendiri* dan
until when younger.sibling should deceive DIRI own and
diri-nya?

DIRI-3

‘When can you stop deceiving *yourself* and him?’

(8) (vii) ‘(non-)pronoun + *sendiri*’

Bangsa Melayu di-katakan tidak mampu berfikir dalam bahasa *mereka*
ethnic Malay PASS-say not able think in language their
sendiri.

own

‘Ethnic Malays are said to be unable to think in *their own* language.’

(9) (viii) *sendiri*

Di samping itu, responden tidak pasti sama ada mereka memiliki
at side that respondent not certain whether they have
sikap bangga terhadap bahasa *sendiri*.

attitude proud towards language own

‘Moreover, the respondents are not certain whether they take pride in *their own* language.’

3. Non-compositional Analysis: Cole and Hermon (2005)

In this section, I briefly review Cole and Hermon’s (2005) study as a representative of non-compositional analyses of Malay anaphoric expressions. As noted at the outset of this paper, Cole and Hermon focus on two patterns, i.e. (i) ‘*diri* + pronoun’ and (ii) ‘*diri* + pronoun + *sendiri*’. They do so presumably for typological considerations rather than descriptive considerations concerning the anaphoric system of a specific language. In particular, pattern (i) in Malay could pose a serious problem to some rather solid typological generalizations about long-distance reflexives if it were actually a reflexive (Cole and Hermon 1998, 2005). They conclude that pattern (i) is in fact not a reflexive, and hence does not affect the relevant typological generalizations. This conclusion is based on the following analysis of pattern (i).

In Cole and Hermon’s analysis, “*dirinya* is not treated by the syntax as a complex DP [...], but rather as a unitary lexical entry” (643). They assume the existence of a paradigm of ‘*diri* + pronoun’ shown in Table 1, which is reminiscent of the English reflexive pronoun paradigm.

Cole and Hermon claim that “*diri* + pronoun is unspecified in the lexicon with regard to the features [α anaphor] and [α pronominal]” (631). In other words, pattern (i) is neither a reflexive nor a pronominal; it has properties of both. This unspeci-

Table 1: The full paradigm of ‘*diri* + pronoun’ (Cole and Hermon 2005:629)

Person	Singular	Plural
First Person	diri saya/diri-ku	diri kami/diri kita
Second Person	diri kamu/diri-mu	diri kamu/diri-mu
Third Person	diri-nya	diri mereka/diri-nya

fied feature analysis is able to account for the hybrid property of ‘*diri* + pronoun’ that we saw in the last section (cf. (1)).

Given this analysis, a similar analysis has to be assumed for true reflexives ‘*diri* + pronoun + *sendiri*’, though Cole and Hermon are not explicit about them. That is to say, the lexicon of Malay should have ‘*diri* + pronoun + *sendiri*’ as a unitary lexical entry with a paradigm similar to Table 1. Each lexical entry must be specified in the lexicon as [+anaphor, –pronominal] to capture its canonical reflexive behaviours.

While Cole and Hermon’s study has bolstered the general theory of long-distance reflexives by explaining away the typologically unusual behaviours of reflexive-looking forms in Malay, i.e. ‘*diri* + pronoun’, it is problematic when taken as an analysis of a phenomenon in a specific language. To begin with, the various other anaphoric expressions presented in section 2 (iii)–(viii) appear to result from combinations of the constituent parts of ‘*diri* + pronoun (+ *sendiri*)’ (i)–(ii). Yet, under Cole and Hermon’s non-compositional analysis (or any non-compositional analysis for that matter), the relation between (iii)–(viii) and (i)–(ii) remains unclear. Furthermore, the pronoun slot in ‘*diri* + pronoun (+ *sendiri*)’ is in fact not restricted to pronouns, but also available for other DPs, as is the case with (iii) ‘*diri* + non-pronoun’ and (iv) ‘*diri* + non-pronoun + *sendiri*’. It is implausible to think that these patterns with non-pronouns are also unitary lexical entries, for the number of non-pronouns is infinite, unlike that of pronouns.

Given these problems, it is more desirable if a compositional analysis is possible that can account for the anaphoric properties of ‘*diri* + pronoun (+ *sendiri*)’ as well as the other anaphoric expressions. I will propose one such analysis in the next section.

4. Compositional Analysis

The proposed compositional analysis has three main ingredients. The first ingredient is the syntax and semantics of *diri*. A proper understanding of *diri* enables us to see how *diri* is involved in anaphoric expressions and why. It also leads us to the second main ingredient, i.e. the null unspecified possessive pronoun *pro*, which plays an important role in reducing the numerous different patterns into just

three and thereby greatly simplifying the description and analysis of the anaphoric expressions in Malay. The last ingredient is the semantics of the intensifier *sendiri* ‘alone, own’. I argue that the semantics of *sendiri* involves a kind of focus semantics and that the local binding property characteristic of anaphors results from it.

4.1. *Diri*

I make the following two claims about *diri*. First, *diri* is an NP that takes a possessor argument. Second, it denotes a function from an individual to that individual’s physical self:

$$(10) \quad \llbracket \text{diri} \rrbracket = \lambda x.x\text{'s physical self}$$

To put this informally, *diri* means ‘someone’s body’. Given that one’s body, constituting his/her physical self, is the entire whole inalienably possessed by him/her, one could analyse *diri* as a pseudo-identity function.⁴ In other words, the formula in (10) can be approximated as in (11), which roughly says that a person’s body is that person himself/herself.

$$(11) \quad \llbracket \text{diri} \rrbracket = \lambda x.x\text{'s physical self} \approx \lambda x.x\text{'s } x \approx \lambda x.x$$

Two possibilities exist regarding the approximations in (11). The first possibility is that they are hypothesized diachronic processes in the grammaticalization of *diri*. That is, there are two *diri* morphemes synchronically, i.e. (10) and the identity function ($\lambda x.x$ in (11)). Only the former but not the latter preserves the meaning of ‘physical self’. The other possibility is that there is only one *diri* morpheme and the approximations in (11) take place in the speaker’s mind/brain as synchronic subconscious inference processes when *diri* is used in anaphoric expressions. In what follows, I will assume the first possibility for simplicity’s sake, and treat *diri* as an identity function, unless otherwise noted. Further study is needed to determine which possibility is actually the case.

There is evidence for the ‘physical self’ meaning of *diri*. First, *diri* can be used in contrast with words such as *jiwa* ‘soul’, *hati* ‘heart’ and *batin* ‘inner self’, words representing the mental self of an individual:

- (12) a. Saya terpaksa “meremajakan semula” *diri dan jiwa* saya.
 I have.to rejuvenate DIRI and soul my
 ‘I had to rejuvenate my body and soul.’

⁴ Kartono (2013) glosses *diri* in Indonesian, Palembangese and Jambi as ‘body’. This is not a precise description of *diri* in Malay (and probably even in these languages). It is *badan* and *tubuh* that refer to ‘body’. *Diri* is not interchangeable with these words, as it is an abstract notion reflecting the Malay perception of the self.

- b. [...] cuba-lah fahami lagi siapa Ade ini, fahami
 try-PARTICLE understand more who Ade this understand
diri dan hati Ade.
 DIRI and heart Ade
 ‘[...] try to understand more who I am, to understand my physical
 self and heart.’ (DBP Corpus)

Second, if only the physical aspect of an individual changes, it is a change in *diri*:

- (13) Ular itu telah kembali kepada *diri* asal-nya, se-orang putera raja.
 snake that PRF return to DIRI original-3 one-CLF prince king
 ‘The snake has transformed itself back into the original self, a prince.’
 (Si Bongsu dengan Kak Nam)

These examples also lend support to the claim that *diri* is an NP. It is smaller than DP because it can be followed by a possessor DP and other modifiers. Moreover, it can be coordinated with another non-DP noun phrase.

4.2. The Null Unspecified Possessive Pronoun *pro*

The second ingredient for the compositional analysis of various anaphoric expressions in Malay is the null unspecified possessive pronoun *pro*. It is comparable to *one’s* in English, and is a part of the pronominal paradigm of the language.

Where anaphoric expressions are concerned, *pro* occurs in the following two contexts: as a possessor argument of *diri* and with the intensifier *sendiri*. With the introduction of *pro*, it becomes possible to reduce the eight different patterns into just three types, as shown in (14).

- | | | | | |
|------|--------|---|---|---------------------------------------|
| (14) | (i) | ‘ <i>diri</i> + pronoun’ | } | ‘ <i>diri</i> + DP’ |
| | (iii) | ‘ <i>diri</i> + non-pronoun’ | | |
| | (v) | <i>diri</i> = <i>diri</i> + <i>pro</i> | | |
| | (ii) | ‘ <i>diri</i> + pronoun + <i>sendiri</i> ’ | } | ‘ <i>diri</i> + DP + <i>sendiri</i> ’ |
| | (iv) | ‘ <i>diri</i> + non-pronoun + <i>sendiri</i> ’ | | |
| | (vi) | <i>diri sendiri</i> = ‘ <i>diri</i> + <i>pro</i> + <i>sendiri</i> ’ | | |
| | (vii) | ‘(non-)pronoun + <i>sendiri</i> ’ | } | ‘DP + <i>sendiri</i> ’ |
| | (viii) | <i>sendiri</i> = ‘ <i>pro</i> + <i>sendiri</i> ’ | | |

When *diri* appears to occur with no noun phrase, as in (v) and (vi), it actually contains the phonologically null DP *pro* as its complement. Likewise, when *sendiri* appears to occur with no noun phrase, as in (vi) and (viii), it actually modifies *pro*. Notice that (vi) can be parsed in two ways: [*diri* + *pro*] + *sendiri* and *diri* + [*pro* + *sendiri*]. Also notice that ‘*diri* + DP + *sendiri*’ is a combination of ‘*diri* + DP’ and ‘DP + *sendiri*’.

Now, the various anaphoric expressions in Malay can be summarized as in Table 2. Emphatic forms are so called because their denotations are virtually

Table 2: The composition of Malay anaphoric expressions

	Non-focussed	Focussed (<i>sendiri</i>)
Non-emphatic	(A) DP [_A DP] <i>prol-/nya/Ali</i> 'one's/his/her/Ali's'	(C) DP + sendiri [_C [_A DP] sendiri] <i>prol-/nya/Ali sendiri</i> 'one's/his/her/Ali's own'
Emphatic (<i>diri</i>)	(B) diri + DP [_B diri [_A DP]] <i>diri prol-/nya/Ali</i> 'one's/his/her/Ali's' physical self'	(D) diri + DP + sendiri [_D [_B diri DP] sendiri]/ [_D diri [_C DP sendiri]] <i>diri prol-/nya/Ali sendiri</i> 'one's/his/her/Ali's own' physical self'

identical to those of the corresponding non-emphatic forms (cf. (11)), but have more expression.⁵ The forms containing *sendiri* are referred to as focussed forms because the semantics of *sendiri* involves focus semantics, as we shall now see.

4.3. *Sendiri*

According to Alsagoff (1992), *sendiri* 'alone, own' (in a non-subject position) requires a local antecedent when used by itself, as in (15).

- (15) Mariam_i kata bahawa Ali_j menjual kereta *sendiri*_{*i/j}.
 Mariam say that Ali sell car own
 'Mariam said that Ali sold *her/his car.' (Alsagoff 1992:41)

Alsagoff's other example (16) and the additional example in (17) show that the antecedent of *sendiri* must c-command the NP modified by *sendiri* too. These properties seem inherent to *sendiri*.⁶

⁵ In this connection, Kartono (2013:50–53) points out that emphatic forms in Indonesian (more specifically '*diri* + pronoun' in a subject position) are used to express respect and empathy. This description seems also valid in Malay.

⁶ One might wonder whether they can be attributed to parts of *sendiri*, as *sendiri* contains *diri*. However, synchronically, *sendiri* should not be analysed into smaller parts. The etymology of *sendiri* is not clear. Zaharani Ahmad (p.c.) suggested to me that it emerged from *se-orang diri* [one-CLF DIRI] 'alone', which is highly probable, given the fact that the related language Minangkabau employs *surang*, the equivalent of *seorang* in Malay, in contexts where Malay uses *sendiri* such as *dirinya sendiri* (Yusrita Yanti, p.c.).

- (16) Ali_i mengusik doktor itu_j di rumah *sendiri*_{i/*j}.
 Ali tease doctor that in house own
 ‘Ali teased the doctor in his own house.’ (Alsagoff 1992:40)
- (17) [Ibu Ali_i]_j menjual kereta *sendiri*_{*i/j}.
 mother Ali sell car own
 ‘Ali’s mother sold her/*his own car.’

In the present analysis, *sendiri* in the examples above is in fact *pro sendiri*. Given this analysis, the role of *sendiri* is to restrict the otherwise unspecified referent of *pro* to that of its antecedent. Crucially, *sendiri* restricts possible interpretations to a reflexive one.

Let us look at the function of *sendiri* (in a non-subject position) in more detail. *Sendiri* plays a dual role, syntactic and semantic. Syntactically, it searches the sentence for the antecedent of the expression it combines with. As we have seen above, the antecedent must locally c-command *sendiri*. Semantically, *sendiri* involves a kind of focus semantics, as Gil (2001) proposes for *sendiri* in Riau Indonesian.⁷ This focus semantics has to do with its meaning translated into English as ‘own’. *Sendiri* induces a set consisting of the potential referents of the expression it combines with, and excludes from this set all members but the referent of the antecedent.

Given the expression in (18a), where DP_a (*a* for “antecedent”) denotes individual *a* ($\llbracket \text{DP}_a \rrbracket = a$), ‘DP_f *sendiri*’ induces a set of the potential referents of DP_f (*f* for “focus”), as in (18b). Call this set *F*. *F* must include *a*, otherwise the expression is ungrammatical. *Sendiri* entails that among the members of *F*, *V* applies only to *a* (18c).

- (18) a. DP_a V [OBJ NP DP_f *sendiri*]
 b. $F = \{a, b, c, d, \dots\}$
 c. $\forall x[V(x) \rightarrow x = a]$

If DP_f is a pronoun whose ϕ -features are compatible with DP_a, then the referent of DP_f becomes identical to that of DP_a ($\llbracket \text{DP}_a \rrbracket = \llbracket \text{DP}_f \rrbracket = a$). A reflexive interpretation is obtained in this case. For example, in (15) with the expression ‘Ali V [OBJ NP *pro sendiri*]’, the ϕ -features of *Ali* and *pro* are compatible with each other, as the latter is unspecified in terms of ϕ -features. Thus, the otherwise unrestricted referent of *pro* is fixed to *Ali*, giving rise to a reflexive interpretation. By contrast, if *pro* is replaced by another pronoun whose ϕ -features are incompatible with those of *Ali* as in (19), the sentence becomes ungrammatical. This is because

⁷ Gil refers to the function of *sendiri* as ‘conjunctive operator’. Moreover, the specific formulation of it to be presented below differs from Gil’s.

Ali is not in F , the set of potential referents of *saya* ‘my’.⁸

- (19) *Ali_a menjual [OBJ kereta saya_f sendiri].
 Ali sell car my own
 (*‘Ali sold my own car.’)

In the examples above, the DP that *sendiri* combines with is a pronoun. However, when *sendiri* combines with a non-pronoun DP as in *Siti sendiri* ‘Siti herself’, it works slightly differently. Specifically, the DP itself becomes the antecedent of *sendiri*. Moreover, the set F consists not of the potential referents of the DP, but of the DP’s actual referent and its alternatives, because the referent of a non-pronoun DP is already fixed to a particular individual. As a consequence of the first difference, *sendiri* does not affect the referent of the DP. Thus, I will not discuss cases involving non-pronouns below, i.e. patterns (iii) and (iv) as well as pattern (viii) with non-pronouns.

5. Accounting for Each Pattern

This section discusses how the compositional analysis proposed in the last section accounts for the various anaphoric expressions. Since one of the patterns in focussed forms (C), i.e. (viii) *sendiri*, has been already explained in the last section, I begin with focussed forms (C). I will then turn to emphatic focussed forms (D), as *sendiri* plays a crucial role here too. Emphatic forms (B), which unlike the other two forms do not involve *sendiri*, are discussed at the end.

5.1. DP + *Sendiri* (C. Focussed Forms)

As seen in section 4.3, *sendiri* restricts possible interpretations to a reflexive one, by limiting the referent of the expression it combines with to that of its antecedent, which must locally c-command it. To repeat the account of pattern (viii), while the referent of *pro* in (20a) is unconstrained, allowing both reflexive and non-reflexive interpretations, that of *pro* in (20b) is restricted to that of *sendiri*’s antecedent, i.e. Ali. As a result, only a reflexive interpretation is available in (20b).

- (20) a. Ali_i menjual kereta *pro*_{i/j}.⁹
 Ali sell car *pro*
 ‘Ali sold his/her/their/etc. car.’ (reflexive and non-reflexive)

⁸ One might wonder whether the sentence is grammatical if *Ali* is used as a substitute for the first person pronoun. Such a sentence is unacceptable or very unnatural at best for an independent reason: when a proper name is used as a substitute for a pronoun, the substitution must occur throughout the sentence consistently.

- b. Ali_i menjual kereta *pro*_{i/*j} *sendiri*.
 Ali sell car *pro* own
 ‘Ali sold his/*her/*their/*etc. own car.’ (reflexive only)

Another pattern of focussed forms, i.e. (vii) ‘pronoun *sendiri*’, differs from (viii) only in that an overt pronoun is used instead of *pro*. Hence, no additional mechanism is necessary to account for (vii). In (21a) without *sendiri*, the referent of the third person pronoun *-nya* can be either that of the local subject *Ali* or that of the matrix subject *Mariam*, or even a discourse referent that does not appear in the sentence. A reflexive interpretation is obtained in the first case whereas a non-reflexive interpretation is obtained in the second and third cases. However, the addition of *sendiri* rules out the latter interpretative possibility and forces a reflexive interpretation, as in (21b).

- (21) a. *Mariam*_i kata bahawa *Ali*_j menjual kereta-*nya*_{i/j/k}.
 Mariam say that Ali sell car-3
 ‘Mariam said that Ali sold her/his car.’ (reflexive and non-reflexive)
- b. *Mariam*_i kata bahawa *Ali*_j menjual kereta-*nya*_{*i/j/*k} *sendiri*.
 Mariam say that Ali sell car-3 own
 ‘Mariam said that Ali sold *her/his own car.’ (reflexive only)

Notice that in my analysis (the relevant use of) *sendiri* always modifies the possessor. This is obvious when the possessor is overt. Where there does not seem to be any possessor, as in pattern (viii), my analysis assumes a null DP possessor *pro*. My analysis thus accords with the following observation by Gil (2001:112): “in Standard Malay/Indonesian, the actual form *sendiri* appears to function as a reflexive only in possessive constructions.” It is worth noting here that while Malay and Modern English allow both reflexive and non-reflexive interpretations only in possessive constructions, some languages allow both interpretive possibilities for non-possessive pronominals as well (cf. Table 3). These languages include Old English (e.g. van Gelderen 2000; König and Siemund 2000; Keenan 2002), Madurese (Davies 2008), Jambi (Mudung Darat and Tanjung Raden dialects) (Cole et al. 2010), and Palembangse (Kartono 2013). Conversely, other languages distinguish between two kinds of possessive pronouns for reflexive and non-reflexive interpretations. Russian is one such language (Kazuhiro Kojima, p.c.). I am not aware of a language that has distinct reflexive and non-reflexive forms only in possessive constructions.

⁹ The same surface string has another structure without the possessor pronoun *pro*. This alternative structure is irrelevant here; it is associated with an interpretation in which the possessor is unimportant (e.g. ‘Ali sold a car.’).

Table 3: Cross-linguistic variation in the reflexive vs. non-reflexive formal distinction

	Non-possessive	Possessive
Old English, Madurese, etc.	–	–
Modern English, Malay	+	–
Russian	+	+

5.2. *Diri* + DP + *Sendiri* (D. Emphatic Focussed Forms)

As seen in section 2, (ii) ‘*diri* + pronoun + *sendiri*’ shows typical reflexive behaviours. It must be bound locally, as in (22b). This is not a property of the expression as a whole but one of *sendiri* alone, as is the case with focussed forms discussed above. (22a) without *sendiri* has both reflexive and non-reflexive interpretations, putting aside the reason why it does for the moment (see section 5.3). With the addition of *sendiri*, the interpretation is restricted to a reflexive one, as in (22b). This restriction can be achieved in two ways, as *diri-nya sendiri* can be parsed in two ways: (a) $[[diri\ -nya]\ sendiri]$ and (b) $[diri\ [-nya\ sendiri]]$. Since *diri* is an identity function (cf. section 4.1), these two parses can be rewritten as (a) $[[\ -nya]\ sendiri]$ and (b) $[[\ -nya\ sendiri]]$ respectively, and end up being denotationally identical to pattern (vii) ‘DP + *sendiri*’ of focussed forms (C).

- (22) a. Ali_i kata [Siti_j mengambil gambar *diri-nya*_{i/j/k}]. (= (1))
 Ali say Siti take picture DIRI-3
 ‘Ali said Siti took a picture of him/herself/her.’ (reflexive and non-reflexive)
- b. Ali_i kata Siti_j mengambil gambar *diri-nya sendiri*_{*i/j/*k}. (= (3a))
 Ali say Siti take picture DIRI-3 own
 ‘Ali said Siti took a picture of *him/herself/*her.’ (reflexive only)

It is expected that local binding is also required for (vi) *diri sendiri*, because it is in fact ‘*diri pro sendiri*’ in my analysis and hence differs from (ii) ‘*diri* + pronoun + *sendiri*’ only in that the former involves a null pronoun. This prediction is borne out, as in (23).

- (23) Ali_i kata Siti_j mengambil gambar *diri pro*_{*i/j/*k} *sendiri*.
 Ali say Siti take picture DIRI own
 ‘Ali said Siti took a picture of *him/herself/*her.’

5.3. *Diri* + DP (B. Emphatic Forms)

There are two emphatic form patterns to consider: (i) ‘*diri* + pronoun’ and (v) *diri*. The first pattern has been the main concern in previous studies. The sec-

ond pattern is in fact *diri pro*, and differs from the first pattern only with regard to the type of pronoun used, overt or null. I discuss these two patterns in separate subsections below, because the choice of the pronoun affects the resultant interpretation considerably.

5.3.1. Pattern (i): *Diri* + Pronoun

Recall that *diri* is an NP that takes a possessor argument (cf. section 4.1). Thus, in this pattern, *diri*'s possessor argument is saturated by the pronoun, resulting in an emphatic pronoun. English does not have a distinct form corresponding to this Malay form, but a similar meaning can be conveyed by prosodic emphasis.

Being an emphatic pronoun, '*diri* + pronoun' behaves pronominally. For instance, *dirinya* in (24) can take as its antecedent the matrix subject *Ali* and a discourse referent indicated by index *k*. What is surprising, however, is that local binding is also possible, as indicated by index *j*.

- (24) Ali_i kata Siti_j selalu memuji [*diri-nya*]_{i/j/k}.
 Ali say Siti always praise DIRI-3
 'Ali says Siti always praises him/her/herself.'

I adopt an account for this local binding property suggested (but rejected) by Cole and Hermon (2005). Under this analysis, the interpretation of '*diri* + pronoun' is determined by the antecedent of the possessor pronoun. This is because *diri* denotes an identity function (cf. (11)), and hence the indices of the possessor and the entire phrase become identical, as shown in (25).

- (25) $\text{dirinya: } y_1$
 $\swarrow \quad \searrow$
 diri: $\lambda x.x$ -nya: y_1

Now, a direct object possessor can be coreferential with the local subject, as in (26). Replacing *anak lelaki* in (26) by *diri*, (27) is obtained, where local binding holds between the embedded subject *Siti* and *dirinya*, making *dirinya* look like a reflexive pronoun.

- (26) Ali kata Siti_j selalu memuji [anak lelaki-nya_j].
 Ali say Siti always praise child male-3
 'Ali says Siti always praises her son.'
- (27) Ali kata Siti_j selalu memuji [*diri-nya*]_j.
 Ali say Siti always praise DIRI-3
 'Ali says Siti always praises herself.'

Of course, a direct object possessor can be coreferential with other DPs in the sentence and discourse referents, in which case a non-reflexive reading is obtained.

Cole and Hermon (2005) reject the analysis above. They expect sentences (28) to pattern with (29) rather than (30) in terms of Condition C, presumably confusing the index of the possessor *-nya* in (29) with that of the whole DP.¹⁰ What the analysis actually predicts, however, is (28) patterning with (30) rather than (29) based on the correct indexing shown in the parentheses after the Malay sentences.

- (28) a. **Diri-nya_i mencium John_i. ([diri-nya_i]_i)*
 DIRI-3 kiss John
 ‘He kissed John.’
 b. **Diri-nya_i di-cium (oleh) John_i. ([diri-nya_i]_i)*
 DIRI-3 PASS-kiss by John
 ‘He was kissed by John.’
- (29) a. *Anjing-nya_i menggigit John_i. ([anjing-nya_i]_j)*
 dog-3 bite John
 ‘His dog bit John.’
 b. *Anjing-nya_i di-pukul (oleh) John_i. ([anjing-nya_i]_j)*
 dog-3 PASS-hit by John
 ‘His dog was hit by John.’
- (30) a. **Dia_i mencium John_i. (dia_i)*
 3 kiss John
 ‘He kissed John.’
 b. **Dia_i di-cium (oleh) John_i. (dia_i)*
 3 PASS-kiss by John
 ‘He was kissed by John.’

5.3.2. Pattern (v): *Diri*

When *diri* appears to be used by itself, the meaning is generic, and not one of simple reflexivity. Recall that *diri* in such cases is in fact accompanied by the null unspecified pronoun *pro* (i.e. *diri pro*). I argue that the generic meaning is due to the unspecified nature of *pro* accompanying *diri*; *pro* means ‘one’s’ rather than ‘my/your/his/her/etc.’. Furthermore, the meaning of *diri* involved in this pattern, in most instances, is not an identity function but ‘physical self’ (10). This is because most instances of this pattern involve a reflexive meaning, which cannot be captured by the composition of an identity function and *pro*. Arguably because of the use of the ‘physical self’ *diri* instead of the identity function *diri*, (v) is much less common than the other patterns discussed above. Its distribution seems to be lexically determined, even though the semantic contribution of *diri* is clear.

¹⁰ Paul (2004) makes a similar mistake in her discussion of Malagasy *ny tenany* [DET self.3(GEN)].

In this respect, (v) resembles *self-* in English (e.g. *self-care*, *self-ignite*).¹¹

Diri (= *diri pro*) can occur as either a noun modifier or a verbal complement. In (31), *diri* modifies an NP. An NP modified by *diri* is interpreted generically as in (31a), and cannot be associated with a particular individual as in (31b).¹² *Gambar diri* could mean ‘selfie’, but like nouns with *self-* in English, the use must be established in the speaker community in order to become fully acceptable.

- (31) a. Siti tidak lalai menjaga [kecantikan/maruah *diri*].
 Siti not careless keep beauty/dignity DIRI
 ‘Siti does not fail to take care of her beauty/dignity.’ (cf. self-beauty, self-dignity)
- b. *Siti mengambil [*gambar diri*].¹³
 Siti take picture DIRI
 For: ‘Siti took a picture of herself.’ (cf. self-picture)

When *diri* is used as a verbal complement, the VP containing it describes an event that is generally self-directed rather than one that can be other-directed as well as self-directed. For example, *memukul diri* ‘to hit oneself’ in (32) describes a self-hitting action conducted for specific purposes (e.g. religious rituals), but not just any kind of self-hitting.¹⁴

- (32) Ali_i kata Siti_j memukul *diri*_{*i/j/*k}.
 Ali say Siti hit DIRI
 ‘Ali said Siti hit herself.’ (cf. self-hit)

(33) shows examples of VPs with *diri* that are commonly used.

¹¹ Chung (1976) analyses *diri* in Indonesian as a clitic attaching to a verb, as it cannot be separated from the verb by a PP, as in (i). Gil (2001) reaches a similar conclusion for *diri* in Riau Indonesian. However, under the present analysis, where *diri* is actually a phrase containing the null possessor DP *pro*, the same data needs a different explanation, e.g., the adjacency requirement on Accusative Case assignment (Ramli 1995).

- (i) a. Apakah dia membunuh *diri* [di kebun]?
 Q he kill DIRI at garden
 ‘Did he kill himself in the garden?’
- b. *Apakah dia membunuh [di kebun] *diri*?
 Q he kill at garden DIRI (Chung 1976:44)

¹² (31b) is acceptable with the irrelevant reading ‘Siti took a standing picture’.

¹³ Examples of *gambar diri* were actually found in the Internet, though they were very rare. All attested examples were in Indonesian, and *gambar diri* had a specialized meaning, namely ‘selfie’, and hence was not associated with a particular individual.

¹⁴ Not all speakers accept (32). It is acceptable only for those who know or can imagine some inherently self-directed hitting action. Nomoto (2011) thus reports a sentence with the same phrase, i.e. *memukul diri*, as ungrammatical.

- | | | |
|------|--|--|
| (33) | a. membunuh diri
kill DIRI
‘to commit suicide’ | c. melibatkan diri
involve DIRI
‘to get involved’ |
| | b. melarikan/melepaskan
make.run/release
diri
DIRI
‘to run away’ | d. menyerahkan diri
yield DIRI
‘to surrender’ |

These phrases suggest that the generic meaning ‘V one’s physical self \approx V one-self’ makes a transitive verb semantically intransitive-like.¹⁵ This quasi-intransitivization explains the inherently self-directed meaning of VPs with *diri*. It also provides an account for what appears to be local binding in (32), because an intransitivized transitive verb has reflexivity encapsulated in the verb meaning. Hence, one can treat the transitive VP *memukul diri* ‘*x* hits one’s physical self’ in (32) as if it means ‘*x* hits *x*’ or, to put it more intransitively, ‘*x* *x*-hits’ ($\lambda x.x$ hits *x*).

6. Conclusion

This paper has shown that a compositional analysis of Malay anaphoric expressions is not only possible but is empirically more desirable than a non-compositional analysis, as it can account for a wider range of anaphoric expressions in a systematic manner. The complex anaphoric expressions comprising ‘body’ plus pronoun in some related languages are known to behave similarly to ‘*diri* + pronoun’ in Malay. These languages include Malagasy (Paul 2004), Madurese (Davies 2008), Jambi (Cole et al. 2010; Kartono 2013), Javanese and Palembangese (Kartono 2013). The proposed analysis extends to these languages straightforwardly. By contrast, in non-compositional analyses, it remains unclear why the relevant binding properties are associated with ‘body’ + pronoun, but not with some other forms, and why the association is consistent across languages. In fact, Paul (2004) suggests a compositional analysis for *ny tenany* in Malagasy. This paper has developed her basic idea into a more complete and plausible hypothesis.

I would like to end this paper with an important empirical finding recently put forward by Kroeger (2014). He points out cases of ‘*diri* + pronoun + *sendiri*’ (pattern (ii)) occurring in a subject position, where it cannot be bound locally.

¹⁵ It is possible that a third *diri* morpheme exists that technically intransitivizes a transitive verb.

- (34) *Diri-nya sendiri* selalu di-utamakan-nya.
DIRI-3 own always PASS-prioritize-3
'Himself is always prioritized by him.' (i.e. 'He always gives priority to himself.')
- (Kroeger 2014:18)

Examples like this can be real counterexamples to the previous generalization that pattern (ii) does not occur in subject positions (Cole and Hermon 2005). (34) is not an apparent counterexample of the type mentioned in footnote 1, where *dirinya* (pattern (i)) happens to be immediately followed by the adverbial *sendiri* meaning 'alone, by oneself', because *sendiri* in (34) cannot be moved to a clause-final adverbial position without changing the meaning. The status of pattern (ii) is extremely important in Malay syntax, as it is one of the few diagnostics for determining the precise phrase structure of a construction (see Kroeger 2014 for a relevant discussion, and Kartini and Nomoto 2012 for an example of an argument for a particular syntactic structure based on pattern (ii)). While a non-compositional analysis like Cole and Hermon's (2005) cannot handle examples like (34), the present analysis should be able to handle them. The key lies in understanding the function of *sendiri* in the subject position, which is something that is glossed over in this study and needs to be explored in future research.

References

- Alsagoff, Lubna. 1992. Topic in Malay: The other subject. Doctoral dissertation, Stanford University.
- Chung, Sandra. 1976. An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7:41–87.
- Cole, Peter, and Gabriella Hermon. 1998. Long distance reflexives in Singapore Malay: An apparent typological anomaly. *Linguistic Typology* 2:57–77.
- Cole, Peter, and Gabriella Hermon. 2005. The typology of Malay reflexives. *Lingua* 115:627–644.
- Cole, Peter, Gabriella Hermon, and Yanti. 2010. Anaphora in traditional Jambi Malay. In *Austronesian and theoretical linguistics*, ed. Raphael Mercado, Eric Potsdam, and Lisa deMena Travis, 327–344. Amsterdam: John Benjamins.
- Davies, William D. 2008. Madurese reflexives with reference to Malay/Indonesian. *Lingua* 118:1603–1616.
- van Gelderen, Elly. 2000. *A history of English reflexive pronouns: Person, self, and interpretability*. Amsterdam: John Benjamins.
- Gil, David. 2001. Reflexive anaphor or conjunctive operator: Riau Indonesian *sendiri*. In *Long-distance reflexives*, ed. Peter Cole, Gabriella Hermon, and C.-T. James Huang, volume 33 of *Syntax and Semantics*, 83–118. San Diego: Academic Press.

- Homer, Vincent. 2009. Backward control in Samoan. In *Proceedings of the Sixteenth Meeting of the Austronesian Formal Linguistics Association (AFLA 16)*, ed. Sandy Chung, Daniel Finer, Ileana Paul, and Eric Potsdam, 45–59.
- Kartini Abd. Wahab, and Hiroki Nomoto. 2012. Passives without ‘by’ in Malay. Paper presented at the 19th Meeting of the Austronesian Formal Linguistics Association (AFLA 19).
- Kartono, Bambang. 2013. A puzzle in binding: Half reflexives and locally bound pronouns. A comparative study of anaphoric systems in Indonesian, Javanese, Palembangnese, City Jambi and Village Jambi. Master’s thesis, Utrecht University.
- Keenan, Edward L. 2002. Explaining the creation of reflexive pronouns in English. In *Studies in the history of English: A millennial perspective*, ed. Donka Minkova and Robert P. Stockwell, volume 1, 325–355. Berlin: Mouton de Gruyter.
- König, Ekkehard, and Peter Siemund. 2000. Intensifiers and reflexives. In *Reflexives: Forms and functions*, ed. Zygmunt Frajzyngier and Traci S. Curl, 41–74. Amsterdam: John Benjamins.
- Kroeger, Paul. 2014. Passive agents in Malay: The binding properties and discourse functions of agentive =*nya*. In *Current trends in Malay linguistics*, 5–29 ed. Siaw-Fong Chung and Hiroki Nomoto, volume 57 of *NUSA*. Jakarta and Tokyo: Universitas Katolik Indonesia Atma Jaya and Tokyo University of Foreign Studies.
- Nomoto, Hiroki. 2011. Teori tambatan dalam bahasa Melayu. *Toukyou Gaidai Tounan’ajiagaku* 16:1–16.
- Paul, Ileana. 2004. NP versus DP reflexives: Evidence from Malagasy. *Oceanic Linguistics* 43:32–48.
- Ramli Haji Salleh. 1995. *Sintaksis bahasa Melayu: Penerapan teori Kuasaan dan Tambatan*. Kuala Lumpur: Dewan Bahasa dan Pustaka.