

Malay tough movement, restructuring and the theory of phases

Abstract

The primary objective of this paper is to discuss *tough*-movement (TM) in Malay where the embedded object undergoes long A-movement to the matrix clause, much like German TM as described by Wurmbrand (2001). However in Malay TM, the embedded clause must exhibit passive morphology. The secondary objective of this paper is to determine the source of this passive morphology and discuss its implications for restructuring. I argue that passive in Malay TM is an example of voice matching seen in restructuring contexts (Wurmbrand 2013, Wurmbrand & Shimamura 2017). However, I depart from Wurmbrand (2013) in an important manner. It is not restructuring per se that requires voice matching. Rather voice matching is only required in a restructuring context when the matrix T probes the embedded object. I then propose that the restructuring voice head can be fully or partially copied from the matrix voice head and that it is only a fully copied voice head that allows for phase extension (Den Dikken 2007). It is argued that this is what allows the matrix T to probe the embedded object without violating the PIC. Hindi long distance agreement is then shown to support this view of restructuring and phase extension.

Keywords: A-movement, *tough* movement, restructuring, complex predicates, Malay

1. Introduction¹

Tough movement (TM)² is characterized by a configuration in which, an otherwise embedded object, shows up as a matrix subject. This is shown in the English example below.

- (1) a. It was easy to please John.³
 b. John_i was easy to please ___i.

Although *John* appears in the matrix position in (1b), it is generally accepted that *John* is interpreted as the internal argument of the embedded verb *please*. While most of the discussion of such constructions has been focused on English due to its hybrid A/A'-properties (eg. Postal 1971), TM has also been much discussed in other Germanic and Romance languages, Wurmbrand (2001) being one of the more prominent examples.

In this paper, I have two objectives. The first objective is to discuss TM in Malay (SOV, Malayic) and show that it exhibits TM in a quite novel manner.⁴

- (2) a. Adalah senang [untuk **me**-masak nasi ayam].

¹ Special thanks to all my language consultants. For Malay: Siti Umairah, Khairiyah Sabina, Nurhelmy Hakim, Puteri Sukima, Fairuz, Tiu Yau-Hwan, Sulaiman. For Hindi: Deepak Alok, Naved Aijaz, Abhishek Radhakrishnan. Thanks also to members of the SynSem reading group at NUS, especially Hiroki Nomoto and Yosuke Sato for discussion. All errors are mine.

² I refer to this phenomenon as *tough* movement (TM) only as a convenient label for the phenomenon.

³ List of abbreviations: 1, 2, 3 – person, ABS – absolutive, ACT – active, AG – agent, C – complementizer, COP – copular, DAT – DATIVE, ERG – ergative, EXPER – experiencer, F- feminine, FUT – future, INF – infinitive, INTR – intransitive, IR – irrealis, L – linker, M – masculine, mid – middle voice, NEG – negation, NPL – non-plural, OBL – oblique, PASS – passive, PST – past, PERF – perfective, PL – plural, RL – realis, TR – transitive, SG – singular,

⁴ All my informants are located in Singapore and Malaysia and are all college educated.

COP easy for ACT-cook chicken rice

‘It is easy to cook chicken rice.’

b. Nasi ayam senang [untuk **di**-masak].

Chicken rice easy for PASS-cook

‘Chicken rice is easy to cook.’

(2a) shows the expletive construction in Malay and (2b) shows TM. Note the voice morphology on the embedded verbs, which are in the active and passive forms respectively. This type of TM formation as the default strategy has not been demonstrated for any language as far as I am aware.⁵ I will argue for the following derivation for Malay TM.

(3) DP TOUGH [CP (untuk) [VP DI-Verb <DP>]]

A-movement

The diagram shows a curved arrow pointing from the DP in the embedded clause [CP (untuk) [VP DI-Verb <DP>]] to the DP in the main clause (3) DP TOUGH. Below the arrow is the label 'A-movement'.

As seen in (3), the fronted DP in Malay TM is related to the base position through long A-movement. This is very similar to Wurmbrand’s (2001) analysis of German TM which is argued to also exhibit A-movement. In addition, I show that the embedded clause lacks a TP node and as

⁵ Montalbetti & Saito (1983) show that although the embedded clause in Spanish TM is usually unmarked for voice, it can sometimes be realized as a passive, a claim repeated in Authier & Reed (2009).

a) Esta vocal es facil de ser nasalizada

This vowel is easy to-be nasalize

‘This vowel is easy to nasalize.’ (Authier & Reed 2009: 9)

However, Authier & Reed (2009: 9) claim that according to Montalbetti (p.c.), this type is only found in ‘stilted written styles’. In contrast, the regular way of forming TM in Malay is using an embedded passive verb.

such at most a vP. The analysis in (3) is motivated by looking at various intervention type effects, the distribution of the voice makers and the negation morpheme *tidak*.

My second objective is to determine the source of the passive morphology. I argue that the Malay data is another instance of the type of matching effects in restructured contexts identified by Wurmbrand (2013) and Wurmbrand & Shimamura (2017) (W&S, henceforth). Among other things, I show that this accounts for why only an internal argument can be raised to the matrix TP of a *tough* predicate.

(4) *Ali_i senang [(untuk) [vP ____i **me**-masak nasi ayam]].
 Ali easy for ACT-cook chicken rice

Lit: ‘Ali is easy to cook chicken rice.’

Given that the embedded clause in TM consists of a vP, an embedded subject should, in principle, be able to A-move to the matrix Spec, TP position. (4) shows that this is not allowed. I argue that there is no PRO in such constructions in the first place which is why (4) is not grammatical. However, this means that the passive morphology in Malay TM cannot be due to suppression of the external argument given that there is no such argument in the first place. Instead, I argue that the passive morphology is a result of voice matching with the matrix verb which is also unaccusative of the type W&S claims is required in restructuring contexts. CROSS-CONTROL CONSTRUCTIONS (CCC) in Malay (Polinsky & Potsdam 2008, Nomoto 2008) are shown to be further evidence for this.

However, one of the implications of the analysis of the Malay data, particularly of the ungrammaticality of (4) is that long A-movement is not a necessary consequence of having a restructured complement. In fact, the Malay expletive *tough* constructions have a restructured

complement, i.e. no embedded external argument PRO, but no A-movement. Neither is there voice matching. What this indicates is that the matching effects seen in restructured contexts is not correlated with restructuring per se but with long A-movement that is possible in such a context. Based on this, I argue that voice matching occurs in restructuring contexts if and only if the matrix T probes the embedded object. I argue that this falls out as an effect of the Phase Impenetrability Condition (PIC) and Phase extension (Den Dikken 2007) made possible through voice head replication. I then further support this view by turning to long-distance agreement (LDA) in Hindi-Urdu which has no A-movement but has a matrix T probing the embedded object in a restructured context (Bhatt 2005, Bhatt & Keine 2017). Crucially I show that in such contexts, there can be no voice mismatch between the matrix verb and the embedded infinitival verb, an observation not made before.

The outline of the paper is as follows. In the next section, I discuss Malay TM and outline its properties and derivation based on standard assumptions. Here, I also discuss some plausible accounts of the passive morphology in Malay TM and rule them out. In section 3, I discuss matching effects in restructuring (W&S) and show that the Malay TM facts are best thought of as another instance of such matching effects. I also discuss the CCC in Malay that further supports the presence of such matching effects. Here, I also motivate the claim that voice matching is correlated with matrix T probing the embedded object, rather than with restructuring per se. In section 4, I outline my assumptions regarding phases and SPELL OUT and show that phase extension can account for why voice matching is correlated with a matrix T probing the embedded object. In section 5, I apply this theory to Hindi LDA and show that it correctly predicts that such contexts do not tolerate voice mismatch. I, then, conclude.

2. Tough Movement in Malay

2.1 The broad typology

Although there are different accounts of English TM, it is generally accepted that there is an A'-dependency in the embedded clause of English TM (Postal 1971, Chomsky 1977, Brody 1993, Hicks 2009, Hartman 2011). We know this from the fact that TM licenses parasitic gaps and can have an intervening embedded clause between the matrix clause and the embedded gap.

- (5) a. Webber's musicals are easy to condemn __ [without even watching *pg*].
 b. *Webber's musicals are likely to be condemned __ [without anyone even watching *pg*]. (Hicks 2009: 542)

- (6) John_i was easy [to convince Sally [to please ___i]].

(5) from Hicks (2009) shows that TM, unlike raising, licenses a parasitic gap. And (6) shows that the gap can occur further embedded. These are then strong evidence for the claim that English TM does have an A'-movement component in the embedded clause and for current purposes, we can assume Chomsky's (1977) analysis of English TM which uses a null operator as shown below.

- (7) John was easy [Op_i to please t_i].

In the embedded clause, a null operator A'-moves to the clause edge, however, *John*, is base-generated in the matrix clause. Other proponents of such an analysis include Rezac (2006) and Keine & Poole (2017), although see Postal (1971), Mulder & Den Dikken (1992), Brody (1993), Hicks (2009), Hartman (2011), Longenbaugh (2016) and Selvanathan (2018) for differing analyses.

On the other hand, German TM is argued to consist primarily of A-movement, Wurmbrand (2001) being the most prominent advocate for such an approach. Notice that unlike English TM, German TM can only have one level of embedding.

(8) a. Dieser texte sind leicht zu lesen
 This text are easy to read
 ‘These texts are easy to read.’ (Wurmbrand 2001: 27)

b. **Dieses Buch ist schwer [Hans zu überzeugen [zu lesen]]*
 This book is hard John to convince to read.
 ‘This book is hard to convince John to read.’ (Wurmbrand 2001: 29)

Thus, (8a) is grammatical, whereas (8b), with the gap further embedded, is not grammatical. I will adopt Wurmbrand’s restructuring analysis for German TM in which, the embedded object undergoes long A-movement to the matrix Spec, TP position.

(9) Dieser text_i sind leicht [zu lesen _____i].



(9) shows the general analysis of (8a). Other languages which have been argued to have TM similar to German are French and Spanish (Wurmbrand 2001, Authier & Reed 2009).

With this background, I present the Malay TM sentences, reproduced from above.

(10) a. Ia/ Adalah senang [(untuk) **me**-masak nasi ayam].
 It/ COP easy for ACT-cook chicken rice
 ‘It is easy to cook chicken rice.’

b. Nasi ayam senang [(untuk) **di**-masak].
 Chicken rice easy for PASS-cook

‘Chicken rice is easy to cook.’

(10a) shows the expletive construction with a *tough* predicate and (10b) shows TM. The expletive construction can either have an overt expletive pronoun or just a bare copula form in the matrix position with an optional non-finite complementizer.⁶ Notably, the embedded clauses in these constructions must have the active and passive voice markers respectively.⁷ Thus, the following are ungrammatical sentences.

- (11) a. *Ia/ Adalah senang [untuk **di**-masak nasi ayam].
 It/ COP easy for PASS-cook chicken rice
 ‘It is easy to cook chicken rice.’
- b. *Nasi ayam senang [untuk **me**-masak].
 Chicken rice easy for ACT-cook
 ‘Chicken rice is easy to cook.’

These sentences show that passive marking on the embedded verb in the expletive and active marking on the embedded verb in TM are both disallowed. Observe the occurrences of these markers in canonical active and passive clauses below.

- (12) a. Ali **me**-masak nasi ayam
 Ali ACT-cooked chicken rice
 ‘Ali cooked chicken rice.’

⁶ *Untuk* is a complementizer that occurs before an infinitival clause. As Cole & Hermon (2005) claim, *yang* is the complementizer that occurs before finite clauses in Malay.

⁷ While the *meN* marker is usually characterized as an active marker, it is also argued to have a progressive aspectual meaning (Soh & Nomoto 2009).

- b. Nasi ayam **di**-masak (oleh Ali)
 Chicken rice PASS-cook by Ali
 ‘Chicken rice was cooked (by Ali).’

Thus, the active clause in (12a) and the passive clause in (12b) are marked with the *meN* and *di* prefixes respectively. For more on Malay voice markers, I refer the reader to Chung (1976), Saddy (1991), Sneddon (1996), Nomoto & Shoho (2007), Cole & Hermon (1998), Cole, Hermon & Yanti (2008), Aldridge (2008), Sato (2011, 2012), Soh & Nomoto (2009, 2010, 2011), Son & Cole (2004), Wouk (1989), Guilfoyle, Hung & Travis (1992), Voskuil (1993), Willet (1993) among many others. Just from the surface forms in (10), we can see that Malay provides the opportunity to make novel generalizations about TM because it has overt voice morphology that English and German do not. Perhaps the most interesting aspect is that the embedded clause in Malay TM must be in passive voice, an option robustly not allowed in English (or German).

- (13) a. The rice was easy to cook.
 b. *The rice was easy to be cooked.

Thus, the obvious question that has to be answered is why must the embedded clause in Malay TM be in the passive form? I turn to this question and other questions about the Malay TM derivation in the next section.

2.2 The properties of Malay TM

The passive morphology in Malay TM is highly suggestive that Malay TM may involve only A-movement, unlike English which has A'-movement. This is supported by the fact that in Malay, an intervening embedded clause can only occur in the expletive construction but not in TM.

- (14) a. Ia/ Adalah senang [untuk me-yakinkan Ali [untuk me-masak
It/ COP easy for ACT-convince Ali for ACT-cook
nasi ayam]]
chicken rice
'It was easy to convince Ali to cook the chicken rice.'
- b. *Nasi ayam_i senang [untuk me-yakinkan Ali [untuk di-masak ____i]]
chicken rice easy for ACT-convince Ali for PASS-cook
For: 'Chicken rice was easy to convince Ali to cook.'

(14a) shows the expletive construction and the string *meyakinkan Ali* 'to convince Ali' is possible. However, (14b) shows that this string cannot intervene between the matrix clause and the lowest clause. Thus, Malay TM behaves on par with German TM in (8) in not having A'-movement in the embedded clause. Rather, there is only A-movement.

In addition, evidence from the distribution of the Malay negation markers, *tidak* and *bukan*, discussed by Kroeger (2014) indicates that the A-movement in Malay TM is necessarily long A-movement. The markers are shown below.

- (15) a. Saya tidak lapar
1SG NEG hungry
'I am not hungry.' (Sudaryono 1993: 88)
- b. Dia bukan tidur, tetapi ber-baring sahaja
3SG NEG sleep but MID-lie.down only
'He is not sleeping, but only lying down.' (Asmah 1982: 145)

Based on a number of semantic and syntactic pieces of evidence, Kroeger (2014) argues that *bukan* attaches high to IP whereas *tidak* attaches lower, adjoined to Spec, vP.⁸ One piece of evidence that *bukan* is higher in the structure than *tidak* comes from the fact that *bukan* must always precede a modal whereas *tidak* can precede or follow a modal.

- (16) a. Saya *tidak* harus makan.
 1SG NEG must eat
 ‘I am not required/obligated to eat.’ (Sudaryono 1993:76)
- b. Saya harus *tidak* makan.
 1SG must NEG eat
 ‘I must not eat.’ (Sudaryono 1993:77)
- (17) a. Saya *bukan* harus makan, tetapi ...
 1SG NEG must eat but
 ‘I am not required to eat, but...’ (Kroeger 2014: 154)
- b. *Saya harus *bukan* makan ...
 1SG must NEG eat (Kroeger 2014: 154)

Thus, in (16), *tidak* can follow the modal, whereas (17) shows that *bukan* must precede the modal. If we assume that the modal is in I (as Kroeger 2014 does), we can see that *tidak* can be in a position lower than I whereas *bukan* cannot. This suggests that *tidak* can be adjoined to vP whereas *bukan* cannot. In addition, when both *bukan* and *tidak* co-occur, *bukan* must come first.

⁸ There are other differences between *bukan* and *tidak*. *Tidak* usually negates adjectival and verbal predicates whereas *bukan* usually occurs with nominal predicates. However, when *bukan* does negate a verb, an additional contrastive meaning is reported to be present (Kroeger 2014: 138).

- (18) Sebenarnya dia *bukan tidak* mampu, melainkan tidak mau.
 Truly 3SG NEG NEG able on.contrary NEG want

‘Actually it is not the case that he cannot (do it), rather he does not want to.’

(Sudaryono 1993:203)

The reverse order between *bukan* and *tidak* is reported to be ungrammatical (Kroeger 2014: 155).

This is further evidence that *bukan* occurs higher structurally than *tidak*. I refer the reader to Kroeger (2014) and the references therein for more evidence for where *bukan* and *tidak* attach in the structure. With this background, it is notable that *tidak* is possible in a clause embedded by *untuk*, while *bukan* is not (Kroeger 2014).

- (19) Terkadang aku harus diam [untuk *tidak/*bukan* memperbesar
 occasionally 1SG must silent for NEG enlarge
 masalah].
 problem

‘Occasionally I need to keep silent in order not to make the problem worse.’

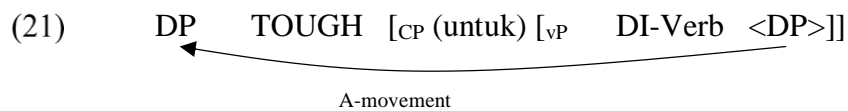
(Kroeger 2014: 167)

If Kroeger (2014) is right in where these negation markers attach structurally, the difference in the availability of the two negation markers in (19) suggests that *untuk* embedded clauses do not have a TP node after all. We see the same pattern in *tough* constructions.

- (20) Nasi ayam senang [untuk *tidak/ *bukan* **di**-masak].
 Chicken rice easy for NEG PASS-cook

‘Chicken rice is easy to not cook.’

Thus, in (20), the embedded clause can have *tidak* but not *bukan*. Given our previous discussion, this indicates that there is no TP node in the embedded clauses. This means that the embedded object must have long A-moved from its base position to the matrix clause without moving to some intermediate Spec, TP.⁹ Thus, the following is the proposed analysis of Malay TM.



The embedded object A-moves from the embedded clause to the matrix Spec, TP. The embedded clause has to at least be a vP (due to the presence of passive morphology and the ability to host *tidak*) but cannot be as big as a TP (due to the inability to allow *bukan*). I will assume that it is a vP. An optional non-finite complementizer *untuk* projects a CP and takes the vP as its complement. As this is a non-finite complementizer, I will assume that this CP is not a phase.

2.3 The source of the passive morphology

Now that we have a derivation of Malay TM, we can now set about answering the question as to why the embedded clause must be in passive form. I will consider two plausible explanations, the first that this is due to a general extraction restriction seen in Austronesian languages and the second that this is due to suppression of an external argument. I will rule both of them out.

⁹ If we assume that passive voice heads are phasal (Legate 2003), then under standard assumptions, the object has to stop at the edge of the vP in order to obey the Phase Impenetrability Condition (PIC, Chomsky 2000). However, I will argue that this type of adjunction is not a possibility with A-movement in general in a later section.

2.3.1 Austronesian DP extraction restrictions

Austronesian languages such as Tagalog are famous for imposing extraction restrictions from a clause depending on voice morphology (eg. Rackoweski 2002, Rackowski & Richards 2005, Aldrige 2008). The following Tagalog examples are from Aldrige (2008).

- (22) a. **B-in-ili** ng babae ang isda.
 -TR.PERF-buy ERG woman ABS fish
 “The woman bought the fish.”
- b. **B-um-ili** ang babae ng isda.
 -IN.PERF-buy ABS woman OBL fish
 “The woman bought a fish.” (Aldrige 2008: 1441)

(22) shows that the verbal morphology (assumed to be a voice head) differs depending on the absolutive marked argument. In (22a), *-in-* is correlated with absolutive marking on the direct object and in (22b), *-um-* is correlated with absolutive marking on the external argument.

Notably, only the absolutive marked argument can be extracted.

- (23) a. Ano ang **b-in/ *um-ili** ng babae?
 what ABS TR.PERF/ *INTR.PERF-buy ERG woman
 “What did Maria buy?” (Aldrige 2008: 1442)
- b. Sino ang **b-um/ *in-ili** ng isda?
 who ABS -IN.PERF/ *TR.PERF-buy OBL fish
 “Who bought the fish?” (Aldrige 2008: 1442)

(23) shows that in wh-movement, the *-in-* prefix allows only the extraction of the direct object and that the *um* prefix allows only the extraction of the external argument. As such one may

wonder if the passive morphology in the embedded clause in TM is a reflection of a similar restriction. But it cannot be. The reason is that, even though Malay does exhibit a similar extraction asymmetry associated with different voice morphology, a *di-* passive is not necessary for extraction of a direct object.

- (24) a. Ali ***(mem)**-beri buku pada Fatimah
 Ali ACT-give book to Fatimah
 ‘Ali gave a book to Fatimah.’ (Cole et al 1998: 231)
- b. Siapa_i ____i ***(mem)**-beri buku pada Fatimah
 Who ACT-give book to Fatimah
 ‘Who gave a book to Fatimah?’ (Cole et al 1998: 231)
- c. Apa_i Ali (***mem**)-beri ____i pada Fatimah
 What Ali ACT-gave to Fatimah
 ‘What did Ali give to Fatimah?’ (Cole et al 1998: 231)

(24a) shows a transitive clause which must have the *meN* prefix. (24b) shows wh-movement of the external argument which is compatible with the *meN* prefix. (24c) shows that wh-movement of the direct object is incompatible with the *meN* prefix. Instead, there cannot be any overt prefix.¹⁰ Thus, extraction of direct objects is indeed incompatible with the *meN* prefix but this does not mean that the passive morphology in TM is due to a similar extraction restriction.

Consider the following.

- (25) *Nasi ayam senang [untuk masak].

¹⁰ There is rich literature discussing this pattern in Malay. I refer the reader to Cole, Hermon & Yanti (2008), Aldridge (2008), Sato (2012) among others for various theories on when and why A'-movement obviates *meN*.

Chicken rice easy for cook

‘Chicken rice is easy to cook.’

(25) shows that the embedded verb in TM cannot be bare either. This shows that the passive morphology in the TM embedded clause is not due to general voice restrictions on direct object extraction. If this was the case, we should expect to see that bare embedded verbs are possible.

2.3.2 Relativized Minimality

An alternative reason why the embedded clause in TM has to be passive is due to considerations of Relativized Minimality (RM, Rizzi 1990). Given that the embedded clause in Malay TM is a vP, when the embedded clause has active morphology, one might expect that there is an intervening PRO in Spec, vP as shown below.

(26) *DP_i TOUGH [_{vP} PRO ACT-V <DP>]

Violation of RM

In such a configuration where the embedded verb is in active form, A-movement of the embedded object leads to a RM violation. In contrast, passivizing the embedded clause suppresses this external argument PRO as shown below.

(27) DP_i TOUGH [_{vP} PASS-V <DP>]

No violation of RM

Thus in (27), A-movement of the embedded object can proceed without inducing a RM violation given that there is no intervening PRO in the structure and the obligatory passive morphology in Malay TM has a simple explanation. As for the motivation behind the movement of the embedded object to matrix Spec, TP, we could make the standard assumption that a passive voice head does not assign case to its direct object but the matrix T does.

However, there is reason to doubt this account as well. For one, if there is a PRO in the embedded clause in Malay TM in the expletive construction, we expect the following to be possible, but it is not.

(28) *Ali_i senang [(untuk) [vP ____i me-masak nasi ayam]].
 Ali easy for ACT-cook chicken rice

Lit: 'Ali is easy to cook chicken rice.'

In (28), the embedded clause remains in active form and the external argument itself, is A-moved to the matrix Spec, TP position. Assuming that PRO can be realized as an overt argument if it has a case licenser, movement of the external argument to the matrix Spec, TP should case license an overt DP in this position. But this is not possible. It is not clear why. However, if there is no PRO in the embedded clause after all, then there is no external argument to target for movement to the matrix position and (28) is expected to be ungrammatical. Thus, (28) supports the view that there is no PRO in the expletive construction after all.^{11, 12}

Evidence from reflexives suggests that there may be a PRO in the embedded clause but upon closer inspection, the evidence does not support such a view either. First, note that the

¹¹ This indicates that the expletive construction has a restructured complement, i.e. alcs an embedded PRO, although another tell-tale sign of restructuring, i.e. long A-movement, has not taken place. In a later section, I will use this as the main piece of evidence for arguing for a distinction between partial and full restructuring contexts.

¹² One may also suppose that perhaps there is PRO but that it is syntactically inert from some reason (Keine to appear). This may explain why the embedded subject cannot be moved to the matrix clause. However, this should also mean that there is no need to suppress this argument through passivization either.

Malay reflexive form requires a syntactic A-binder (Arka & Manning 1998 (A&M), Cole et al., Sato 2012).

- (29)
- a. Saya men-yerahkan diri saya ke polisi.
 1SG ACT-surrender SELF 1 to police
 ‘I surrendered myself to the police.’ (A&M: 3)
- b. *Diri saya men-yerahkan saya ke polisi
 Self 1 ACT-surrender 1SG to police
 ‘Myself surrendered I to the police.’ (A&M: 3)
- c. Diri saya saya serahkan ke polisi
 Self 1 1SG surrender to police
 ‘I surrendered myself to the police.’ (A&M: 7)

(29a) shows a direct object reflexive form that is bound by the surface subject. (29b) shows that the reverse order between the reflexive and the binder is not possible. This sentence cannot be ruled out as a general restriction on reflexives occurring as sentence subjects (like in English) because SELF-forms in Malay can occur as sentence subjects as seen in (29c). (29c) shows what is known as the bare passive in Malay (Dardjowidjojo 1978, Alsagoff 1992, Sneddon 1996, Cole & Hermon 2005, Aldridge 2008, Sato 2012). In such constructions, the general consensus is that this is not a topicalization structure, but one where the object reflexive is A-moved to Spec, TP with the external argument remaining low, perhaps cliticized to the verb. In so far as this is correct, (29c) shows that the *diri*-SELF form can occur as a sentence subject and that the problem with (29b) is that there is no syntactic binder for the reflexive at any point in the derivation (Belletti & Rizzi 1988).

With this background, we can now look at the following expletive sentence.

- (30) Adalah senang [untuk me-mukul diri kita]
 COP easy for ACT-beat self 1.PL
 ‘It was easy to beat ourselves.’

The grammaticality of (30) suggests that there is a PRO in the embedded clause, an apparent contradiction to the conclusion above. However, the data in (30) is also compatible with the view that there is no PRO in the embedded clause. First, note that the antecedent of the reflexive is by default interpreted as the implicit experiencer of the matrix predicate. This can be seen also in the following examples.

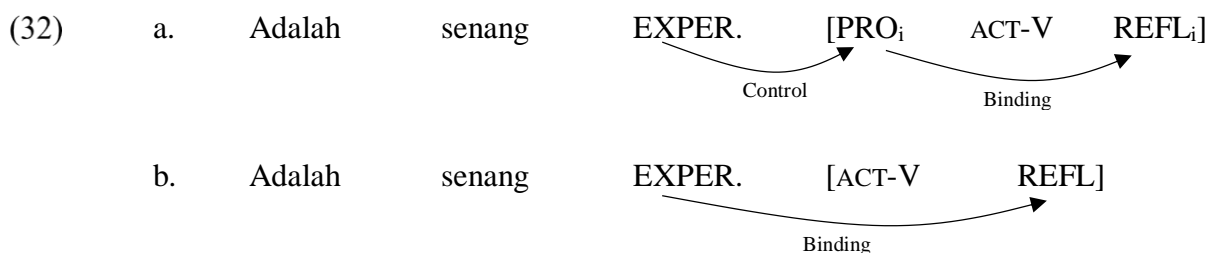
- (31) a. Adalah senang [untuk Ali] me-mukul diri nya
 COP easy for Ali ACT-beat self 3
 ‘It was easy for Ali to beat himself.’
 a. *Adalah senang [untuk Ali] me-mukul diri kita
 COP easy for Ali ACT-beat self 1.PL
 For: ‘It was easy for Ali to beat ourselves.’

In (31a), the matrix PP introduces an experiencer *Ali* which is then interpreted as the obligatory antecedent of the reflexive.¹³ (31b) shows that the antecedent of this reflexive cannot be

¹³ Given that *untuk* is like English *for* in being ambiguous between a complementizer and a preposition, one might wonder if *Ali* in (31) is actually an embedded subject. However, this cannot be. Because unlike English *for*, *untuk* cannot license a DP when it is a complementizer.

- a. It is easy [_{PP} for the rich] [_{CP} for the poor to do the hard work]. (Chomsky 1973)
 b. Adalah senang [_{PP} untuk Ali] [_{CP} untuk (*Sam) me-masak nasi ayam.

independent of the matrix experiencer. These facts are compatible with the idea that the matrix experiencer obligatorily controls the embedded PRO which then binds the reflexive. But these facts are also compatible with the idea that a null matrix experiencer serves as the antecedent of the reflexive without the need for any intervening PRO. These two options are shown below.



In (32a), the implicit argument controls the PRO (Landau 1999) and PRO binds the reflexive. In (32b), there is no PRO but rather the experiencer directly binds the reflexive. I assume that this experiencer is present even when it is not overtly realized by a PP. Now consider the following.

- (33) Diri kita senang [untuk di-pukul]
 REFL 1.PL easy for PASS-beat
 ‘We were easy to beat.’ Lit: Ourselves were easy to beaten.

(33) shows TM with the embedded object reflexive realized as the matrix subject. And this sentence is possible with the implicit first person experiencer binding the reflexive. If we assume

cop easy for Ali for Sam act-cook chicken rice


For: ‘It is easy for Ali for Sam to cook the chicken rice.’

(a) shows that in English both the prepositional *for* and the complementizer *for* can co-occur. We can assume that the second occurrence of *for* is the complementizer and it is this head that licenses the embedded subject. Malay can also have two occurrences of *untuk* as seen in (b). However, the second occurrence of *untuk* cannot have an overt DP after it. This indicates that complementizer *untuk* does not license an overt DP. *Ali* in (31) must be part of a PP.

that a reflexive is licensed if it is bound at some point in the derivation following Belletti & Rizzi (1988) and that either the experiencer or the PRO is the reflexive binder, the grammaticality of (33) is expected under (32b) and not (32a). This is because when the embedded clause is passive, there is no PRO in the structure that could have bound the reflexive before it A-moves to the matrix Spec, TP. On the other hand, regardless of whether the embedded clause is active or passive, the structure still has an experiencer argument in the matrix clause that can bind the reflexive in its base position. Thus, I propose that the reflexive data, like the other data we have seen in this section, is compatible with the view that there is no PRO in the Spec, vP of the embedded clause of the expletive *tough* construction shown below after all.

- (34) Adalah senang [untuk **me**-masak nasi ayam].
 COP easy for ACT-cook chicken rice
 ‘It is easy to cook chicken rice.’

The main implication from this is that passivization should not be required in the embedded clause in order to allow TM. This is because in the absence of PRO, the embedded object can be raised to the matrix Spec, TP position without violating RM as shown below.

- (35) DP_i TOUGH [_{vP} ACT-V <DP>]

 No violation of RM

I conclude that the reason why the embedded clause in Malay TM must be in passive form is not due to RM either.¹⁴

¹⁴ Also note that a RM analysis will not explain why there is obligatory co-construal between the implicit matrix experiencer argument and the implicit embedded agent. This key fact has to be accounted for in any satisfactory account behind why the embedded clause in Malay TM must be in passive voice.

Thus far, we have discussed Malay TM and seen that the embedded clause must be in passive form. I have argued that the embedded object in Malay TM undergoes A-movement to the matrix Spec, TP from its base position and that the embedded clause is a vP.¹⁵ The passive morphology in TM however, does not have an explanation yet. It is not due to general extraction restrictions like those in Tagalog and Malay. Neither is it due to RM. We need to look elsewhere.

3. Matching effects in restructuring

3.1 Passive in Malay TM as voice matching

In so far as the arguments above are right, the obligatory passive morphology in the embedded clause of Malay TM is still unaccounted for. In this section, I propose that the source of the passive morphology in Malay TM is the same as the matching effects that we see in restructuring contexts in other languages. The generalization, adopted from W&S, is as follows.

- (36)
- | | | | | |
|----|------------------|--------------------------|---|--------------|
| a. | *DP _i | Pred _{restruc.} | [_{vP} ACT-V <DP _i >] | *No matching |
| | | Non-active | Active | |
| b. | DP _i | Pred _{restruc.} | [_{vP} DI-V <DP _i >] | Matching |
| | | Non-active | Non-active | |

(36a) shows two possible configurations with a matrix restructuring predicate which is unaccusative, like a *tough* predicate. In such a configuration, the embedded clause predicate must also be non-active, i.e. passive in the case of Malay TM. Thus, only (36b) is allowed.¹⁶

¹⁵ As mentioned above, I assume that the CP layer projected by optional *untuk* is syntactically inert.

¹⁶ Given that Malay has bare passives as shown earlier, one may wonder why the embedded clause in Malay TM cannot be bare. I propose that this is because even in the bare passive, an external argument is present, albeit cliticized, thus it does not count as truly non-active for the purpose of matching.

Wurmbrand (2001) originally proposed that the embedded clause in German TM is a VP and not a vP. She makes this claim because she needed to explain why a verb that otherwise looks like it is in the active form (and thus capable of assigning accusative case) still triggered A-movement from the embedded object position. If the embedded clause is only a VP without a v head, then the embedded object has to move to the matrix Spec, TP to get case. However, Chung (2004), and W&S show that not all cases of lexical restructuring have an embedded clause which can be analyzed as having only a VP. For example, Chamorro (V initial) has restructuring where the embedded clause has passive morphology similar to what we have seen in Malay TM.

- (37) **Ma**-hassu **ma**-na'na'lu i lepblu ni
 NPL.RL.IN.PASS-think NPL.RL.IN.PASS-return the book OBL
 ma'estra siha.
 teacher PL
 'The teachers remembered to return the books.' (Chung 2004: 219)
 (Lit. 'The books were remembered to be returned by the teachers.')

Chung (2004) and W&S analyze (37) as a context of restructuring where the embedded object triggers agreement (in the intransitive paradigm) on the matrix verb.¹⁷ Notably, the embedded clause must have passive morphology, as the following shows. Thus, while (37) is grammatical, the following is not.

- (38) ***Ma**-hassu **mu**-na'na'lu i lepblu ni
 NPL.RL.IN.PASS-think INF.TR-return the book OBL

¹⁷ A complex agreement paradigm and flexible word order obfuscates the pattern of agreement seen in Chamorro restructuring. I refer the reader to Chung (2004) for more details.

ma'estra siha.
 teacher PL (Chung 2004: 220)

In contrast to (37), the embedded clause in the long A-movement context above cannot be in active form. Voice mismatch rules this out. Based on such languages, W&S propose that restructuring is a context in which there is obligatory voice matching. In other words, the matrix predicate and embedded predicate in a restructuring context both have voice heads and crucially they must match. Thus, when the matrix predicate in restructuring is non-active (passive/unaccusative), the embedded clause must also be non-active (passive/unaccusative).¹⁸ See W&S for more cross-linguistic evidence for voice matching in restructuring contexts.

Malay also has another construction which supports the view that long A-movement is correlated with voice-matching. This has been called the CROSS CONTROL CONSTRUCTION (CCC, Polinsky & Potsdam 2008) and is shown below.

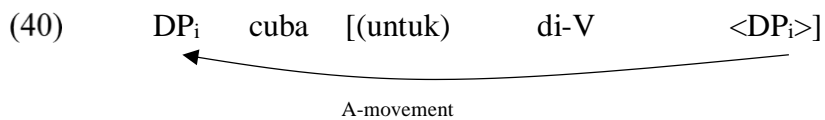
- (39) Mat Rempit cuba [(untuk) di-tangkap (oleh polis)]
 Motor cycle gang try for PASS-catch by police
- i. ‘The motor cycle gang tried to be caught by the police.’
 ii. ‘The police tried to catch the motor cycle gang.’ (Nomoto 2008: 1)

(39) shows a passive embedded clause under *cuba* ‘try’.¹⁹ This sentence is reported to have two interpretations. The first is the interpretation where the motor cycle gang was acting in a self-destructive manner and tried to be caught by the police. I will assume with Nomoto (2008) that

¹⁸ In this perspective, W&S claims that the embedded clause in German TM appears active but is actually a default spell-out of voice morphology.

¹⁹ See Nomoto (2008) for an extensive list of predicates that exhibit CCC.

this is an instance of subject control. The relevant interpretation for us is the second, CCC, one in which the surface subject *Mat Rempit* ‘motor cycle gang’ is interpreted as the embedded object. Although Polinsky & Potsdam (2008) and Nomoto (2008) propose different analyses for this construction, they are unified in assuming long A-movement of the embedded object from its base position to the matrix Spec, TP. I will thus assume this is indeed the case in the CCC reading. Thus (39), on the CCC reading, has the following derivation.



This is essentially the same derivation that has been proposed for Malay TM which also has an embedded passive clause. I will thus assume that CCC readings are also a case of restructuring where the embedded clause is only a vP and has no PRO. If this is right, we expect to see matching effects in (39) in the form of passive morphology *di-* in the matrix clause but we do not. But rather than being a contradiction of matching effects in restructuring, I propose that in such constructions, the matrix verb *is* in the non-active form and thus indeed matching with the embedded passive VOICE head which is also non-active. We can see it in the following.

- (41) Mat Rempit men-cuba [(untuk) di-tangkap (oleh polis)]
 Motor cycle gang ACT-try for PASS-catch by police
- i. ‘The motor cycle gang tried to be caught by the police.’
- iii. *‘The police tried to catch the motor cycle gang.’ (Nomoto 2008: 1)

Unlike adjectival *tough* predicates, *cuba* ‘try’ is a verb that can host the active *meN* prefix. (41) shows this. In this case, there is definitely mismatch between the matrix and embedded clause VOICE heads. And as one might predict in the voice matching analysis, in this case, the CCC

reading is no longer possible. (41) only has the control reading. This indicates that in (39), the matrix verb is not in the active form and that long A-movement (a signal of restructuring) is not possible with mismatched VOICE heads as expected.²⁰

We can thus conclude that in contexts of long A-dependencies in Malay (TM and CCC), the voice features of the matrix and embedded voice heads have to be the same.²¹ We thus have a reason for why the embedded clause in Malay TM must be a passive. Since the matrix predicate in TM is unaccusative, i.e. non-active, the embedded clause must also be the same. Thus, Malay TM exhibits a pattern that is part of a larger generalization supported within Malay and cross-linguistically. The following shows voice matching in restructuring contexts in schematic form.

(42) a. *... V_[α]-Pred_{matrix}. [_{VP} V_[β]-Pred_{embed}. ...] *No matching

²⁰ One reason why the matrix verb is bare rather than with the *di* prefix may have to do with the fact that *cuba* and the *di* prefix are generally incompatible. We can see it in the following examples.

- a) Ali men-cuba nasi ayam
 Ali ACT-try chicken rice
 ‘Ali tried chicken rice.’
- b) *Nasi ayam di-cuba (oleh Ali)
 chicken rice PASS-try by Ali.

For: The chicken rice was tried by Ali.

Cuba ‘try’ does not usually take a direct object but it does in the sentence in (a). However, even in such a context, *cuba* is not compatible with the type of argument demotion/ promotion that takes place with the *di* passive. This can then explain why *di* is not possible in (39). What this also indicates is that what matters is that the voice head be active or non-active for matching purposes. Further distinctions between non-actives are not relevant.

²¹ How this feature duplication is explained will be taken up in the next section.

b. ... $V_{[\alpha]\text{-Pred}_{\text{matrix}}}$ [$v\text{P}$ $V_{[\alpha]\text{-Pred}_{\text{embed}}}$...] Matching

(42) shows restructuring contexts with a matrix and embedded predicate where the voice value (active/ passive) of the matrix and embedded voice heads must be the same.²²

3.2 Explaining feature matching in restructuring

In this section, I will review W&S's account of voice matching in restructuring contexts and show that it can explain several aspects of the Malay TM construction.

3.2.1 Wurmbrand & Shimamura (2017): Voice restructuring

W&S's approach to restructuring, which is couched in a valuation model of AGREE (Pesetsky & Torrego 2007), is to posit a specific restructuring voice head that occurs in the embedded clause in a restructuring context. One of the key aspects of this voice head is that it comes unspecified for voice and phi features pertaining to the external argument. In the course of the derivation, these are valued by the matrix voice head which does have values for these features. This is shown below.

(43) [$v\text{P}$ $V_{[\text{voice: PASS}][\text{i}\phi: \text{AG}]}$ [V_{matrix} [$v\text{P}$ $V^R_{[\text{voice: } _][\text{i}\phi: _]}[V_{\text{embed}} \dots]]$]]

(43) shows a restructuring context of long A-movement.²³ Recall that in this context, the matrix and embedded clauses must be non-active. Here, the matrix voice head comes valued for voice features (voice: PASSIVE) and phi features. The phi-features encode the features of the implicit agent (iφ: AG) (Legate 2012, W&S). The embedded clause contains the special restructuring

²² W&S describes several languages where a clear voice mismatch in restructuring exists, eg. German TM. As mentioned earlier, I will assume with them that this is a default voice valuation and not real active.

²³ This is a simplified representation of W&S. They make the distinction between VoiceP and vP as part of the split voice domain. I will abstract away from this distinction as this is not relevant for our purpose. See W&S for details.

voice head (v^R) which comes unvalued for both voice and phi features. In W&S's account, the lower voice head is valued by the higher voice head to yield the following.

$$(44) \quad [_{VP} \quad \underline{V}_{[voice: PASS][i\phi: AG]} \quad [_{V_{matrix}} \quad [_{VP} \quad \underline{V}^R_{[voice: \underline{PASS}][i\phi: \underline{AG}]} \quad [_{V_{embed} \dots}]]]]$$

Underlining shows the values that are copied by the embedded voice head v^R from the matrix v . Thus, the lower voice head receives the same values as the matrix voice head for voice and phi features. This account succeeds in capturing three facts about long A-movement. First, the embedded voice head comes unvalued for phi features and as such it is unable to project a PRO, thus explaining why there is no intervention effect by a putative PRO in long A-movement contexts. Second, voice matching is accounted for, given that the lower voice head copies its voice value from the higher voice head. Third, this resolves the question as to how the agent interpretation of the embedded verb is derived. Since the embedded voice phi features are copied from the matrix voice head, this ensures the obligatory control interpretation.

This analysis can, as shown by W&S, account for the voice matching in Chamorro. The following are reproduced from above.

- (45) **Ma**-hassu **ma**-na'na'lu i lepblu ni
 NPL.RL.IN.PASS-think NPL.RL.IN.PASS-return the book OBL
 ma'estra siha.
 teacher PL
 'The teachers remembered to return the books.' (Chung 2004: 219)
 (Lit. 'The books were remembered to be returned by the teachers.')
- (46) ***Ma**-hassu **mu**-na'na'lu i lepblu ni
 NPL.RL.IN.PASS-think INF.TR-return the book OBL

ma'estra siha.

teacher PL (Chung 2004: 220)

Thus as seen above, when the matrix clause is in passive form, the embedded clause must also be in passive form. W&S's theory can also explain the following more complex data regarding the type of control interpretations that are possible in Chamorro.

- (47) Si Rita ha-ayuda i lalahi [gumoddi i chiba
 Rita 3SG.RL.TR-help the boys INF.TR.tie the goat
 ni esti na tali].
 OBL this L rope

'Rita helped the boys tie up the goat with this rope.' (Chung 2004: 213)

(47) shows an infinitive clause which is not a restructuring context in Chamorro, as such there can be an embedded, external argument PRO. In this clause, observe that the matrix non-agent can control the embedded PRO. However, when the embedded clause is a restructuring complement, this type of non-subject control is no longer possible.

- (48) *Si Rita ha-ayuda i lalahi [**ma**-gumoddi i chiba
 Rita 3SG.RL.TR-help the boys 3PL.RL.TR-tie the goat
 ni esti na tali].
 OBL this L rope

'Rita helped the boys tie up the goat with this rope.' (Chung 2004: 213)

(48) shows that the embedded clause with the realis marker identified by Chung (2004) as identifying a restructuring context. Notably, non-subject control in this context is not allowed. W&S's theory explains these facts without further stipulation. Recall that v^R does not allow an

external argument as it copies its phi-features from the matrix voice head. Thus, PRO only exists in the embedded clause in (47) but not (48). This means that the only way to get a control interpretation in the restructuring context is by copying the phi features of the matrix voice head. Since only the agent phi features are present on the voice head, only this can be copied to the v^R .²⁴ On the other hand, the non-restructuring infinitive allows an embedded PRO as it does not have a v^R and as such allows non-subject control in the usual way (Landau 2013).

W&S's account can also explain the voice matching effects in Malay TM and the CCC construction quite straightforwardly. Recall the relevant constructions.

- (49) a. Nasi ayam senang [(untuk) **di**-masak].
 Chicken rice easy for PASS-cook
 'Chicken rice is easy to cook.'
- b. Mat Rempit cuba [(untuk) **di**-tangkap (oleh polis)]
 Motor cycle gang try for PASS-catch by police
 'The police tried to catch the motor cycle gang.' (Nomoto 2008: 1)

(49a) shows Malay TM and (49b) shows the CCC construction. These are both contexts of long A-movement which exhibit voice matching as discussed above. In these constructions, the embedded voice head, v^R , comes unvalued for phi and voice features. The values for these features are copied from the matrix voice head. This results in the same voice value as the matrix clause, i.e non-active (which is spelled out as the passive). Copying of the phi features results in

²⁴ Given that the matrix clause is active, the matrix voice head cannot come valued with the phi features of the agent. I assume, following W&S, that in this case, the phi features of the external argument are copied to the matrix voice head which is then copied by the embedded v^R .

an interpretation where the implicit agent of the embedded verb is co-construed with the implicit argument in the matrix clause. In the case of TM, it is the implicit matrix experiencer and in the case of CCC, it is the implicit matrix agent.

- (50) Nasi ayam $V_{[\text{voice: PASS}][\text{i}\phi: \text{EXP}]}$ -senang [(untuk) $v^R_{[\text{voice: PASS}][\text{i}\phi: \text{EXP}]}$ -masak].
 Chicken rice easy for PASS-cook
 ‘Chicken rice is easy to cook.’

(50) shows the derivation for Malay TM but the analysis for CCC will be similar. As before, underlining shows the values that are copied by the embedded voice head v^R . Here, the voice and phi values are copied from the matrix voice head.²⁵ Thus, the embedded clause has passive morphology and the embedded agent is necessarily co-construed with the matrix experiencer.²⁶

To summarize the discussion so far, we have seen that W&S’s v^R account of restructuring can capture a number of salient properties of long A-movement. In this analysis, the restructuring voice head, v^R , and how its features are valued can explain why there is voice matching in Malay TM and how the right interpretation for these sentences comes about. In the next section, I will discuss a non-trivial problem for W&S’s account. This lies in their assumption that a restructured context and voice matching have a one-to-one correspondence. I show that there cannot be such

²⁵ This raises the question as to why German TM does not exhibit passive morphology in the embedded clause. For this, I adopt W&S’s analysis that German spells out default active morphology in the infinitive clause.

²⁶ One of the difficult aspects of the interpretation of the CCC construction has been the obligatory co-construal of the implicit matrix agent with the implicit embedded agent. Polinsky & Potsdam (2008) proposes a lexical semantic and Nomoto (2008) proposes downward theta role assignment by the matrix voice head to account for this co-construal. In the proposed account, this co-construal falls out from phi-feature copying.

a correspondence in Malay. Instead, I argue that voice matching arises only when the matrix T probes into the embedded vP.

3.2.2 A problem with W&S: Restructuring without matching effects

Note that in W&S's account, the restructuring voice head is a special head that restricts the generation of an embedded external argument PRO. This head can also explain voice matching effects. However, this fails to explain why there cannot be an embedded PRO even in the Malay expletive construction shown below.

- (51) Adalah senang [untuk **me**-masak nasi ayam].
 COP easy for ACT-cook chicken rice
 'It is easy to cook chicken rice.'

(51) shows the expletive construction where TM has not taken place and there clearly is no voice matching. This active voice cannot be a default voice valuation as it is able to assign case to the direct object. In W&S's account, this translates to an absence of the restructuring voice head, v^R . However, recall that I have argued that even in these expletive constructions, there cannot be an embedded PRO. This is because it is the lack of an embedded PRO that accounts for the ungrammaticality of the following sentence.²⁷

²⁷ Ideally, one would like to compare this with how subject-to-subject raising works in Malay. However, Malay does not have such raising predicates (Yosuke Sato p.c.) and as such it is not possible to see these cases. However, from English, we know that subject-to-subject raising does not typically involve matching effects.

- a) John seemed to like Mary.
- b) John seemed to be liked by Mary.

I will thus assume that (52) cannot be ruled out by appealing to voice mismatch.

- (52) *Ali_i senang [(untuk) [_{vP} ____i me-masak nasi ayam]].
 Ali easy for ACT-cook chicken rice
 Lit: ‘Ali is easy to cook chicken rice.’

If v^R is why the embedded clause does not have a PRO, then Wu’s analysis predicts that voice matching is always correlated with lack of PRO. This is because v^R by hypothesis comes unvalued for phi *and* voice features and it copies both these values from the matrix voice head. But this cannot be the case in (51). In this sentence, there is no embedded PRO but yet voice matching does not obtain. In other words, in this sentence, v^R is not already valued for phi features (since there is no PRO) but is already valued for voice features (since there is no voice matching).²⁸ This type of split is not expected in W&S’s account.²⁹

²⁸ One might suppose that there is an embedded PRO in (51) but that the ungrammaticality of (52) is due to some other reason. One might postulate that a *tough* predicate imposes certain semantic restrictions on what type of surface subject it can have which restricts the application of A-movement. However, subject-to-subject raising predicates such as *seem* and *likely* in general do not impose such semantic restrictions. In addition, at least historically, English allowed the embedded PRO to be raised as a subject (Anderson 2001).

- a) He...found the natives...very hard to believe that the fact was possible. (1726-7 Swift *Gulliver* III.x; OED)

Anderson (2001) provides corpus data such as (a) that indicates that an embedded subject can be raised in a tough construction in late 18th century English. The fact that such sentences are not possible now suggests that the subcategorization requirements of the matrix predicate has changed in Modern English.

²⁹ There is one type of split discussed in W&S where v^R comes unvalued for phi features but has default voice values which is often spelled out as active. W&S proposes that languages like German have such a v^R . However, this cannot be the case for the Malay expletive construction. This is because if Malay was capable of inserting a default voice value in v^R , we expect to see this in TM as well. But this is not the case. In addition, this voice head is clearly

What these indicate is that what is a defining characteristic of a restructured complement in Malay is that the embedded clause lacks an external argument PRO. Voice matching itself is required only if long A-movement takes place. In terms of the theory I am developing, in Malay TM where long A-movement takes place, v^R copies both its phi features and voice features from the matrix voice head. However, in the expletive construction even where there is no A-movement, there is a v^R that comes unvalued for phi features *but is already valued for voice features*. This accounts for why the embedded clause in (51) does not have a PRO but also no matching voice value. I propose the following analysis of the expletive construction in Malay with respect to feature valuation.

(53)	Adalah	v _{[voice: PASS][iφ: EXP]} -senang	[v^R _{[voice: ACT][iφ: EXP]} -masak	nasi ayam].
	COP	easy		ACT-cook	chicken rice

(53) shows an expletive construction which also has a v^R in the embedded clause. In this case, only the phi features of v^R comes unvalued which prevents the generation of an embedded PRO. These phi features are copied from the matrix voice head which gives us the required interpretation. However, this v^R comes valued for voice features, in this case active. Since this is an active voice head, this voice head can assign accusative case to the embedded object.³⁰

Based on this discussion, I propose that v^R (at least in Malay) comes in two types.

capable of assigning accusative case to the embedded object, something not expected of a voice head spelled out with default active voice (W&S).

³⁰ This is a further argument against Burzio's Generalization (Burzio 1986) which states that the ability of the v head to have an external argument is connected to its ability to assign accusative case. However, Burzio's generalization has been shown to be untrue for many languages, Hindi being a prominent example (Mahajan 2000, Bhatt 2005).

- (54) a. Partial restructuring voice: $v^R_{[\text{voice: } \alpha][\text{i}\phi: ___]}$
 b. Full restructuring voice: $v^R_{[\text{voice: } ___][\text{i}\phi: ___]}$

The restructuring voice head that occurs in the expletive construction is in (54a) and comes unvalued only for phi features which restricts the presence of a subject PRO. Voice values are fully specified. In the case of the expletive construction in (51), it is active. This ensures that the embedded object can be assigned case. The full restructuring head in (54b) comes unvalued for voice and phi features. These values are copied from the matrix voice head. This is the voice head that occurs in Malay TM. I assume that TM in Malay is triggered by the matrix T probing the embedded object to check its case features and that this is also what triggers movement of the embedded object to the matrix Spec, TP. What the Malay facts then indicate is that full v^R is required if and only if the matrix T probes the embedded object.³¹ Partial v^R is used in a restructuring context if and only if the matrix T does not probe the embedded object.

In the expletive construction in (51), the embedded clause has partial v^R that comes valued with active voice values. One might wonder if it is possible for this to come valued with passive voice values. The following is a candidate for such a derivation.

- (55) a. Adalah senang [untuk di-pukul (oleh Ali)]
 COP easy for PASS-beat by Ali
 ‘It was easy to be beaten (by Ali).’
 b. ... $V_{[\text{voice: PASS}][\text{i}\phi: \text{EXP}]}$ -senang [$v^R_{[\text{voice: PASS}][\text{i}\phi: \text{EXP}]}$ -pukul ...

(55) shows an expletive construction which has a passive embedded clause, but here, the matrix

³¹ I will suggest an explanation for this in a later section using the notion of domain extension (cf. den Dikken 2007).

T has not probed the embedded object. This appears to be what partial v^R with inherent passive values might look like, shown schematically in (55b). But this cannot be the case. In fact, the embedded clause in (55) cannot have a v^R at all. The reason for this comes from its interpretation. Recall that if the embedded voice head was partial v^R , it would come unvalued for phi features which can then be valued by the matrix voice head's phi features which corresponds to the implicit matrix experiencer. This should lead to obligatory co-construal between the matrix experiencer and the embedded agent. This is shown in (55b). However, this interpretation is not possible in (55a). This sentence only has the interpretation where the matrix experiencer is co-construed with the embedded patient.³² This must mean that embedded clause has a voice head that comes with all of its values specified.³³ Thus, the optionality of long A-movement in Malay *tough* predicates comes down to whether there is a full v^R or partial v^R in the embedded clause.

Wurmbrand (2001) also discusses optionality of restructuring effects in German and Japanese but proposes a different characterization. I will discuss German as it is more parallel with the Malay TM facts. Consider the following.

³² I will assume that in (55a), there is an object PRO that remains in situ as there is no Spec, TP in this structure for it to move to. The interpretation of this sentence comes about by the matrix experiencer controlling the object PRO. This also provides us a potential explanation for why (55a) with a partial v^R is not possible. In such a structure, there is no way for the object to receive an interpretation, given that Malay does not allow an object *pro*.

³³ This indicates that a predicate like *senang* c-selects a vP with no external argument whether it be a vP which is headed by v^R or an inherently passive voice head.

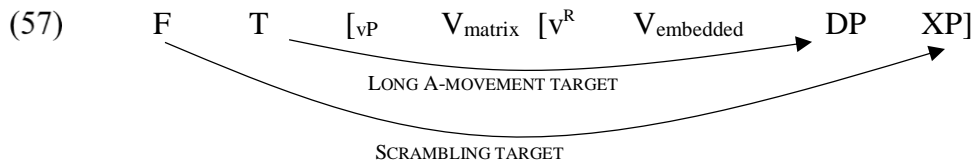
- (56) a. dass der Traktor und der Lastwagen zu reparieren versucht
 that [the tractor and the truck].NOM to repair tried
 wurden
 were
 ‘that they tried to repair the truck and the wagon.’ (Wurmbrand 2001: 19)
- b. dass versucht wurde den Traktor un den Lastwagen
 that tried was [the tractor and the truck].ACC
 zu reparieren
 to repair.
 ‘that they tried to repair the truck and the wagon.’ (Wurmbrand 2001: 38)

(56a) shows a restructuring verb, unaccusative *try* in German, where long A-movement has taken place. Thus, the logical embedded object is realized in the matrix clause with nominative case and triggers phi agreement with the matrix auxiliary. (56b) shows the variant where A-movement has not taken place. Here, the embedded object is realized in the embedded clause with accusative case. The matrix auxiliary shows singular agreement. Thus, the German data is parallel to Malay TM in that in both cases, there is apparent optionality of long A-movement.

While I have analyzed both TM and the expletive construction in Malay as having a restructured complement, albeit different types, Wurmbrand (2001) proposes that (56) shows that the German unaccusative *try* can optionally take a restructuring complement or not. When it does, long A-movement takes place and when it does not, the object remains in situ. As evidence for this, Wurmbrand shows that scrambling which is only possible out of a restructuring infinitive in German is possible in (56a) but not in (56b). I will not repeat these well-known data

here. The generalization in German is that scrambling is possible out of an infinitive clause if and only if long A-movement has taken place.

The distinction between partial v^R and full v^R and the claim that full v^R is present if and only if matrix T probes the embedded object allows us to think about the optionality of German long A-movement in a different way than that suggested by Wurmbrand (2001). Once we allow for independent valuation of the voice and phi values of a voice head, we can extend the same partial v^R and full v^R analysis to the German sentences in (56). So, why is scrambling only possible with full v^R in this perspective? To see why, I will first assume that scrambling is also triggered by a matrix head probing the scrambled element. This is shown below.



(57) shows a restructuring complement prior to long A-movement where the matrix T probes the embedded DP object. As argued above, this must mean that the v^R is a full v^R . Since scrambling is possible in this context, this indicates that a full v^R is needed if *any* matrix head probes *any* phrase within the embedded clause. On the other hand, partial v^R which does not allow long A-movement also disallows any other matrix head from probing any phrase within the embedded clause. Thus, scrambling is also disallowed when there is no long A-movement. Based on this, I conclude that the optionality of A-movement in German restructuring contexts is compatible with the same analysis given to Malay optionality. In other words, both the infinitival clauses in (56) lack an embedded external argument PRO, i.e. they are both restructured clauses. We do not have to posit widespread ambiguity for every predicate that allows for long A-movement and has

a non-A-movement variant as well. Where they differ is that (56a) has an infinitival clause with a full v^R that allows a matrix head to probe into the embedded clause whereas the infinitival clause in (56b) has partial v^R which makes the embedded vP opaque for any such probing.³⁴

4 Phase extension and A-movement

In this section, I will propose an explanation as to why full v^R is required if and only if a matrix T probes the embedded object.³⁵ I will suggest that full v^R duplicates the matrix voice head which effectively extends the phasal domain of the embedded vP to the matrix vP. This is an instance of phase extension (cf. Den Dikken 2007). This is what prevents the spell out of the lower vP (including the embedded object) before the matrix T is merged. In order to see how this account works, I will first outline some assumptions regarding phases and the nature of A-movement.

4.1 Assumptions of phases and spell out

I will make the following assumptions for what follows.

- (58)
- a. vP (passive, unaccusatives and active) and finite CP are phases.
 - b. The complement of a lower phase (VP, TP) is sent to SPELL OUT when a higher phase head is merged.
 - c. Elements in the spelled out phrase are inaccessible for probing by elements outside the spelled out phrase.

³⁴ Certain restructuring predicates do not have expletive variants. For example, German only has a TM variant but not the expletive variant (Wurmbrand 2001). In the proposed analysis, this reduces to whether a restructuring predicate allows an embedded clause with partial v^R or whether it always requires full v^R .

³⁵ The discussion is outlined with a matrix T probing the embedded object but it would apply to any matrix head probing an embedded XP in a restructuring context.

- d. A-movement is successive cyclic through Spec, TP.
- e. A-movement cannot use the edge of phases as escape hatches.

(58a) is standard. Chomsky (2000) argues that unaccusatives and passive vPs are not strong phases but I will follow Legate (2003) who argues that such vPs are also strong phases. (58b) describes my assumption about when SPELL OUT occurs. This is adopted from W&S and can avoid a look ahead problem observed by Boskovic (2007) that plague the somewhat standard assumption that the complement of a phase head is sent to spell out when the phase itself is completed. Consider what happens in cases of wh-movement in a sentence like the following.

(59) What_i did John_j [_{VP} <what_i> [_{VP} t_j [_{VP} eat <what_i>]]]?

In (59), the wh-object moves to the edge of the vP and then to CP. However, if VP (which contains the object) is spelled out when vP is completed, the object will be sent to spell out along with the VP. In order to avoid spell out, the object must be moved out of the vP. But the problem is that only C (which is merged much later) can indicate whether the wh-object needs to be moved at all. This is the look ahead problem. However, if VP is spelled out only when the higher phase head, C, is merged as assumed here, the look-ahead problem disappears. If the C that is merged has WH features, then the object can move out of VP before VP is sent to SPELL OUT. (56c) is standard and follows from the phase impenetrability condition (PIC, Chomsky 2000).

(58d) and (58e) are assumptions to do with A-movement in particular. (58d) indicates that long distance A-movement is successive cyclic in that if there is an intermediate Spec, TP position in the structure, A-movement must land there (Landau 1999, Chomsky 2000, Boskovic 2002, Abe 2016). The following shows the movement steps involved in a sentence shown below.

(60) John_i seems [_{TP} <John_i> to [_{VP} <John_i> like Mary]].

The crucial step is shown in the intermediate copy where the embedded subject moves to the intermediate Spec, TP before moving to the matrix Spec, TP position. This movement can be thought of as successive cyclic movement in the sense of Boskovic (2002).

(58e) may well be the most controversial assumption made here. I assume that A-movement (unlike A'-movement) does not move to the edge of phases to escape SPELL OUT. (59) shows a context of A'-movement which requires adjunction of the wh-object to the vP in order to escape SPELL OUT. This adjunction is sometimes thought of as a need to check D features (Sato 2012) or to check EPP features (Aldridge 2008) in the context of Malay/ Indonesian. Whatever it is, I assume that this is not possible for A-movement. One evidence for this assumption is that A-movement is only possible across a non-finite clause boundary.³⁶

- (61) a. John seems [_{CP} to like Mary].
 b. *John seems [_{CP} that likes Mary].

(61) shows this distinction. Chomsky (2000) explains (61b) through the ACTIVATION CONDITION. Since the embedded T checks the case feature of *John* in the embedded position, this makes it inactive for probing by the higher T. In (61a), the embedded T does not check the case features of the embedded subject and thus the subject remains active for further probing by the matrix finite T. I will follow Saito (2004) and Boskovic (2007) and assume that the ACTIVATION CONDITION does not exist. This means that we need an alternative explanation for why (61b) is bad. I propose that this arises from the fact that A-movement cannot freely use the edge of a

³⁶ Lubukusu has been argued to allow A-movement across a finite clause boundary (Diercks 2012). In such languages which are few, conditions on adjunction, spell out and phase hood of certain phrases may be different.

phase to escape. I assume this applies to CPs and vPs.³⁷ This can explain why (61a) but not (61b) is ungrammatical as discussed directly below.

(62) John [_{vP} seems [_{TP} <John_i> to [_{vP} <John_i> like Mary]]]

(62) is the derivation of (61a) which has a non-finite clause boundary. Here, the embedded subject moves to the embedded Spec, TP. When the matrix vP is completed, *John* does not need to be at the edge of the matrix vP because the matrix VP will only spell out when the matrix C is merged (assumption (58b)). Thus, the matrix T can probe the embedded subject without violating the PIC.³⁸ The situation is different with a finite clause boundary. i.e. a phasal C.

(63) *John [_{vP} seems [_{CP} that [_{TP} <John_i> [_{vP} <John_i> like Mary]]]

(63) shows the derivation of (61b). In this structure, the embedded subject also moves from the embedded Spec, vP to the embedded Spec, TP. Then, comes the CP phase. The complement of this CP phase, i.e. the embedded TP, is spelled out when the matrix v is merged. This explains why the matrix T cannot probe the embedded subject, i.e. it is already spelled out by the time matrix T is merged. However, this must mean that *John* cannot be moved to the edge of the CP as this would allow the matrix T to probe it and (63) would be a grammatical sentence. Saito

³⁷ This contrasts with Keine (2016, 2017) who takes these data to argue that CPs and vPs are different. Later when we discuss Hindi LDA, we will see that there is no movement of the embedded object and yet the matrix T can probe into the embedded vP which should have been spelled out. This assumption I make here about the inability of the embedded object to use the edge of vP as an escape hatch during A-movement will make possible a uniform analysis for Malay TM and Hindi LDA.

³⁸ This is also compatible within a feature inheritance model (Chomsky 2005). When the matrix C is merged, it will transfer its features to matrix T which will then probe the embedded subject in the embedded Spec, TP position which will trigger its movement to the matrix Spec, TP. Then, the matrix VP will be spelled out.

(2004) suggests that this ban on adjunction in A-movement may be due to the fact that adjunction of this type is A'-movement which then leads to an improper movement violation when the DP moves to the matrix Spec, TP. I will assume that that this is the right way to think about this restriction.³⁹ However, I extend this restriction to the edge of vP as well.

Moreover, the assumptions made about phases and spell out here ensure that adjunction to vP is not needed to escape a spell out domain for typical cases of A-movement either.

(64) [CP [TP John_i was [vP [vP beaten <John_i>]]]]

(64) shows a passive clause where the embedded object A-moves to Spec, TP. Based on the assumptions I have outlined, the object does not need to move to the edge of vP before moving to Spec, TP. This is because the VP will be spelled out when C is merged, but this takes place only after T has been merged. This means that T is able to probe the object in the object's base position without incurring a PIC violation.

Long subject-to-subject raising also follows in the same vein.

(65) a. John seems to be certain to like Mary.
 b. John [vP seems [TP <John_i> to be [vP certain [TP <John_i> to [vP <John_i> like Mary]]]]].

(65b) shows the derivation of the sentence in (65a). In this derivation too, *John* never needs to adjoin to the edge of any intermediate vP. This is because *John* needs to move to every embedded Spec, TP successive cyclically. Let's consider each step of the derivation. First, *John* is base-generated in the lowest Spec, vP. It then moves to the lowest Spec, TP. When the

³⁹ Although see Keine (to appear) for a contrary view.

intermediate vP is completed (immediately containing *certain*), the lowest VP is spelled out but *John* can remain in the lowest Spec, TP because the intermediate VP is not spelled out yet. When the intermediate T is merged, it probes *John* which triggers movement of *John* to the intermediate Spec, TP again. When the highest v is merged, the intermediate VP will be spelled out, but by now, *John* has again already left the intermediate VP. When the highest T is merged, *John* moves to the highest Spec, TP again. Notice that at no point in the derivation does *John* need to move to the edge of any vP to escape spell out. The presence of specifiers of TP to which the DP must move to for independent reasons ensures that the DP will always leave the complement of a phase before it is spelled out. Crucial ingredients of this analysis are assumptions (58b) and (58d).

4.2 Full v^R as phase extension

We are now ready to explain why full v^R and matrix T probing the embedded object are correlated. Note that on my assumptions, in long A-movement in restructuring contexts, matrix T should never be able to probe the embedded object. Consider the following to see why.

(66) ...T [_{vP} v-easy [_{vP} v-cook DP]]

(66) shows the partial schema of the Malay TM sentence in (2b). In this structure, matrix T has to probe the DP but the DP is in a VP that has been spelled out. Recall that a complement of a phase is spelled out when a higher phase head is merged. This means that the lowest VP containing the object is spelled out when the higher v head is merged. When T is merged, it should no longer be able to access the embedded object due to the PIC. We cannot solve this problem by adjoining the embedded object to the intermediate vP because, as I have argued above, this type of adjunction is not available for A-movement. (65) shows that in long A-

movement without a finite clause boundary, what we need are intermediate Spec, TPs which allow a DP to escape a spell out domain without having to adjoin to the phase edge. However, in Malay TM (and in restructuring in general) such an intermediate TP node is not available either. Thus, the assumptions I have outlined have led to an impasse where it is not clear how matrix T could probe the embedded object without violating the PIC. This is, of course, by design.

This is where full v^R comes in. I propose that what full v^R does is to replicate the matrix voice head in the lower clause and in doing so extends the embedded vP boundary all the way to the matrix vP. Thus in (66), even though there are two vPs which usually indicate two phases, by making the two v heads identical, the higher vP boundary is treated as the only one vP phase in the structure. In other words, when there are two distinct voice heads in the structure, there are two phasal boundaries between the matrix T and the embedded object. However, with full v^R , there is now effectively only one phasal boundary between the matrix T and the embedded object. This means that the embedded VP in (66) will only be spelled out as part of the matrix VP when the matrix C is merged. When matrix T is merged, the embedded VP is not spelled out yet and as such, T can probe the embedded object in its base position without a PIC violation.

However, in order for this type of phase extension to take place, it is crucial that the embedded voice head is a full copy of the matrix voice head (at least in Malay). This can only happen if there is a full v^R in the embedded clause. If the embedded clause has partial v^R , then the matrix and embedded voice heads are not identical and as such there can be no phase extension. In this case, by the time the matrix T is merged in the structure, the embedded VP containing the embedded object is spelled out. This means that matrix T cannot probe the embedded object. Thus, phase extension provides an explanation for why full v^R is correlated with matrix T probing

the embedded object. This follows from the assumptions about phases and SPELL OUT outlined earlier.⁴⁰

While the notion of phase extension is borrowed from Den Dikken (2007), the way it is implemented here is quite different. Den Dikken (2007) was primarily concerned with accounting for predicate inversion in the context of certain types of copular clauses and Scandinavian object shift. He defines phase extension as follows.⁴¹

(67) **Phase extension**

Syntactic movement of the *head* H of a phase α up to the head X of the node β dominating α *extends* the phase up from α to β ; α loses its phasehood in the

⁴⁰ Towards the end of the preparation of this manuscript, I became aware of Sugimura (2011) who also employs phase extension in restructuring but her focus is Japanese restructuring predicates. Nonetheless, there is direct precedent for my approach here. However, she proposes that Japanese employs head movement of V to T to accomplish domain extension while in my approach there is no head movement. Thus, although her approach is more faithful to how Den Dikken (2007) formalizes phase extension, my approach is applicable to even languages which do not have head movement, such as Malay.

⁴¹ This idea of domain extension exists in various forms in the literature. Chomsky (1993) argues that head movement can extend the domain of the lower head, although he does not couch his analysis in terms of phases. Gallego & Uriagereka (2006) argue for ‘phase sliding’ which differs from den Dikken’s phase extension in that there is no ‘transfer’ of phasehood from the lower head to the higher one. I remain open to the idea that these (or other) implementations of how a lower domain is extended could be the right way to think about why full v^R is needed for the matrix T to probe the embedded object. What is crucial for me is that full v^R leads to such domain extension (whatever its specifics). Den Dikken’s theory is immediately applicable for my purposes.

process, and any constituent on the edge of α ends up in the domain of the derived phase β as a result of Phase Extension. (Den Dikken 2007: 1)

The crucial part of his theory for us is that when a lower phase head moves to a higher head, the lower phrase ceases to be a phase.

(68) [XP X+Z [ZP <Z> ...]]

In (68), Z is the phase head which projects ZP which would be counted as the phase. However, Z head-moves to X. This head movement ceases ZP's status as a phase, but now XP is the phase.

In other words, the phrase that contains the unpronounced copy of the phasal head no longer counts as a phase. However, the higher phrase which contains the pronounced copy of the phase head counts as a phase. The mechanism required for our purposes is the same as shown below.

(69) Full v^R : T [vP V_{[voice: PASS][iφ: EXP]} [_v^RP V^R_{[voice: PASS][iφ: EXP]} DP]]

(69) shows the voice heads in Malay TM. By having a full v^R in the embedded clause, the higher phase head is fully replicated in the embedded clause. Representationally, this looks the same as if head movement had taken place from the lower v to the higher v . The only difference is that both copies of v can potentially be pronounced, due to the fact that this is not actually a movement chain. Crucially for us, the lower vP ceases to be a phase and there is now only one phase boundary between the matrix T and the embedded object in its base position. This is what allows matrix T to now probe the embedded object without violating the PIC. If the embedded v^R is only partial, then there is no identity between the two voice heads and thus no phase extension of the lower vP. The matrix and embedded vP remain distinct phases. This means that there remain two phase boundaries between the matrix T and the embedded object in its base position. Matrix T will thus not be able to probe the embedded object because the lowest VP containing

the embedded object will be spelled out by the time T is merged. This is what we get in the Malay expletive *tough* construction.

In summary, I have argued that the passive morphology in Malay TM is an instance of voice matching, following W&S. However, contrary to W&S, I have argued that voice matching is correlated with a matrix T probing the embedded object, not restructuring per se. On the basis of this, I motivated the notions of partial v^R and full v^R and full v^R is argued to lead to phase extension which is what allows the matrix T to probe the embedded object in its base position without violating the PIC. When there is only partial v^R , the embedded clause is still restructured (i.e. lacks an external argument PRO), but the PIC prevents probing of the embedded object by the matrix T. In my account, voice matching occurs in restructuring contexts when and only when full v^R is present. This is contrary to W&S who proposes that (full) v^R is a necessary condition for restructuring. I believe the way I have characterized v^R , extends W&S's theory in a principled way to account for the Malay *tough* construction data while retaining the empirical coverage for the cross-linguistic restructuring facts. I will thus assume that this is the right way to think about restructuring with and without long A-movement in Malay TM.⁴²

In the next section, I discuss long distance agreement (LDA) in Hindi. This has been argued by Bhatt (2005) to be a restructuring context which involves matrix T probing the

⁴² German TM allows long A-movement with default active voice valuation in the embedded clause. This difference between German and Malay indicates that German has a pseudo-full v^R and not full v^R in the sense that German full- v^R is merged, already valued with default voice (spelled out as the active form) which cannot assign accusative case. I propose that the difference between Malay and German is that phase extension is possible in German with this pseudo-full v^R whereas Malay requires full v^R .

embedded object for agreement without long A-movement of the embedded object. I show that such contexts exhibit a ban on voice mismatch which indicates that Bhatt is right in analyzing these as restructuring and also that there is a v^R in restructuring contexts.

5. Hindi long distance agreement

Long distance agreement (LDA) in Hindi takes place with certain types of predicates which take an infinitival clause. I will assume with Bhatt (2005) and Keine (2017) that LDA is an instance of the matrix T probing the embedded object for phi agreement.

- (70) Ram-ne [rotii khaa-nii] chaah-**ii**
 Ram-ERG bread.F eat-INF.F want-PFV.F.SG
 ‘Ram wanted to eat bread.’ (Mahajan 1989, from Bhatt 2005: 761)

In (70), *want* (a common restructuring verb) exhibits LDA with the embedded object.^{43, 44} LDA is not possible with certain other predicates.

- (71) Anjum-ne saddaf-ko [citt^{hi} likh-ne]-ko kah-**aa**
 Anjum-ERG Saddaf-DAT note.F write-INF.OBL-ACC say-PFV.M.SG

⁴³ In Hindi, T agrees with the highest morphologically unmarked argument. Thus, if (70) is in present tense, the subject of the Hindi clause will be in unmarked nominative, and in this case T will agree with the matrix subject and LDA is no longer possible. See Butt (1995, 2014), Bhatt (2005) and Davison (2013) for more details.

⁴⁴ This agreement is optional for many speakers and as such can be realized with default masculine agreement as well. In addition, LDA in such constructions also exhibits a matching effect where the embedded clause must show the same agreement morphology as the matrix verb. Due to space, I will not deal with these aspects of Hindi LDA in detail in this paper but will remark on these in a footnote after the main analysis. See fnnt 49.

‘Anjum told Saddam to write a note.’ (Butt 2014: 177)

Thus, the verb *tell* in Hindi in (71) which also takes an infinitival complement does not allow LDA between matrix T and the embedded object. Here, only the default masculine agreement form is allowed.⁴⁵

Bhatt (2005) provides evidence from NPI licensing to argue that LDA is possible only in restructuring infinitives, which he takes to be a context in which there is no embedded external argument PRO.

(72) ek-bhii larke-ne [Sita-kii kitaab nahi: par:h-nii] chaah-ii
 one-PSI boy-ERG Sita-GEN.F book.F NEG read-INF.F want-PFV.F.SG

‘Not even a single boy wanted to read Sita’s book.’ (Bhatt 2005: 780)

(72) shows that with *want*, a restructuring infinitive that allows LDA, an NPI is licensed as the matrix subject. Note however that the NPI licensor, the negation marker, occurs in the embedded clause. In contrast, *tell*, which does not allow LDA, does not allow such NPI licensing.

(73) *ek-bhii lar:ke-ne Sita-se [kitaab nahi: par:h-ne]-ko kah-aa
 one-PSI boy-ERG Sita-INST.F book.F NEG read-INF.OBL-ACC say-PFV.M.SG

(Bhatt 2005: 780)

(73) shows that a negation in the embedded infinitival does not license a matrix subject NPI licensor when with the *tell* verb.⁴⁶ Bhatt (2005) argues that NPI licensing in Hindi is an indicator

⁴⁵ Such predicates also require the infinitival clause to be accusative case-marked which suggests that these are gerunds (Butt 1995).

⁴⁶ See Bhatt (2005) for control sentences that show that a matrix negation licenses a NPI subject with a *tell* verb as well as other details regarding NPI licensing in Hindi.

of whether the infinitive is transparent for processes which are otherwise bound within the clause, i.e. restructuring. I will adopt Bhatt's argument that this is indeed the crucial difference between *want* and *tell* in Hindi.⁴⁷

Furthermore, I will assume following Bhatt (2005) that in Hindi LDA, the embedded object remains in situ and does not need to move. This assumption is not universally held. For example, Keine (2013) argues that the embedded object must move to the edge of the embedded clause in order for LDA to take place. However, I will follow Keine (2017) and Bhatt & Keine (2017) and propose that A-movement is not necessary for LDA in Hindi. One convincing piece of evidence that Keine (2017) and Bhatt & Keine (2017) provide for this is idiom chains.

- (74) a. raam-ne bhains ke aage biin bajaa-yii
 Ram-ERG buffalo in.front.of flute.F play-PFV.F.SG
 'Ram did something futile.'
- b. #biin_i raam-ne bhains ke aage t_i bajaa-yii
 flute.F Ram-ERG buffalo in.front.of play-PFV.F.SG
 'Ram did something futile.' (Keine 2017: 178-179)

Keine (2017) first shows that the idiom in (74a) does not allow A-movement, i.e. local scrambling (Mahajan 1990), of the embedded object. Thus, (74a) has the idiomatic meaning but (74b) is reported to have only the literal meaning of playing the flute in front of a buffalo. What

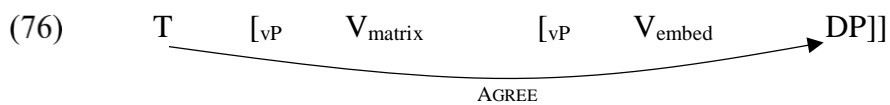
⁴⁷ Davison (2013) provides several arguments that the embedded clause with *want* verbs do have an external argument PRO which would be problematic under my analysis, as this would mean that *want* is not a restructuring verb in Hindi after all under my definition. However, Butt (2014) argues convincingly against Davison's evidence and claims that there is no PRO in the embedded clause of *want* verbs.

idioms like this do is allow us to reason as follows: If we put these idioms in the embedded clause of a construction that otherwise allows LDA, the presence of LDA and the idiomatic meaning indicates that the embedded object need not move in order to allow LDA. This is exactly what we find.

- (75) raam-ne [bhains ke aage biin bajaa-nii] chaah-ii
 Ram-ERG buffalo in.front.of flute.F play-INF.F.SG want-PFV.F.SG
 ‘Ram wanted to do something futile.’ (Keine 2017: 179)

(75) shows that the idiomatic meaning is possible with LDA. This then shows that the embedded object can be in situ and still trigger LDA.⁴⁸

So far, we have seen evidence for the claim that LDA in Hindi occurs in restructuring contexts and that the embedded object remains in situ. This leads to a PIC problem that is also noticed by Keine (2017). If the vP is a phase, then how is it that the matrix T can probe the embedded object which is separated from it by two phase boundaries. This is the same problem we encountered in Malay TM, made more obvious here by the fact that the embedded object in Hindi LDA need not move at all. This is shown schematically below.



Even on our assumptions, by the time matrix T is merged, the embedded VP would have been spelled out and the configuration in (76) would reflect a PIC violation. Based on this, Keine (2017) suggests that vPs should not be considered to be phases. But I believe that the approach to

⁴⁸ Bhatt & Keine (2017) claim that with such idioms, LDA is preferred even though the default masculine agreement is also possible.

restructuring pursued here is potentially better as it can unify Hindi LDA with A-movement in Malay TM. Recall that I have argued that in order for matrix T to probe the embedded object in a restructuring context, the embedded voice head must be full v^R . Consequently, if the embedded voice head in (76) is also a full v^R head, this would be a copy of the higher v head. This extends the phase boundary of the embedded vP such that there will only effectively be one phase between the matrix T and the embedded object. And since in my assumptions, the complement of a lower phase only spells out when a higher phase head is merged, T can probe the embedded object for agreement without violating the PIC. Thus, whether long A-movement takes place or LDA takes place in a restructuring context, the underlying reason for why both of these are possible has a unified explanation, i.e. the presence of full v^R and thus phase extension.

This approach can also account for the following sentence from Keine (2017).

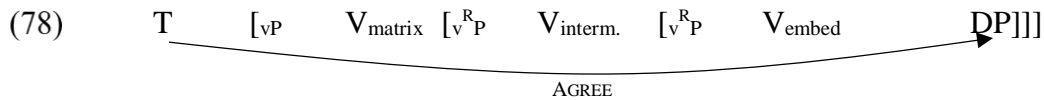
(77)	raam-ne	[[bhains	ke aage	biin	bajaa-nii]	shuruu
	Ram-ERG	buffalo	in.front.of	flute.F	play-INF.F.SG	start
	karnii]	chaah-ii				
	do-INF.F	want-PFV.F.SG				

‘Ram wanted to start to do something futile.’ (Keine 2017: 180)

In this sentence, there is LDA and the idiomatic meaning. However, there are three vP boundaries in the way. Notably, the highest verb and the intermediate verbs are both independently restructuring verbs.⁴⁹ By hypothesis, this means that the intermediate voice head and the lowest voice head are both full v^R as shown below.

⁴⁹ The status of *shuruu* ‘start’ as a restructuring verb can be seen in the following example from Hook (1979: 30)

a)	us-ne	[merii	aaraam-kursii	tor:-nii]	shuruu	ki-ii
----	-------	--------	---------------	-----------	--------	-------



(78) shows that the matrix T can probe the embedded object contained within the lowest VP.

However, this does not violate the PIC because in the approach I propose, the phasal boundary of the embedded vP is extended all the way to the matrix vP since all three voice heads are identical. Thus, regardless of the fact that there are three vP boundaries between the matrix T and the embedded object, phase extension through full v^R delays the spell out of the lowest VP long enough such that the matrix T can probe the embedded object for phi agreement.

One of the important predictions that this approach to Hindi LDA makes pertains to voice mismatch. Recall that I have argued that full v^R is correlated with matrix T probing the embedded object. Since Hindi LDA is also an instance in which matrix T probes the embedded object in a restructuring context, if the v^R approach to restructuring is right, we predict that voice mismatch should also be disallowed in Hindi LDA. This is because the embedded voice head has to be identical to the matrix voice head and as such it must have the same voice value as the matrix voice head. Since in these LDA sentences, the matrix voice head is active, the embedded clause must also be in active voice. An embedded passive with LDA should be disallowed. The

he-ERG my.F rest-chair.F break-INF.F start do-PFV.F.SG

‘He began breaking my easy chair.’

If Bhatt (2005) is right that only restructuring verbs allow LDA, then *shurui* ‘start’ which allows LDA must also be considered a restructuring verb.

evidence indicates that this is indeed correct.⁵⁰ First, observe that in the absence of LDA, a *want* verb can have a passive embedded clause.

- (79) Ram ne [maar khaana] chah-aa
 Ram ERG beat.INF PASS want-PFV.M.SG
 ‘Ram wanted to be beaten.’

(79) shows a restructuring verb with a passive embedded clause, a voice mismatch. This is not a problem for my theory because there is no DP that the matrix T could agree with and as such there is no full v^R in the embedded clause. In this case, the matrix verb spells out with default masculine agreement. Note that *want* can alternatively take a non-restructuring finite clause complement (Butt 2014). Thus, the following is also possible.

- (80) Ram ne chah-aa [ki maar-aa jaaye]
 Ram ERG want- PFV.M.SG C beat- PFV.M.SG PASS
 ‘Ram wanted to be beaten.’

Here, *want* takes a finite embedded clause introduced by the complementizer *ki*. As shown by Butt, in these cases LDA is never possible even if there is an embedded object. Thus, these are never restructured contexts. What (79) and (80) show is that in the absence of LDA, there is no need for voice matching between the matrix and embedded clauses with a *want* predicate. The matrix clause is active and the embedded clause can be passive. However, the possibilities are different when there is LDA.

⁵⁰ This then is highly suggestive that Keine’s (2017) claim that vPs are not phases may not be correct as his approach does not account for this ban on voice mismatch.

- (81) a. Ram-ne [sita-ko kitaab den-ii] chah-ii
 Ram-ERG sita-DAT book.F give-INF.F want-PFV.F.SG
 ‘Ram wanted to give Sita the book.’
- b. *Ram-ne [kitaab den-ii jaaye] chah-ii
 Ram-ERG book.F give-INF.F PASS want-PFV.F.SG
 ‘Ram wanted to be given the book.’

(81a) shows LDA with the embedded object is possible with an active distransitive embedded clause under *want*. This is expected. (81b) however shows that the embedded clause cannot be in passive voice. This is unexpected under any current analysis of Hindi LDA. But this is exactly what we predict to be the case if LDA in Hindi requires the presence of full v^R . The presence of LDA indicates that phase extension has taken place which must mean that the embedded voice head must be full v^R . Thus, its voice value must be the same as the matrix voice head, i.e. active. Notably, an embedded passive is possible when *want* takes a finite embedded clause.

- (82) Ram-ne chah-aa [ki kitaab dii jaaye]
 Ram-ERG want-PFV.M.SG C book.F give-F PASS
 ‘Ram wanted to be given the book.’

(82) shows the finite variant of *want*. In this case, there cannot be LDA, i.e the matrix verb must be in default masculine form even though the embedded verb shows agreement. In addition, the embedded clause can be in the passive form. Thus, what we see is that the ban on passive morphology on a clause embedded by *want* is observed only when there is LDA. I take this ban on voice mismatch in Hindi LDA constructions to be confirmation that Hindi LDA does involve restructuring and that it is the presence of full v^R that allows matrix T to probe the embedded

object in restructuring contexts, thus providing a uniform analysis for long A-movement in Malay TM and long distance agreement in Hindi.⁵¹

⁵¹ I have not addressed here two things pertaining to Hindi LDA: 1) the optionality of LDA, and 2) the matching effect we see in where the embedded infinitival clause must show the same phi agreement as the matrix. I will just make some brief comments about these here. With regard to (1), while Bhatt (2005) proposes that default masculine agreement is a sign that the embedded clause is not restructured (see also Keine (to appear) for a similar claim), there is reason to believe that this is not the case. Recall that the licensing of a matrix subject NPI with an embedded negation was taken as a sign that the embedded clause is restructured. This was shown to be the case with LDA. However, this type of licensing is also seen with default agreement.

a) ek-bhii larke-ne [Sita-kii kitaab nahi: par:h-naa] chaah-aa
 one-PSI boy-ERG Sita-GEN.F book.F NEG read-INF.M want-PFV.M.SG
 ‘Not even a single boy wanted to read Sita’s book.’

Thus, (a) shows that both default agreement contexts are also restructuring contexts and this means that the optionality of LDA has to be explained some other way. Furthermore, somewhat surprisingly, it appears that even the default case must be an instance of LDA, because voice mismatch is not tolerated.

b) *Ram-ne [kitaab diya jaa-naa] chah-aa
 Ram-ERG book.F give PASS-INF.M want-PFV.M.SG
 ‘Ram wanted to be given the book.’

This indicates that the general consensus that the default case examples are not examples of LDA may not be right. The matrix T still probes the embedded object when there is default agreement but there is some other reason why actual phi-agreement does not obtain. This is a puzzle that I will have to leave to future research.

With respect to (2), I will assume, following Bhatt (2005) and Keine (2016) that infinitive agreement is dependent on matrix T agreement and that co-valuation occurs, as described by Bhatt.

6 Conclusion

In this paper, I have described Malay TM and shown that it is formed by A-movement of the embedded object to the matrix Spec, TP position. Unlike German TM which also uses this strategy, the embedded clause in Malay TM must be in passive voice. I have then provided an account for this passive morphology using W&S's claim that restructuring requires a voice restructuring head, dubbed v^R . I have further argued that we need to make a distinction between full and partial v^R in order to capture the fact that full v^R is required if and only if matrix T probes the embedded object. I have then suggested that this is the right way to think about the optionality of long A-movement in Malay TM and German long passives. I have then outlined a set of assumptions that allow us to account for this correlation, namely through the notion of domain extension, or more specifically, phase extension. The nature of phases and spell out were argued to be critical to determine when full v^R is required. Full v^R was argued to have a delaying effect on spell out. Finally, I apply this theory to Hindi LDA which is also another instance of a matrix T probing the embedded object. It was shown that a unified account of long A-movement and LDA is possible based on the assumptions outlined here.

The analysis forwarded in this paper has several implications. I will briefly mention some. First, the paper provides evidence for a particular variant of multiple spell out. In this version, A-movement does not use a phase edge to escape PIC violations. This is perhaps the most surprising implication. Whether this can be extended to all cases of A-movement or to just movement out of a restructured complement has to be investigated more thoroughly. If German scrambling is A'-movement, then even A'-movement may be subject to similar restrictions under the right configuration. The paper also provides independent evidence for the notion of phase

extension where spell out can be delayed under certain circumstances. This suggests that phases are identified dynamically and not statically which is in line with recent approaches (Bobablijk & Wurmbbrand 2005, Den Dikken 2007, Takahashi 2011). In so far as the arguments above are sound, this also provides further support for W&S's claim that the voice domain is split into voice and v , and that the former comprises of two features, voice value (active/ passive) and phi value (agent specification). Partial v^R indicates that these values do not need to be correlated and can at least be one way independent. Finally, this paper reaffirms the connection between AGREE and MOVE. Given that we are able to provide a uniform explanation for Hindi LDA and Malay TM, where only the latter consists of movement, AGREE and MOVE have to be treated as having the same mechanism underlyingly.

References

- Abe, J. 2016. How to probe expletives. *Studia Linguistica* 72(1): 76-112
- Aldridge, E. 2008. Phase-based account of extraction in Indonesian. *Lingua* 118: 1440–1469.
- Alsagoff, L. 1992. *Topic in Malay: the other subject*. Stanford University. PhD Diss.
- Anderson, Deborah. 2001. Structural ambiguity in early English tough constructions: Are child grammars deficient or simply different from adult grammars? In *Working Papers in English and Applied Linguistics*, volume 8, pp. 1–24. Research Centre for English and Applied Linguistics, University of Cambridge.
- Arka, I.W. & Manning, C. 1998. Voice and grammatical relations in Indonesian: A new perspective. *Proceedings of the LFG '98 Conference*. Stanford, CA: CSLI Publications.
- Asmah Hj. Omar. 1982. *Nahu Melayu mutakhir* [Modern Malay grammar]. 2nd edn. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Authier, J.M & Lisa Reed. 2009. French Tough-Movement revisited. *Probus* 21: 1-21
- Belletti, A & Luigi Rizzi. 1988. Psych-verbs and theta theory. *Natural Language and Linguistic Theory* 6(3): 291-352
- Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language & Linguistic Theory* 23: 757–807.
- Rajesh Bhatt & Stefan Keine. 2017. Long-distance agreement. *The Companion to Syntax. Second Edition*, ed. by M. Everaert & H. van Riemsdijk. Hoboken, NJ: Wiley-Blackwell.
- Brody, Michael. 1993. Theta-theory and arguments. *Linguistic Inquiry* 24:1–23.
- Bobaljik, Jonathan & Susi Wurmbrand. 2005. The domain of agreement. *Natural Language and Linguistic Theory* 23:809-865.

- Bošković, Željko. 2002. A-movement and the EPP. *Syntax* 5(3): 167-218
- Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:589–644.
- Burzio, L. 1986. *Italian syntax: A government-binding approach*. Dordrecht: Reidel.
- Butt, Miriam. 1995. *The structure of complex predicates in Urdu*. Stanford: CSLI.
- Butt, Miriam. 2014. Control vs complex predication: Identifying non-finite complements. *Natural Language and Linguistic Theory* 32: 165-190
- Chomsky, Noam. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, eds. Stephen Anderson and Paul Kiparsky, 232–286. New York: Academic Press.
- Chomsky, Noam. 1977. On wh-movement. In *Formal syntax*, eds. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. by Kenneth Hale and Samuel J. Keyser, 1-52. Cambridge: MIT Press.
- Chomsky, N. 2000. *Minimalist inquiries: The framework. Step by step: Essays on minimalist syntax* in honor of Howard Lasnik, eds. R. Martin, D. Michaels & J. Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Chomsky 2001
- Chomsky, Noam. 2005. *On phases*. Ms., MIT, Cambridge, Mass.
- Chung (1976)
- Chung, S. 1976. An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7:41–87.
- Chung, Sandra. 2004. Restructuring and verb-initial order in Chamorro. *Syntax* 7.3:199-233.
- Cole, P. & Hermon, G. 1998. The typology of WH movement: WH questions in Malay. *Syntax* 1:221–258.
- Cole, P. & Hermon, G. 2005. Subject and non-subject relativization in Indonesian. *Journal of*

- East Asian Linguistics* 14:59–81.
- Cole, P., Hermon, G. & Yanti. 2008. Voice in Malay/Indonesian. *Lingua* 118:1500–1553.
- Dardjowidjojo, Soenjono. 1978. Sentence Patterns of Indonesian, University of Hawaii Press, Honolulu.
- Davison, Alice. 2013. Non-finite complements and modality in *de-na* ‘allow’ in Hindi-Urdu. *Natural Language & Linguistic Theory* 32: 137-164
- Dikken, Marcel Den. 2007. Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33:1–41.
- Diercks, Michael. 2012. Parametrizing Case: Evidence from Bantu. *Syntax* 15: 253-286.
- Gallego, A & Juan Uriagereka. 2006. Sub-extraction from subjects. ms.
- Guilfoyle, E., Hung, H. & Travis, L. 1992. Spec of IP and Spec of VP: two subjects in Austronesian Languages. *Natural Language and Linguistic Theory* 10:375–414.
- Hartman, Jeremy. 2011. Intervention in tough-constructions. In *Proceedings of the 39th Meeting of the North East Linguistic Society*, eds. Suzi Lima, Kevin Mullin, and Brian Smith, 387–397. Amherst, MA: GLSA.
- Hicks, Glyn. 2009. Tough-constructions and their derivation. *Linguistic Inquiry* 40:535–566.
- Hook, Peter Edwin. 1979. *Hindi Structures: Intermediate Level*, No. 16 in Michigan Papers on South and South-East Asia, Center for South and South-East Asian Studies, University of Michigan, Ann Arbor.
- Keine, Stefan. 2013. On the role of movement in Hindi/Urdu long-distance agreement.

- Proceedings of NELS 42*, Vol. 1, ed. by S. Keine & S. Sloggett. Amherst, MA: GLSA, pp. 273–284.
- Keine, Stefan. 2016. *Probes and their horizons*. PhD dissertation. U. Mass.
- Keine, Stefan. 2017. Agreement and vP phases. *A Schrift to Fest Kyle Johnson*, ed. by N. LaCara, K. Moulton & A.-M. Tessier. University of Massachusetts Amherst: Linguistics Open Access Publications 1, pp. 177–185.
- Keine, Stefan. (to appear). Selective Opacity. *Linguistic Inquiry*
- Keine, Stefan and Ethan Poole. 2017. Intervention in *tough*-constructions revisited. Ms
- Kroeger, P. 2014. External negation in Malay/ Indonesian. *Language* 90(1): 137-184
- Landau, I. 1999. *Elements of Control*. PhD Dissertation, MIT, Cambridge, Mass.
- Landau, I. 2013. *Control in Generative Grammar*. Cambridge Uni Press.
- Legate, J. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34:506–516.
- Legate, J. 2012. Subjects in Acehnese and the nature of the passive. *Language* 88.3:495-525.
- Longenbaugh, Nicholas. 2016. Non-intervention in *tough*-constructions. In *Proceedings of the 46th Meeting of the North East Linguistic Society*.
- Mahajan, Anoop. 1989. Agreement and agreement phrases. In *MIT working papers in linguistics*, eds. Itziar Laka and Anoop Mahajan, Vol. 10, 217–252.
- Mahajan, Anoop. 1990. *The A/A-bar distinction and Movement Theory*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA. Distributed by MIT Working Papers in Linguistics.
- Mahajan, Anoop. 2000. Towards a united treatment of wh-expletives in Hindi and German. In

- Wh-Scope Marking, ed. by Uli Lutz, Gereon Müller & Arnim von Stechow, 317-332
Amsterdam: John Benjamins.
- Montalbetti, Mario and Mamoru Saito (1983). On certain (tough) differences between Spanish and English. In *Proceedings of the Thirteenth Annual Meeting of the North East Linguistics Conference*, Peter Sells and Charles Jones (eds.), 191–198. Amherst, MA.: GLSA publications.
- Mulder, Rene & Marcel den Dikken. 1992. Tough parasitic gaps. In *Proceedings of the 22nd Meeting of the North East Linguistic Society*.
- Nomoto, H. 2008. A unified analysis of funny control. Paper presented at 12th ISMIL. Leiden.
- Nomoto, H & I. Shoho. 2007. Voice in relative clauses in Malay: A comparison of written and spoken language. in *Corpus-based perspectives on linguistics*. ed. Yuji Kawaguchi, Toshihiro Takagaki, Nobuo Tomimori and Yoichiro Tsuruga, 353 – 370. John Benjamins
- Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture*, eds. Simin Karimi, Vida Samiian and Wendy Wilkins, 262-294. Amsterdam: Benjamins.
- Polinsky, Maria & Eric Potsdam. 2008. The syntax and semantics of wanting in Indonesian. *Lingua* 118: 1617 - 1639
- Postal, Paul. 1971. *Cross-over phenomena*. New York: Holt, Rinehart and Winston.
- Rackowski, A., 2002. The Structure of Tagalog: Specificity, Voice, and the Distribution of Arguments. Ph.D. dissertation, MIT.
- Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:565–599.

- Rezac, Milan. 2006. On tough-movement. In *Minimalist essays*, ed. Cedric Boeckx, 288–325. Amsterdam: John Benjamins.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge: MIT Press.
- Roberts (1993)
- Saddy, D. 1991. WH-scope mechanisms in Bahasa Indonesia. *MITWPL* 15, ed. L. Cheng and H. Demirdache, 183–218. Department of Linguistics and Philosophy.
- Saito, Mamoru. 2004. Ellipsis and Pronominal Reference in Japanese Clefts. *Nanzan Linguistics* 1: 21-50
- Sato, Yosuku. 2012. Successive cyclicity at the at the syntax-morphology interface: Evidence from Standard Indonesian and Kendal Javanese. *Studia Linguistica* 32-57.
- Selvanathan, N. 2018. Tough movement and predication. ms.
- Sneddon, J. 1996. *Indonesian: A comprehensive grammar*. Routledge: London.
- Soh, Hooi Ling & Nomoto, Hiroki. 2009. Progressive Aspect, the verbal prefix meN and stative sentences in Malay, *Oceanic Linguistics* 48: 148-171.
- Soh, HL & Nomoto, H. 2010. Degree achievements, telicity and the verbal prefix meN in Malay. paper presented at the 14th *ISMIL*. Minneapolis.
- Soh, HL & Nomoto, H. 2011. The Malay verbal prefix meN- and the unergative/ unaccusatives distinction. in *Journal of East Asian Linguistics* 20: 77-106
- Son, Minjeong & Peter Cole. 2004. Event decomposition and the syntax and semantics of -kan in Standard Indonesian. in *Proceedings of NELS 34*, U.Mass
- Sudaryono. 1993. *Negasi dalam Bahasa Indonesia: Suatu tinjauan sintaktik dan semantik* [Negation in Indonesian: A syntactic and semantic review/survey]. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

- Sugimura, M. 2011. *Domain Extension: A Study of restructuring predicates in Japanese and the role of head movement*. PhD Dissertation. McGill
- Takahashi, Masahiko. 2011. Some consequences of case-marking in Japanese. Doctoral dissertation, University of Connecticut.
- Voskuil, J. 2000. Indonesian voice and A-bar movement. *Formal issues in Austronesian linguistics*, eds. I. Paul, V. Phillips & L. Travis, 195–213. Dordrecht: Kluwer.
- Willet, Marie L. 1993. *Object preposed construction in Malay*. Master's Thesis. Memorial University of Newfoundland.
- Wouk, Fay. 1989. *The use of verb morphology in spoken Jakarta Indonesian*. PhD Diss, UCLA.
- Wurmbrand, Susi. 2001. *Infinitives: Restructuring and Clause Structure*. Mouton de Gruyter. New York.
- Wurmbrand, Susi. 2013. *Complex predicate formation via Voice incorporation*. ms.
- Wurmbrand, Susi & Koji Shimamura. 2017. in *The Verbal Domain*. ed(s) Roberta D'Alessandro, Irene Franco, and Angel Gallego. Oxford Scholarship Online: 179-204.