

Bahasa Indonesia: a window on typology

Lisa deMena Travis
McGill University

1. INTRODUCTION

Bahasa Indonesia (BI) presents a challenge to researchers, particularly typologists, who have become comfortable investigating the syntax of other linguistic relatives within the Western Malayo-Polynesian (WMP) language family such as Malagasy and Tagalog. When studying Malagasy and Tagalog (M/T), the temptation is to look at those properties that distinguish these two languages from English (French, German, Italian, ...: E/F/G/I) and try to account for why M/T exhibit one cluster of characteristics while E/F/G/I exhibit another. However, if one ever creates an account which predicts that property x entails property y , Bahasa Indonesia is bound to be a counter-example.

In this paper, I expose my own attempt at such a typology. In section 2, I outline some ways in which ways BI patterns with M/T making it tempting to say that the BI should fall with M/T in any language typology. In section 3, I show that certain other characteristics that make M/T distinct from E/F/G/I are missing in BI. After an interim summary (section 4), I outline a typology in section 5 that I have proposed to account for the differences between M/T and E/F/G/I and then speculate in section 6 how BI can be made to fit within this typology. I suggest that certain definable syntactic domains in BI are E/F/G/I-like while other domains are M/T-like. The hypothesis that I conclude with is that typologies can vary language internally.

2. SIMILARITIES BAHASA INDONESIA LIKE MALAGASY/TAGALOG

The best-known characteristics that distinguish WMP languages such as M/T from other well-studied languages can be grouped into three main areas: the voice system, the WH-extraction system, and the binding system. I introduce each of these in turn.

2.1 Voice system

While the exact characterization of the voice system of languages such as Malagasy and Tagalog is quite controversial, it is clear that this system differs from the voice system of most Indo-European languages in a variety of ways. First, in many WMP languages, the voice system and the verbal morphology that marks it is very rich. In some languages, there is special morphology to promote instrumentals, benefactives, or purpose phrases to a designated syntactic position of subject/topic.¹ In this paper, I highlight two other characteristics that distinguish the voice system of both BI and M/T from the voice system of E/F/G/H/I — the form that the verbal morphology takes, and the status of the non-promoted Agent.

2.1.1 Verbal morphology

Not only is the verbal voice morphology in WMP richer than in other languages, we find a surprising patterning. While in English active verbs are unmarked and passive verbs are marked, the tendency goes in the opposite direction in both BI and Malagasy as

¹ Part of the controversy revolves around whether the designated syntactic position is a subject position or topic position, whether it is an A-position or an A' position, and whether it is nominative or absolute. I will refer to it as a nominative subject acknowledging that this labeling prejudices certain other issues. Other analyses will be discussed briefly at the end of this section.

shown below. In the active voice (i.e. when the highest argument is the subject) as in (1a) and (2a), the verb has additional morphology in both languages. When the Theme is in the subject position, we often get the root form of the verb as shown in (1b) and (2b).²

- | | | |
|--------|-------------------------------------|------------------------|
| (1) a. | Dia me-lihat perempuan itu | BAHASA INDONESIA (SVO) |
| | he <i>MEN</i> -see woman that | |
| | 'He sees that woman.' | |
| b. | Perempuan itu dia lihat. | |
| | woman that he \emptyset -see | |
| | 'That woman was seen by him.' | |
| (2) a. | Mahita ity vehivavy ity izy | MALAGASY (VOS) |
| | <i>MA</i> -see this woman this he | |
| | 'He sees this woman.' | |
| b. | Hitany ity vehivavy ity | |
| | \emptyset -see-he this woman this | |
| | 'He saw this woman.' | |

From an English perspective, it is odd to have the form that correlates with the passive (i.e. with Theme as subject) be the unmarked form.³

2.1.2 ‘Ergative’ Agent)

Another surprising characteristic of the voice system is that the non-promoted Agent appears not in a *by*-phrase as in English, but in a regular argument position (or remains a term in the parlance of Relational Grammar). We have already seen this in (1b) and (2b) above. Two other examples are given below.

² In the glossing of the examples I name the verbal morphemes and case-markers (e.g. *MEN*, *MA*) rather than give them grammatical labels.

³ This reverse in the morphological marking is one reason that many have treated these languages as being ergative. In an ergative system, the (a) examples would be the anti-passive forms and the (b) examples would be the simple transitives. See e.g. Aldridge 2004, Gerdts 1988, MacLachlan 1996.

In BI, the non-promoted Agent appears directly before the verb while the non-promoted Agent in Malagasy appears directly after the verb, attached to the verb morphologically by a process dubbed N-bonding by Keenan (2000). In neither case, however, is the Agent realized as an oblique in a prepositional phrase as in an English passive.

2.2 Similar extraction facts

WH-constructions also distinguish BI and M/T from E/F/G/I type languages. Below we will see three ways in which BI and M/T differ from languages like English. Two have to do with how extraction is encoded and one has to do with what can be extracted.

2.2.1 WH-structures as (pseudo-) clefts

WMP languages including BI and M/T often use as the main question formation strategy a cleft or pseudo-cleft sort of structure. Recently this has received attention in the literature (see Law 2005, Paul 2001, Potsdam 2004). The consensus is that the WH- construction has the form and interpretation given below.

- (5) a. [0_v XP_{WH}] [_{DP} RELPART [_{IP} ...]]

b. [is XP_{WH}] [the XP that [...]]

c. The thing that Rabe washed is what?

What looks like a fronted wh-XP is, in fact, a predicate with a zero verb form (copula). The subject of the clause is a headless relative marked by a relative particle (RELPART) of some sort. The example below shows how the strategy used for relativization in BI is also used to form a WH-question. In each case the particle *yang* is present.

- (6) a. Orang yang duduk dekat jendela BAHASA INDONESIA
person COMP sit near window
'The person who is sitting near the window'

b. Siapa yang duduk dekat jendela
who COMP sit near window
'Who is the one sitting near the window?'
'The one who is sitting near the window is who?'

In Malagasy, the formation of (pseudo-)clefts uses the particle *no* as shown in (7a). (7b) shows that the same particle is used in the formation of WH-questions.

2.2.2 Subjects extract best

Another characteristic that M/T and BI share and which distinguishes them from English is the restriction on extraction made famous in Keenan and Comrie's (1977) NP Accessibility hierarchy. Put most simply, subjects are the arguments that extract the most easily. Below we see this first in BI and then in Malagasy. (8a) is a construction with the Agent as the subject. With this verb form, the Theme cannot be relativized as (8b) shows. In order to relativize the Theme, a different construction must be used. This is shown in (9) where we now have a form of the verb which ensures that the Theme is the subject. With this verbal form, the Theme may now relativize.

- (8) a. Dia me-lihat perempuan itu BAHASA INDONESIA
he *MEN-see* woman that
'He sees that woman.'
- b. * [Perempuan [yang dia me-lihat t] itu] menangis.
woman that he *MEN-see* that *MEN-cry*
'That woman that he saw cried.'
- (9) a. Perempuan itu dia lihat.
woman that he \emptyset -see
'The woman was seen by him.'
- b. [Perempuan [yang dia lihat] itu] menangis.
woman that he \emptyset -see that *MEN-cry*
'That woman that he saw cried.'

A similar set of examples can be given for Malagasy. Here we start with a grammatical WH-construction. The Agent has been extracted successfully in (10a). With the same form of the verb, however, the Theme cannot be extracted as shown in (10b). As with BI, in order to extract the Theme, it first must be promoted to subject. This is

shown in (10c) where the change in the form of the verb indicates a change in the choice of subject.

- (10) a. Iza no manasa lamba MALAGASY
 who PRT *MAN.wash* clothes
 ‘Who is washing clothes?’
- b. * Inona no manasa Rabe
 what PRT *MAN.wash* Rabe
 ‘What is Rabe washing?’
- c. Inona no sasan-dRabe
 what PRT wash-*AN-N-Rabe*
 ‘What is being washed by Rabe?’

What these examples have shown is that subjects extract straightforwardly and the extraction of objects is restricted.

2.2.3 *Different system for adjuncts*

Another less known way in which BI and M/T pattern similarly and different from English is in the way adjuncts are extracted. What is interesting is that adjuncts differ from both subjects and objects in WH-constructions. They differ from objects in that they can extract where objects cannot. And they differ from subjects in the mechanism that is used for extraction. We saw above that WH-constructions generally use a specific particle in A' extraction constructions. In many of the WMP languages, however, adjuncts extract without using this particle. We show this in BI below using data from Cole, Hermon, and Tjung (to appear). First, in (11), we see an example of a relative clause where the subject has been relativized and, as expected, we find the particle *yang*. In

(12a) and (12b), however, where adjuncts have been relativized, instead of *yang* we find more contentful material indicating either location or time in place of the particle.

- (11) Orang yang duduk dekat jendela
person COMP sit near window
'The person who is sitting near the window'
- (12) a. Saya ingat hotel **di mana/tempat** saya pernah menginap
1SG remember hotel in which place 1SG once MEN-stay
'I remember the hotel where I once stayed.' (YT)
- b. Saya masih punya foto **waktu/ketika** saya masih kecil
1SG still have photo time when 1SG still small
'I still have the pictures when I was a little boy.' (YT)

Tagalog also distinguishes subject movement from adjunct movement. Below we see that, as in BI, subject extraction is accompanied by a particle, in this case *ang*. When an adjunct has been extracted, however, there is no particle as (14) shows (data adapted from Richards and Rackowski 2004).

- (13) Sino ang binigy-an ng lalaki ng bulaklak
who ANG give-AN NG man NG flower
'Who did the man give the flower to?'
- (14) Kailan binigy-an ng lalaki ng bulaklak ang kalabaw
when give-AN NG man NG flower ANG water-buffalo
'When did the man give a flower to the water buffalo?'

Malagasy also distinguishes between adjunct and argument (subject) extraction but in a less obvious way. Unlike BI and Tagalog, both subject and adjunct extraction require a particle, in the case of Malagasy this particle is *no*.

- (15) a. Iza no manasa lamba
 who PRT *MAN.wash* clothes
 ‘Who is washing the clothes?’
- b. Aiza no manasa lamba Rabe
 where PRT *MAN.wash* clothes Rabe
 ‘Where is Rabe washing the clothes?’
- c. Oviana no manasa lamba Rabe
 when PRT *MAN.wash* clothes Rabe
 ‘When is Rabe washing the clothes?’

In a certain construction called the Bodyguard Construction (see e.g. Paul 2000), however, adjuncts and only adjuncts may appear as WH-words to the left of a fronted subject followed by the particle *no*.

- (16) a. Aiza ianao no mipetraka
 Where you PRT *MI.live*
 ‘Where do you live?’
- b. * Iza ato no mipetraka
 who here PRT *MI.live*
 ‘Who lives here?’

The languages, then, share the characteristic of having a different strategy to extract adjuncts.

2.3 Similar binding facts

A final area where BI and M/T differ significantly from E/F/G/I is in terms of anaphor binding. Most striking is that these languages allow a non-promoted Agent to bind an anaphor in the subject position. We see this for BI in (17) with data from Arka and

Manning (2006) and in (18) for Malagasy. In each case a Theme has been promoted to the subject position yet is able to be bound by the Agent.

- (17) Diri saya saya serahkan ke polisi BAHASA INDONESIA
Self 1 1SG surrender to police
'I surrendered myself to the police.'

(18) Novonoiny ny tenany MLAGASY
PST.kill-N-3SG the body-N-3SG
'S/he killed himself/herself.'

2.5 Summary

We have now seen several ways in which BI and M/T show characteristics quite distinct from E/F/G/I. These characteristics cluster around certain parts of the syntax and these clusterings have suggested particular accounts. There are three things that are special — the predicates, the subjects, and the Agents. The **predicates** encode robust voice systems that tend to be unmarked with Theme as subject. The **subjects** behave peculiarly blocking movement of other arguments across them and allowing anaphors. **Agents** remain in an argument position when not promoted to subject. Organizing the language particular properties in this way has lead research in two particular directions. Some researchers have proposed that the subject position is in fact an A'-position (similar to a topic) (see e.g. Pearson 2001 for Malagasy and Sells 2000 for Tagalog). Others have proposed that these characteristics point to an ergative analysis of these languages (see e.g. Aldridge 2004 for Tagalog). To my knowledge, however, no one has proposed either of these analyses for BI. This is surprising on one hand given the similarities we

have just seen. On the other hand, as we are about to see, there are ways that BI is much closer to an English type language.

3 DIFFERENCES■BAHASA INDONESIA LIKE ENGLISH

As similar as BI seems to languages like Malagasy and Tagalog, there are also significant ways in which it patterns more like English. These similarities have posed problems for any account of M/T which capitalizes on the differences. Below I present five ways that BI patterns with English rather than with M/T.

3.1 Predicate initial

Perhaps the most striking English-type characteristic of BI is that, like English, it is subject initial (SVO) while Malagasy is VOS and Tagalog is verb-initial with free word order of the elements following the verb. This difference in word order is apparent in all of the examples given above and will become one of the more important differences in the typology that I will be presenting.

3.2 Contains BY-phrases

Another characteristic of BI, and one that exemplifies its intermediate status, is that it famously has two types of passive (see Chung 1976). We saw one type of passive in (1b) and (3) above. This passive is similar to the M/T passive where the Agent remains an argument. The other type of passive is exemplified in (19) below. In this sort of the passive, as in English, the Agent, if it appears at all, appears in a *by*-phrase.

Note further that this sort of passive is also more like English in that it is marked by extra morphology on the verb.

3.3 Object extraction

Above we discussed extraction in BI and M/T and I showed that subjects are the most extractable arguments. However, in BI, under certain circumstances, objects may extract. This type of extraction has been discussed in Chung (1976), Cole, Hermon, and Tjung (to appear) and in Cole, Jonczyk, and Lilly (1999) for Javanese. Data are given below from Cole and Hermon (1998). While the authors point out that the bare form of the verb in (20a) suggests that extraction is from an Theme as subject construction (i.e. passive), the placement of the adverb after the Agent shows that the Agent is in the subject position and extraction is indeed from object position. (20b) shows what the word order would be if extraction was from the subject position.⁴

- (20) a. **Object Relativization** BAHASA INDONESIA
Inilah buku_i [yang **saya** sudah Ø-baca _i]
this.LAH book COMP 1SG already read
'This is the book that I have read.' (YT)

⁴ There are other tests including the case form of the Agent that confirm this analysis. The interested reader can find more details in the articles cited.

b. **Subject Relativization**

Inilah buku_i [yang _i sudah **saya** Ø-baca]
this.LAH book COMP already 1SG read
'This is the book that Badu has read' (YT)

3.4 Derived objects□

BI and M/T also differ on the form an applicative takes. I have argued elsewhere (Travis 2001) that Malagasy has no derived objects. First we see below a case of a derived object from BI. In an "object creating rule" (Chung 1976), with additional morphology on the verb (in this case *-kan*), the object of a preposition becomes the direct object of the verb. In Relational Grammar terms, a 3 (oblique) becomes a 2 (object). *Ali*, the goal, appears as a prepositional object in (21a) but as a direct object in (21b) (data adapted from Baker, 1988; p. 234, due to Chung 1976).

- (21) a. Saya mem- bawa surat itu kepada **Ali** BAHASA INDONESIA
I MEN- bring letter the to Ali
'I brought the letter to Ali.'
- b. Saya mem- bawa- kan **Ali** surat itu
I MEN- bring- to Ali letter the
'I brought Ali the letter.'

In M/T, however, while it is common to use verbal morphology to promote an oblique to an argument position, the promotion is not to object position but to subject position. In Relational Grammar terms, then, a 3 (oblique) becomes a 1 (subject). We can see this in the Malagasy example below. In (22a), *ny ankizy* 'the children' is the object of a preposition, but with a different form of the verb, *ny ankizy* becomes the subject as in (22b).

- (22) a. Nividy lamba ho an' **ny ankizy** ny lehilahy MALAGASY
 NI-buy clothing for the children the man
 'The man bought clothing for the children.'
- b. Nividian'ny lehilahy lamba **ny ankizy**
 NI-buy-DET man clothing DET child
 'The man bought clothing for the children.'

There is no verb form for a verb root such as *vidy* 'buy' which will promote the goal to the direct object position.

3.5 DIRECT vs. INVERSE VP languages (Pearson 2000)

The final way I discuss in which BI patterns more like E/F/G/I than like M/T has to do with a distinction pointed out by Pearson (2000). He notes that not all VO languages organize their VP internal elements in the same way. He divides VO languages into two types — DIRECT languages (like English) and INVERSE languages (like Malagasy). Some of the relevant distinctions are given below.

First, in direct languages, in double object constructions (DP DP), the indirect object precedes the direct object, while in inverse languages the order is reversed. This is exemplified below with an example from Malagasy (DO IO) followed by a translation into English (IO DO).

- (23) Nanolotra ny dite ny vahiny ny zazavavy (MP: (3))
 PAST.offer DET tea DET guest DET girl
 'The girl offered **the guests** the tea.'

Further, as already noted in Rackowski (1998), adverbs in Malagasy occur in the opposite order from adverbs in English. An example is given in Malagasy below, again to be contrasted with its English translation.

- (24) Manasa tsara ny lambany foana i Ketaka (MP: (7))
 PRES.wash well DET clothes.3 always DET Ketaka
 ‘Ketaka **always** washes his clothes well.’

Using an example with double objects in BI below, we can see that BI patterns with English rather than with Malagasy suggesting that it, like English, is a direct language.

- (25) **IO DO**
 Saya mem- bawa- kan **Ali** surat itu
 I *MEN-* bring- to Ali letter the
 ‘I brought Ali the letter.’

4 INTERIM SUMMARY

The table below summarizes the characteristics of BI that we have just seen. In the case of Properties 1-6, we can see that BI patterns like M/T and in the case of Properties 7-11, BI patterns like E/F/G/I.

(26)

	1MORPH	2CLEFT	3SUBJEX	4ADJ#	5ERG	6NOMAN	7PRED1	8DERO	9DIR/INV	10OBJEX	11BYDP
<i>M/T</i>	rich	✓	✓	✓	✓	✓	✓	no	Inv	no	no
<i>BI</i>	rich	✓	✓	✓	✓	✓	no	✓	Dir	✓	✓
<i>E/F/G/I</i>	poor	no	no	no	no	no	no	✓	Dir	✓	✓

We can now easily see why BI often stands in the way of accounts that try to capture the clustering of properties that are found in M/T that distinguish these languages from languages such as English. In looking only at the top and bottom rows, there is an impressive array of differences. However, BI appears to fall somewhere between the two language types.

One especially interesting part of BI syntax is that, not only does it appear to be both M/T-like in some ways and E/F/G/I-like in others, but it also in one given characteristic seems undecided. In M/T, Themes of transitive verbs become subjects without the demotion of the Agent (Property5). In E/F/G/I, Themes of transitive verbs can only become subject if the Agent is demoted (Property11). BI can use both of these strategies. However, when it uses the English-type passive strategy, it loses its M/T-type binding property. Recall that in both M/T and BI, nominative Theme anaphors can be bound by an Agent within the VP (Property6). We saw an example of this in (17) above. In BI, however, with an English-type passive, we get English-type binding. In the example below we can see that with a *by*-phrase, we no longer can have a nominative anaphor.

- (27) ?* Dirinya di-serahkan ke polisi oleh Amir
self.3 PASS-surrender to police by Amir
'Himself was surrendered to the police by Amir.'

Now armed with these observations, let us turn to a possible typology to distinguish M/T from E/F/G/I and see if we can fit BI within this typology.

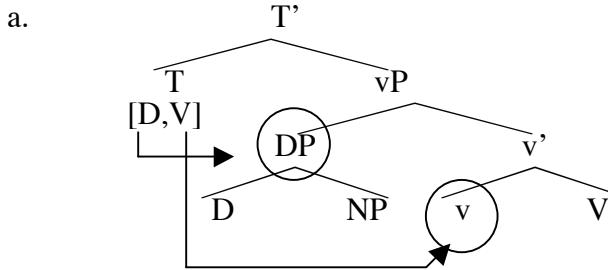
5 A TYPOLOGY

In other work (Travis 2005, in press), I have proposed a language typology which is based on the feature theory of movement as in the Minimalist Program. I claimed that languages can differ on what level category is targeted by a particular feature. Put most simply, given the features D and V, there will be two types of languages — languages that move DPs and V⁰s (like English) and languages that move D⁰s and VPs (like Malagasy). I have called with A-type and B-type languages respectively.

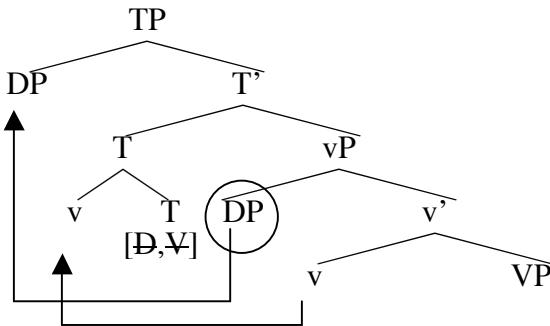
Here I give a necessarily brief introduction to this typology. Let us start by looking at English/French, an A-type language. I assume, similar to the system outlined in Chomsky (1995), that all languages have both V and D uninterpretable features that are generated as part of the inflectional domain and that need to checked before the interfaces. In English/French, the checking is followed by movement of the relevant element. We can see this in the tree below. The D and V, following Chomsky (1995), are housed in T and checked by the highest DP argument and the v respectively. Whatever feature ensures movement will see to it that the DP moves in both English and French, and that the v moves in French.⁵

⁵ We will return to the question of what happens to V-movement in English below.

(28) TYPE A \square



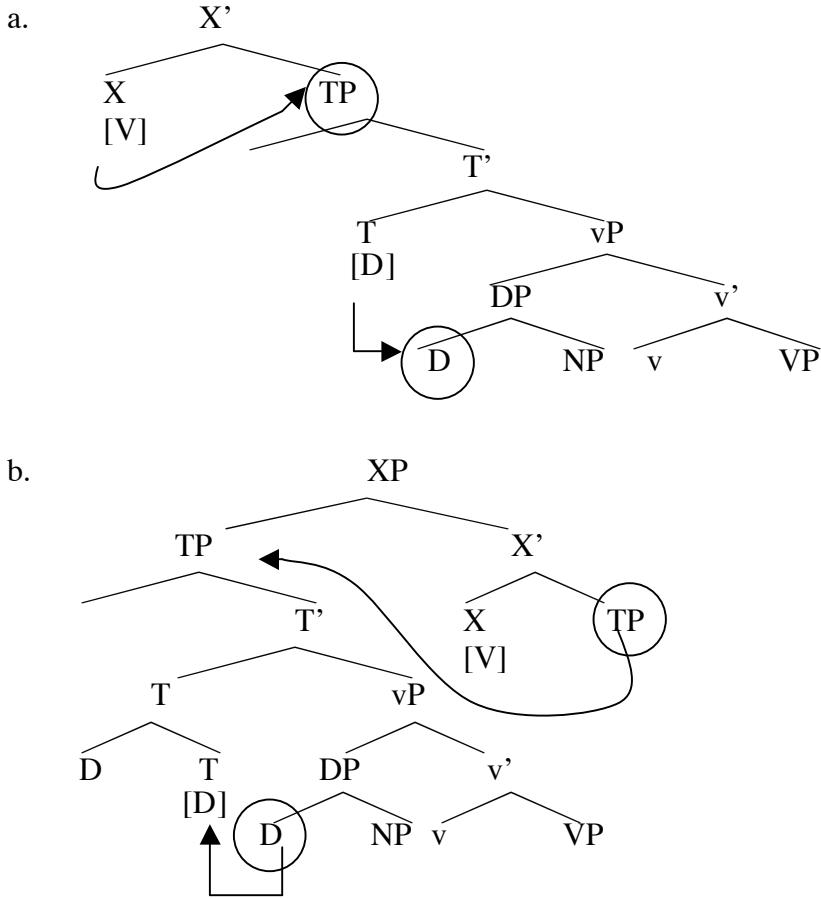
b.



As pointed out in Chomsky (2001), nothing in the computational system tells us why an XP moves to satisfy the D feature while an X moves to satisfy the V feature. He proposes that this distinction follows the fact that head movement is driven by “phonological properties, conditioned by the phonetically affixal character of the inflectional categories” (Chomsky 2001: 38). Chomsky, then, solves the problem by moving head movement out of the syntactic component. Taking a different direction, however, I suggest that languages differ on exactly this choice. Below we see a sketch of a B language where again, the inflectional domain houses the two features, but now it is the V feature that triggers XP movement and the D feature triggers X movement.⁶

⁶ There are details that are missing from this presentation of the analysis. One question is why in these languages the features are housed in different functional heads. At this point I don't have a good answer to this question. It just seems to be the fact and the reason behind it is still unclear. There is also the problem of how the subject comes to be predicate external if there is no DP movement. I return to this below. For other details, I send the reader to two other papers (Travis 2005 and in press).

(29) TYPE B \Box



The typology sketched above was designed to account for languages like Malagasy and Tagalog and to oppose them to languages like English and French. As such, it is meant to account for the types of phenomena discussed above that distinguish M/T from E/F/G/I. First, M/T are predicate initial rather than subject initial languages (Property7). This difference is clearly achieved in the structures given above. Further, WH-constructions are cleft constructions where, once again, it is not an argument that is fronted but rather is predicate (see (5) above) (Property2). We account for the DIRECT/INVERSE distinction in a similar fashion. Pearson (2000) proposes that in inverse languages VP shells are fronted within the VP explaining the inverse order of elements

within the VP. It is expected, then, that inverse languages will be B languages where a verbal feature targets XP movement (Property9).

The status of subjects clearly has to be different in B languages since their function cannot be created through DP movement. I assume, using ideas of Keenan (2000) and Pearson (2004) that subjects are base-generated in a position outside of the TP and identified by verbal morphology (similar to clitic-left dislocation (CLD) structures in Italian as presented in Cinque 1990). This explains why the verbal morphology of these languages is so rich and different from E/F/G/I voice morphology (Property1). A CLD analysis also explains the binding facts (Property6).⁷ As we see below, clitic left dislocated elements in Italian show connectivity effects meaning that they can be anaphors bound by something within the clause. This is the same behavior as subjects in M/T.

- (30) A *? lei/se stessa, Maria non ci pensa.
of her/herself, Maria not-there-thinks
'Maria doesn't think of herself.'

D-movement is meant to explain the ergative external arguments (Property5). I assume that the checking of the D feature by the D head of the highest argument has as its overt reflex the termhood of a non-subject external argument. Since the verbal morphology in M/T has a very different function from the voice morphology in E/F/G/I, we don't expect that these languages would have *by*-phrases (Property11). Since DPs do not move, we don't expect derived objects (Property8).

⁷ It is still not clear to me why the unmarked form is used to make Themes into subjects. This fact has a better explanation in accounts which treat these languages as being ergative.

This language typology is also supposed to account for the extraction facts. Since there is no XP movement and since WH-constructions are formed by clefts (predicate-initial) structures, we expect the extraction facts to be different. The proposal is that with subject extraction, the verbal morphology plus the relevant particle creates a headless relative without movement and the WH-XP is a fronted predicate. Because the formation of these constructions is dependent on the verbal morphology, we expect subjects to be the only extractable arguments (Property3 and Property10). Adjuncts, however, can be predicates that take events as their subjects. Therefore, the extraction of adjuncts can create a different construction (Property4).

Obviously, if this typology accounts for Properties 1-11, BI creates a problem. We have seen that for Properties 1-6, BI patterns like a B language. For Properties 7-11, however, BI patterns like an A language. Before turning to a possible solution, I look at the properties again in order to emphasize the variable behavior even within one part of the grammar. While BI always uses the cleft construction for WH- (Property 2) and has different adjunct extraction (Property 4) like a B language, and it is always subject first, has derived objects, and direct VPs like an A language (Properties 7, 8, 9), with respect to many of the other properties, BI seems to shift back and forth. Sometimes non-subject external arguments are ergative and sometimes a *by*-phrase (Properties 5 and 11). Sometimes (depending on the morphology of the verb) objects can extract (Properties 1, 3, and 10). Further, when the structure is clearly A language like (i.e. it has a *by*-phrase) then binding is also A language like (Property 6).

6 BAHASA INDONESIA WITHIN THE TYPOLOGY

Now let us look at possible ways that we can situate BI within this typology. I start by assuming that BI is a B language to explain the six characteristics that it shares with M/T. Now we have to explain the five characteristics that it shares with E/F/G/I. The proposals I make are quite tentative at this point and represent on-going research.

6.1 V-Initial

The clearest problem is that BI isn't predicate initial as would be predicted (shown in the trees in (29)). This is actually the easiest problem to solve. We have already seen that A type languages should have V movement into the inflectional domain, however, English is famous for not having this sort of movement (see e.g. Emonds 1978, Pollock 1989). Lack of V movement in English does not mean that English has no V feature. Rather, the assumption is that there is movement that is non-overt.⁸ We could even say that we should expect to covert movement in B languages since it is attested in A languages. In fact, in my previous work, I have had to assume that most D movement in B languages is covert. Ironically, BI has the clearest case of overt D movement. Looking again at the first example introduced above, we can see that the 'ergative' agent appears before the verb while other non-subject arguments must appear after the verb. We might think that the word order of (31b) is the result of non-movement of the verb over the Spec,vP rather than movement of the D (pronoun) to a position higher in the structure. However, the well-known generalization is that only pronouns can appear in front of the verb. If we

⁸ I do not intend to enter the debate as to how covert movement is represented.

assume (following Postal 1969) that pronouns are determiners, we can explain this restriction by viewing the word order in (31b) as derived by D-movement.⁹

- (31) a. Dia me-lihat perempuan itu
he *MEN*-see woman that
'He sees that woman.'
- b. Perempuan itu dia lihat.
woman that he \emptyset -see
'That woman was seen by him.'

My conclusions, then, would be that predicate fronting in BI is covert (except in the case of WH-constructions where the WH-predicate fronts. D-movement, however, is overt.

6.2 Object extraction

Next, let us look more closely at the fact that objects sometimes can extract. Note that this fact does not undermine the observation that subjects extract most easily. Without changing the verb form, the only argument that can extract is the subject. If the object extracts, it is crucial that the verb form has no overt morphology. As a first step in explaining this apparent anomaly I suggest that this lack of morphology loosens the restriction on extraction in BI and as well as in M/T. To see that this is also true of Tagalog, we look at the examples below from Cena (1979). There are certain counterexamples to the subjects-only extraction restriction in Tagalog, but choice of

⁹ In my other papers, I have assumed that Celtic languages also show evidence for D-movement as in Baker and Hale (1990). There are still problems, however, since it seems that Ds never move leaving their complement NPs behind while Vs do move leaving complements behind.

predicate is key. In the example below we have the predicate *katulong* ‘help’ which has no voice morphology. In this case, we can extract the object.¹⁰

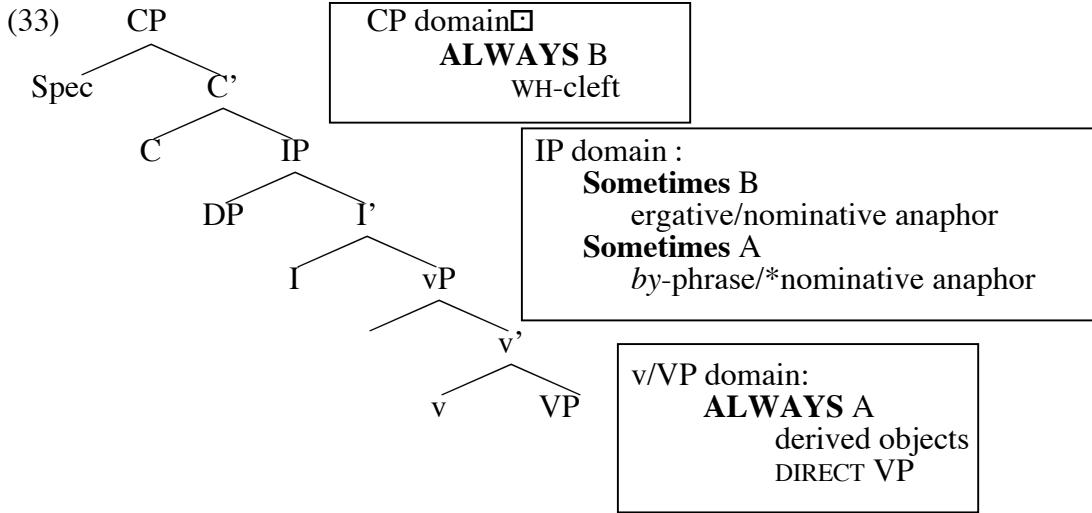
- (32) a. Katulong ng doktor **ang anak** TAGALOG
 help GEN doctor T child
 ‘The child is helping the doctor.’
- b. Madismaya ang doktor na katulong **ang anak**
 disappoint T doctor LINK help T child
 ‘The doctor whom the child is helping was disappointed.’

While this observation still requires an account, at least BI is looking more similar to M/T. In cases where a verb can appear in a form with no voice morphology, the subjects-only restriction on extraction is lifted.

6.3 Typology and domains

Now we are left with three more characteristics to account for: (i) presence of derived objects, (ii) direct order of VP internal constituents, and (iii) the presence of *by*-phrases. To begin to attack these issues, I look again at the characteristics and where they appear on the phrase structure. We can see below that the characteristics that are always B like are high in the structure while the characteristics that are A like are low in the structure. If we divide the structure into three domains — the CP-, the IP- and the VP-domains — the top domain (CP) is **always** B, the bottom domain (VP) is **always** A, and the middle domain (IP) appears to vary.

¹⁰ Along the same lines, in the recent perfective form of the verb, again *ka-* is added to the root, there is no voice morphology, and any argument can be extracted.



While much more work needs to be done on developing this account, the preliminary results can be interpreted as indicating that the typology sketched above may vary even within a language. It appears that BI is an A language within the VP, variable within the IP domain, and a B language within the CP domain.

5. CONCLUSION

This paper ends on a very speculative note and with an outline of a research direction. The hypothesis is that not only do languages divide along the lines created in the typology introduced above, but that specific language domains can also divide along these lines. This has the effect that parameters must be fine-tuned enough to allow languages to choose different values for a parameter depending on the domain. This becomes an empirical issue requiring more careful data work. The overall conclusion, however, is that any work on the typology of languages and of Austronesian languages in particular needs to pay close attention to the dual behavior of Bahasa Indonesia. This

behavior which seems to place it between two language types should offer important information on the specifics of language typology.

REFERENCES

- Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages, Cornell University: Doctoral dissertation.
- Arka, Wayan, and Christopher D. Manning. 2006. Voice and grammatical relations in Indonesian: A new perspective. In *Voice and Grammatical Functions in Austronesian Languages*, eds. Peter K. Austin and Simon Musgrave: Center for the Study of Language and Information.
- Baker, Mark, and Kenneth Hale. 1990. Relativized Minimality and Pronoun Incorporation. *Linguistic Inquiry* 21:289-297.
- Cena, Resty. 1979. Tagalog Counterexamples to the Accessibility Hierarchy. *Studies in Philippine Linguistics* 31:119-124.
- Chomsky, Noam. 1995. *The Minimalist Program*. vol. 28: Current Studies in Linguistics. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: a life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- Chung, Sandra. 1976. An Object Creating Rule in Bahasa Indonesia. *Linguistic Inquiry* 13:39-77.
- Chung, Sandra. 1976. On the Subject of Two Passives in Indonesian. In *Subject and Topic*, ed. Charles Li, 57-98. New York: Academic Press.
- Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. vol. 17: Linguistic Inquiry Monograph. Cambridge, MA: The MIT Press.
- Cole, Peter, and Gabriella Hermon. 1998. Subject and Non-Subject Relativization in Indonesian. Paper presented at Second Symposium on Malay/Indonesian Linguistics, Ujung Pandang, Indonesia.
- Cole, Peter, Elizabeth Jonczyk, and Jason Lilley. 1999. A note on extraction from object in Javanese and other Javanic languages. In *Proceedings of the sixth meeting of the Austronesian Formal Linguistics Association (AFLAVI)*, eds. Carolyn Smallwood and Catherine Kitto.
- Cole, Peter, Gabriella Hermon, and Yassir Tjung. to appear. The Formation of Relative Clauses in Jakarta Indonesian: Data From Adults and Children. In *Selected Studies on Indonesian/Malay Linguistics.*, eds. A. van Engelenhoven and H. Steinhauer. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Emonds, Joe. 1978. The verbal complex V'-V in French. *Linguistic Inquiry* 9:151-175.
- Gerdts, Donna. 1988. Antipassives and Causatives in Ilokano: Evidence for an Ergative Analysis. In *Studies in Austronesian Linguistics*, ed. Richard McGinn, 323-345. Athens, Ohio: Ohio University Center for International Studies.

- Keenan, E.L., and B. Comrie. 1977. Noun Phrase Accessibility. *Linguistic Inquiry* 8:63-99.
- Keenan, Edward L. 2000. Morphology is Structure: a Malagasy Test Case. In *Formal Issues in Austronesian Linguistics*, eds. Ileana Paul, Vivianne Phillips and Lisa Travis, 27-47. Dordrecht, the Netherlands: Kluwer Academic Press.
- Law, Paul. 2005. Questions and clefts in Malagasy. Paper presented at Austronesian Formal Linguistics Association (AFLA XII), UCLA.
- MacLachlan, Anna. 1996. Aspects of Ergativity in Tagalog, McGill: Doctoral dissertation.
- Paul, Ileana. 2000. Malagasy Clause Structure, McGill University: Doctoral dissertation.
- Paul, Ileana. 2001. Concealed pseudo-clefts. *Lingua* 111:707-727.
- Pearson, Matt. 2004. Voice morphology, Case and argument structure in Malagasy. paper presented at AFLA 11, ZAS, Berlin.
- Pearson, Matthew. 2000. Two types of VO languages. In *The Derivation of VO and OV*, ed. Peter Svenonius, 327-363. Amsterdam: John Benjamins Company.
- Pearson, Matthew. 2001. The Clause Structure of Malagasy: A Minimalist Approach, University of California, Los Angeles: Doctoral dissertation.
- Pollock, Jean-Yves. 1989. Verb Movement, UG, and the Structure of IP. *Linguistic Inquiry* 20:365-424.
- Postal, Paul M. 1969. On So-called Pronouns in English. In *Modern Studies in English*, eds. David Reibet and Sanford Schane, 201-224. Englewood Cliffs, NJ: Prentice-Hall.
- Potsdam, Eric. 2004. Wh-questions in Malagasy. In *Austronesian Formal Linguistics Association (AFLA XI)*, ed. Paul Law, 244-258.
- Rackowski, Andrea. 1998. Malagasy Adverbs. In *The Structure of Malagasy, Volume II*, ed. Ileana Paul, 11-33. Los Angeles: UCLA Occasional Papers in Linguistics.
- Rackowski, Andrea, and Norvin Richards. 2004. Phase edge and extraction: a Tagalog case study. unpublished ms., MIT.
- Sells, Peter. 2000. Raising and the Order of Clausal Constituents in the Philippine Languages. In *Formal Issues in Austronesian Linguistics*, eds. Ileana Paul, Vivianne Phillips and Lisa Travis, 117-143. Dordrecht, the Netherlands: Kluwer Academic Press.
- Travis, Lisa deMena. 2001. Derived Objects in Malagasy. In *Objects and other Subjects: Grammatical Functions, Functional Categories and Configurationality*, eds. William D. Davies and Stanley Dubinsky, 123-155. Dordrecht: Kluwer Academic Publishers.
- Travis, Lisa deMena. 2005. VP, D⁰ movement languages. In *Negation, Tense and Clausal Architecture: Cross-linguistic Investigations*, eds. Raffaella Zanuttini, Héctor Campos, Elena Herburger and Paul Portner. Washington, DC: Georgetown University Press.
- Travis, Lisa deMena. in press. Through the looking glass: Malagasy in Wonderland. In *Clause structure and adjuncts in Austronesian languages*, eds. Hans-Martin Gärtner, Paul Law and Joachim Sabel. Berlin: Mouton de Gruyter.