ON THE LEFT PERIPHERY IN INDONESIAN*

Catherine Fortin
Carleton College
cfortin@carleton.edu

In this paper, I argue that the complex array of question formation processes seen in Indonesian most readily receives a unified analysis when questions are viewed as a type of focus construction (following e.g. Kader 1981 for Malay; Horvath 1986 for Hungarian; Rizzi 1997 for Italian; and Bošković 2002 for Slavic multiple-\textit{wh}-fronting languages). I identify a set of five Foc(us) heads in Indonesian, each with a unique constellation of properties, that together yield all types of constituent questions, \textit{yes-no} questions, and non-interrogative focus constructions attested in Indonesian. I also claim that a number of asymmetries observed in Indonesian questions can be accounted for with this type of analysis, and discuss many of the directions for future research needed to fully develop this analysis. Lastly, I argue that in the overall typology, Indonesian is a ‘Unique Focus’ (Stoyanova 2008) language, albeit one with certain unexpected properties.

1. Overview

My goal in this brief paper is to outline an empirically adequate, theoretically sound, unified analysis of both \textit{wh}-questions and \textit{yes-no} questions in Indonesian. Indonesian has a robust array of question formation processes, and an equally robust array of asymmetries in these processes, both described in Section 2, that must be accounted for in a satisfactory analysis. The main idea defended here, proposed in Section 3, is that Indonesian \textit{wh}-questions and \textit{yes-no} questions are formed with the same mechanisms, and together represent a special type of focus construction. In support of this idea, I describe the many empirical properties that questions share with non-interrogative focus constructions in Indonesian. I emphasize that this paper represents the outline of the analysis, and do not purport to advance here a complete analysis, nor one that necessarily extends to Malay, which displays some significantly different properties. Consequently, I devote a substantial portion of the paper, Sections 4 and 5, to identifying directions for future research.

2. How to Ask Questions in Indonesian

2.1. \textit{wh}-questions

In Indonesian, generally speaking, both left-peripheral \textit{wh}-phrases (1a) and \textit{wh}-in-situ (1b) are possible.

---

* I thank my native speaker consultants, Kathy Triyana, Jingga Morry, Jingga Inlora, and Nancy Surachman. For useful comments and discussion on the material presented here, I thank participants of AFLA 16, including Mark Donohue, Darlene Hilton Bird, Ed Keenan, Paul Kroeger, Eri Kurniawati, Paul Law, Tim McKinnon, Hiroki Nomoto, Yuko Otsuke, Ileana Paul, Norvin Richards, Yosuke Sato, and Lisa Travis.
(1) a. **Apa** yang akan Ali beli?  b. Ali akan membeli **apa**?
   what COMP FUT Ali buy Ali FUT meng-buy what
   ‘What will Ali buy?’  ‘Ali will buy what?’

There is no matrix/embedded asymmetry with respect to the possibility of *wh*-in-situ. There are, however, several other significant asymmetries.

First, left-peripheral *wh*-phrases can be (but need not be; this point will be discussed further below) marked with –kah (2a). Researchers (e.g. Cheng 1997) frequently refer to –kah as a question marker, and I follow this convention in the glosses. –kah cannot appear on in-situ *wh*-phrases (2b). The limited distribution of –kah can be interpreted as evidence that the morpheme is uniquely affiliated with a left-peripheral position, and not with the *wh*-phrase itself, although it does encliticize to the *wh*-phrase.¹ More empirical support will be advanced for this claim below.

(2) a. **Apa**(+)kah yang Ali beli?  b. Ali membeli **apa**(*)kah*)?
   what(++)QUES COMP Ali buy Ali meng-buy what(++)QUES

Second, left-peripheral *wh*-DPs – even subjects – must be set off from the remainder of the clause by **yang** (3a), which left-peripheral *wh*-adjuncts resist (3b). For the moment, I suggest that **yang** is simply a complementizer that is in complementary distribution with a phonologically null complementizer; this point will also be returned to below.

(3) a. **Siapa** *(yang)* membeli mobil baru?
   who *(COMP)* meng-buy car new
   ‘Who bought a new car?’

b. **Kapan/Dari siapa** *(yang)* Ali membeli mobil baru?
   when/from who *(COMP)* Ali meng-buy car new
   ‘When/From who did Ali buy a new car?’

Third, only left-peripheral *wh*-phrases are sensitive to island effects (4a); in situ *wh*-phrases easily appear within islands of all types, such as adjunct clauses (4b).

(4) a. **Apa** yang Ali jadi terlalu gemuk  
   what COMP Ali be too chubby  
   [CP karena dia makan]?
   ‘What did Ali get fat because he ate?’

¹ –kah cannot float free at the left edge of the clause:

(i) **Kah** Ali membeli **apa**?
   QUES Ali meng-buy what
   ‘What did Ali buy?’
b. Ali menjadi terlalu gemuk [CP karena dia makan apa]?  
   Ali meng-be too chubby [CP because 3SG eat what]  
   ‘Ali got fat because he ate what?’

The fourth, and perhaps the most widely-discussed (Cole & Hermon 1998, inter alia), asymmetry involves the verbal prefix meng-, which marks transitive verbs in active voice clauses. This prefix cannot appear on a verb whose complement or subcomplement appears left-peripherally (5b), although it does not prevent a wh-complement or wh-subcomplement from appearing in situ (5a). Furthermore, meng- does not block a left-peripheral wh-adjunct (6).

(5) a. Ali membeli apa?  
   ‘Ali bought what?’  

   b. Apa yang Ali (meng-)beli?  
   ‘What did Ali buy?’

(6) Kapan/Dari siapa Ali membeli mobil baru?  
   ‘When/From who did Ali buy a new car?’

To account for the asymmetries illustrated in (4)-(6), Cole & Hermon (1998), in their analysis of Singaporean Malay, propose that in-situ wh-phrases are unselectively bound by an operator Merged directly into [Spec, CP], and do not undergo covert movement (cf. Huang 1982). In my view, the simplest explanation for these asymmetries invokes an asymmetry in movement, as the meng- related asymmetry illustrated in (5)-(6) can be construed as a type of island: namely, meng- renders some projection of the verb bearing it an island. Thus, I adopt the Cole & Hermon-style non-movement analysis for Indonesian wh-in-situ questions, and for questions containing left-peripheral wh-phrases, I assume a standard wh-movement analysis (following e.g. Kader 1976 for Malay).

The big picture, as far as wh-questions go, is this. Left-peripheral wh-phrases have fronted, and thus are constrained by islands; in situ wh-phrases, which are unselectively bound and undergo no movement, are not. The question marker (–kah), which cannot appear on in-situ wh-phrases, is associated with a left-peripheral position. Fronted wh-DPs require a special complementizer, yang, while all other fronted wh-phrases require a phonologically null complementizer.

2.2. yes-no questions

Indonesian also displays several strategies for forming yes-no questions. Two such strategies are pragmatically neutral; a third strategy involves focusing of a constituent. The two types of pragmatically neutral yes-no questions, under which no changes in word order obtain, are shown

---

2 Analyses of meng-’s blocking effect on movement are numerous; see, e.g., the works cited above, and references therein. However, the precise cause of meng-’s blocking effect remains controversial, and a stand won’t be taken here.
in (7). The first strategy involves question intonation only (7a); under the second strategy, a question particle, \textit{apa(kah)}, appears left-peripherally (7b).

(7) a. Siti sudah pulang? b. Apa(kah) Siti sudah pulang?
Siti already go.home QUES Siti already go.home
‘Did Siti go home?’ ‘Did Siti go home?’

Under the third strategy, a constituent is fronted and marked with –kah, yielding a non-pragmatically neutral interpretation (8). In (8a), for example, a vP is focused; in (8b), the subject is focused.

(8) a. [Sudah pulang]+kah Siti? b. Siti+kah yang sudah pulang?
already home +QUES Siti Siti+QUES COMP already go.home
‘Did Siti ALREADY GO HOME?’ ‘Did SITI already go home?’

Overall, then, a common characteristic of Indonesian questions is the optional fronting of a constituent. We’ve seen that both \textit{wh}-phrases (in constituent questions) and non-\textit{wh}-phrases (in yes-no questions) are able to, but are not required to, overtly front. Although it is not demonstrated here for lack of space, fronted non-\textit{wh}-phrases are subject to the same island constraints (including the blocking effect of \textit{meng}-) that \textit{wh}-phrases are. Both types of fronted phrases can be (and, in the case of non-\textit{wh}-constituents, must be) marked with –kah.

The availability of such a wide number of possibilities for forming questions in Indonesian means that developing an empirically adequate, unified analysis will be difficult, although such an analysis arguably is desirable. In the next section, I briefly review one previous unified analysis of Indonesian questions, that of Cheng 1997, which I argue to be unsuccessful. I then propose a different analysis which capitalizes on the common characteristic identified above: the optional fronting and –kah-marking of a constituent.

3. A Unified Analysis of Indonesian Questions


Cheng (1997) seeks to propose an analysis of \textit{wh}-questions which accounts for their differing properties crosslinguistically. Primarily concerned with understanding why an individual language will choose to exploit a \textit{wh}-in-situ or \textit{wh}-movement strategy, she proposes the Clausal Typing Hypothesis (CTH) (9), which states that all questions need to be ‘typed’ as such at S-Structure. Typing is accomplished with \textit{wh}-movement to, or a \textit{wh}-particle in, the left periphery.

(9) \textit{Clausal Typing Hypothesis} (Cheng 1997: 22, ex. 9)
Clauses need to be [overtly] ‘typed’… either a \textit{wh}-particle in C is used or … fronting of a \textit{wh}-word is used.

The CTH predicts that overt \textit{wh}-movement and question particles are in complementary distribution, a prediction which appears to be falsified by the Indonesian facts. To maintain this
prediction, Cheng argues that left-peripheral wh-phrases in Indonesian do not undergo wh-movement. Left-peripheral wh-adjuncts (6) are base-generated topics, while left-peripheral wh-DPs (5b) are base-generated pivots of clefts, an analysis which has the further benefit of capturing the apparent structural similarities between these questions and relative clauses (10).

(10) Buku yang Ali beli itu mahal sekali.
    book COMP Ali buy DEM expensive very
    ‘The book that Ali bought is very expensive.’

Cheng’s analysis, particularly of those questions containing wh-adjuncts, faces non-trivial difficulties. On the empirical side, we’ve seen that certain wh-adjuncts – namely, PPs – are permitted to appear in situ; these are undergenerated by Cheng’s analysis. On the ‘conceptual’ side, the assumption that wh-adjuncts are ‘topics’ appears to ignore a serious semantic incompatibility (see e.g. Molnár 2006: 206 for discussion): ‘topics’ are standardly taken to encode only old or given information, which wh-phrases are not. Fortunately, it further appears that the larger theoretical motivation for the Clausal Typing Hypothesis does not hold up: Bruening (2007) cites typological data on about 500 languages, showing there to be no correlation cross-linguistically between wh-in-situ and question particles, meaning that nothing a priori prevents wh-movement and wh-in-situ strategies from co-existing within a language.³

3.2. Proposal: questions are focus constructions

The unified analysis of Indonesian questions that I seek to advance in this paper relies not on ‘topics’, as Cheng’s does, but on a different information-structural construct: focus. I adopt Lambrecht’s (1994: 207) definition of focus as new information, ‘the portion of the proposition which cannot be taken for granted… the unpredictable or pragmatically non-recoverable element of an utterance’, and I propose that questions in Indonesian are best analyzed as focus constructions. This proposal follows a long tradition of viewing questions as a special type of focus construction, including Horvath 1986 for Hungarian, Rizzi 1997 for Italian, Bošković 2002 for Slavic multiple-wh-fronting languages, and Stoyanova 2008 for Berber, Somali, and Irish. I begin with the hypothesis that the –kah observed in yes-no questions and the – kah observed in wh-questions are the same –kah, which is a focus marker. I also hypothesize that the fronting seen in both types of questions is focus fronting.

Certain aspects of wh-movement in Indonesian/Malay have previously been analyzed as involving focus. Saddy (1991) assumes that fronted Indonesian wh-arguments, although not wh-adjuncts, are focused, although the reason for this division is unclear. Kader (1981) characterizes Malay –kah and –lah as ‘focus morphemes’, which, by marking a constituent, enable that constituent to participate in a process of ‘Focus Fronting’.⁴ Similarly, Cole, et al. (to appear: 17) describe Malay –kah as ‘the interrogative focus particle’, and –lah as ‘the affirmative focus

³ I am grateful to an anonymous AFLA reviewer for directing me to this work.
⁴ Kader notes two further ‘focus morphemes’ in Malay: –tah, which appears in rhetorical questions, and pun, which has a range of meanings including ‘also, too, even’. To the best of my knowledge, Indonesian does not display –tah at all. I do not discuss pun here, which in Indonesian (as in Malay) behaves quite differently from –kah and –lah.
particle’, but do not delve more deeply into the ramifications of this classification for question formation in Malay.

3.2.1. The framework

My proposal is couched within the Minimalist Program (Chomsky 1995), including the Probe-Goal theory of agreement (Chomsky 2001). Under this theory of agreement, feature-checking obtains long-distance; a head with an unchecked feature (the probe) can agree with a goal bearing a matching feature in its c-command domain. Movement to a specifier position, such as *wh*-movement, is thus not taken to be driven by the need to check (for example) a *wh*-feature, since the *wh*-feature is checked at a distance. Movement to a specifier is thus exclusively by the need to check driven by an EPP feature on a head.

I further assume the particulars of Rizzi’s (1997) Split CP Hypothesis, under which the left periphery of the clause (the CP) is split into a number of projections concerned with information structure and illocutionary force, as illustrated in (11).

(11) \[ \text{ForceP} > \text{TopP}^* > \text{FocP} > \text{TopP}^* > \text{FinP} > \text{IP} \]

In Rizzi’s framework, which was influenced most heavily by Italian, English, and other Indo-European languages, TopP – the functional projection related to topics, or old/given information – is able to iterate, but FocP – the functional projection related to focused constituents, or new information – is not. Multiple topics are permissible in an Italian clause (12a), but only one focus phrase is allowed (12b).

(12)a. *A GIANNI IL LIBRO daro (non a Pierre, l’articolo) 
    TO GIANNI THE BOOK give.FUT.1SG 
    ‘To Gianni the book I’ll give (not to Pierre, the article)’ 
    (Rizzi 1997: 290, ex. 22)

b. Il libro, a Gianni, domani, glielo darò senz’altro (Italian) 
    the book, to Gianni, tomorrow, to.3SG give.FUT.1SG for.sure 
    ‘The book, to Gianni, tomorrow, I’ll give it to him for sure’ 
    (Rizzi 1997: 290, ex. 21)

Rizzi (1997) further argues that [Spec, FocP] is the landing site for *wh*-movement in Italian. He notes that while a *wh*-phrase can co-occur with a topic (13a), a *wh*-phrase and a focused phrase cannot co-occur (13b). Moreover, only one *wh*-phrase is permitted; there is no possibility for a second, in situ, *wh*-phrase (14).

(13)a. A Gianni, che cosa gli hai detto? 
    to Gianni, what to.3SG have.2SG tell.PAST 
    ‘To Gianni, what did you tell him?’ 
    (Rizzi 1997: 291, ex. 24a)
b. *A GIANNI che cosa hai detto (non a Piero)?
   to GIANNI, what have.2SG tell.PAST
   ‘TO GIANNI what did you tell (not to Piero)?’ (Rizzi 1997: 291, ex. 25a)

(14) *Che cosa hai dato a chi?
   what have.2SG give.PAST to who.ACC
   ‘What was given to whom?’ (Stoyanova 2008: 3)

3.2.2. Extending the framework cross-linguistically

Since Rizzi’s original proposal, it has been recognized that if wh-movement is focus movement, whether or not FocP can iterate must be parametrized. As shown above, Italian has ‘unique focus’ (Stoyanova 2008): it permits a maximum of one focused phrase, whether or not it is wh-, per clause. On the other hand, Serbo-Croatian, Bulgarian, and Russian permit any number (in principle) of fronted wh-phrases (15). If wh-movement is movement to [Spec, FocP], this implies that FocP is able in these languages to iterate without limit, in order to host the multiple fronted wh-phrases (Bošković 2002).

(15) Koj kogo kak e tselunal? (Bulgarian)
   who.NOM who.ACC how is kiss.PAST
   ‘Who kissed whom how?’ (Bošković 2002: 366)

   English appears to be somewhere in between Italian and the Slavic multiple wh-fronting languages. Like Italian, only one wh-phrase overtly fronts, but unlike Italian, the number of wh-phrases per clause is not limited to one (16).

(16) a. Who gave what to whom when?
   b. * Who what to whom when gave?

   In Indonesian, like Italian, no more than one wh-phrase - even in situ - is permitted (17).

(17) a. *Siapa yang membeli apa?
   who COMP meng-buy what
   ‘Who bought what?’

   b. * Kamu membeli apa dari toko yang mana?
   2SG meng-buy what from store COMP which
   ‘What did you buy from which store?’

---

5 Alternatively, FocP in the Slavic languages has multiple specifiers, or otherwise permits attraction of multiple phrases to a single head.
Indonesian thus far appears to be another example of a language with unique Focus, in which iteration of FocP (or multiple adjunction to [Spec, FocP], or multiple specifiers of FocP) is not possible.

3.3. Indonesian (non-\textit{wh}) Focus

In order to bolster my contentions that (i) Indonesian has unique Focus and (ii) Indonesian question formation involves focus movement, I now describe less controversial instantiations of focus in Indonesian, and demonstrate that question formation behaves similarly. Indonesian has a non-interrogative focus marker, –lah, which has a distribution and behavior similar to that of –kah, which I’ve suggested is also a focus marker.\footnote{Kader (1981: 303) characterizes Malay –lah as ‘the focus morpheme in declarative and imperative sentences’, in contrast to –kah, which appears in interrogatives.} First, like –kah, –lah is associated with a left-peripheral position, and encliticizes to a constituent which is treated as new information (18a). Also like –kah, –lah cannot appear on constituents in situ (18b).

(18) a. Kemarin{+lah} Ali membeli mobil baru.
   yesterday{+FOC} Ali buy car new
   ‘It was YESTERDAY that Ali bought a new car (not last Tuesday).’

   [door DEM]{+FOC COM} Ali (*meng-)open
   ‘It is this door that Ali opened.’

   b. [Pintu yang mana]{+kah} *(yang) Ali (*mem-)buka?
   [door COM which]{+QUES COM} Ali (*meng-)open
   ‘Which door did Ali open?’

   (20) a. [Kemarin]{+lah} (*yang) Ali mem-beli mobil baru.
   yesterday{+FOC COM} Ali buy car new
   ‘It was yesterday Ali bought a new car.’
    ‘When did Ali buy a new car?’

Next, as is the case with multiple wh-phrases, multiple focused phrases cannot occur (21). Similarly, a focused phrase cannot co-occur with a wh-phrase, whether or not that wh-phrase is marked with –kah (22a) or even fronted (22b).

    ‘It was YESTERDAY that Ali opened THIS DOOR (not that window last Tuesday).’

(22) a.*Kemarin+lah siapa+kah yang menelpon?
    ‘It was YESTERDAY (not last Tuesday) that who phoned?’

b.*Kemarin+lah kamu menelpon siapa?
    ‘It was YESTERDAY (not last Tuesday) that you phoned who?’

This pattern of data indicates that –kah and –lah are in complementary distribution, although the clash does not appear to be semantic; instead, it appears that the two markers are competing for the same structural position, which I take to be, in Rizzi’s framework, Foc. It further indicates that wh-phrases and non-wh-focused phrases in complementary distribution; since at most one (of either type) can appear in a clause, it again appears that the two types of phrases are competing for the same structural position, which I take to be [Spec, FocP]. I conclude that Indonesian is, indeed, like Italian, a language with unique Focus.

3.3.1. Indonesian focus: the fuller picture

I claimed above that –lah and –kah are both instances of Foc. Clearly, they differ in some respects: as has already been made evident, –lah lacks –kah’s interrogative interpretation. Still, both –kah and –lah can mark both fronted wh-phrases and non-wh-phrases, and the four possible combinations yield four distinct focus constructions. The first three are familiar; the fourth, which combines the non-interrogative focus marker with a wh-phrase, is a rhetorical question.

<table>
<thead>
<tr>
<th>combination</th>
<th>construction</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>–kah + wh-phrase</td>
<td>constituent question</td>
<td>(2a)</td>
</tr>
<tr>
<td>–kah + non-wh-phrase</td>
<td>yes-no question</td>
<td>(8a)</td>
</tr>
<tr>
<td>–lah + non-wh-phrase</td>
<td>non-interrogative focus construction</td>
<td>(18a)</td>
</tr>
<tr>
<td>–lah + wh-phrase</td>
<td>rhetorical question (Cheng 1997)</td>
<td>(23)</td>
</tr>
</tbody>
</table>

Table 1: Recipes for Indonesian Focus
3.4. Indonesian Questions as Focus Constructions

I argue, then, that the Indonesian lexicon contains at least two types of Foc: -kah, which is an interrogative Foc, and –lah, which is a non-interrogative Foc.\(^8\) As we’ve seen, both of these markers appear only in conjunction with fronted phrases. Focused phrases must be marked with one of the two focus markers if the phrase itself is not a wh-phrase, although fronted wh-phrases can appear without –kah. Within the Minimalist Program, the ‘optional’ application of an operation is hypothesized to be impossible; as such, the ‘optional’ wh-movement seen in Indonesian, poses a problem. Given the probe-goal theory of agreement, however, and the concomitant hypotheses that all movement to a specifier is driven by an EPP feature on that specifier’s head and all variation is rooted in the lexicon (Chomsky 2001), this apparent ‘optionality’ can straightforwardly be accounted for. To begin, both –kah and –lah have a strong Foc feature which is checked, long-distance, against some constituent in its c-command domain bearing a matching feature. Both –kah and –lah also contain an EPP feature which drives movement of the focused constituent to [Spec, FocP].

The Indonesian lexicon further contains two phonologically null interrogative Foc, analogous to -kah. Both null Foc contain, like –kah/–lah, a focus feature that is checked long distance against a constituent bearing a matching feature in its c-command domain. One of these null Foc differs from –kah/–lah in lacking an EPP feature, with the consequence that the focused constituent remains in situ; the second null Foc does bear an EPP feature, which drives movement of the focused constituent to its specifier.\(^9\) It is unclear at this point (at least to me) whether the non-interrogative Foc, -lah, also has a phonologically null analog lacking an EPP feature, which would generate non-interrogative focus constructions in which the focused constituent remains in situ (e.g. *I saw SAM (not Pat) at the party*). Given that non-interrogative focus constructions in which the focused constituent remains in situ should display a distinctive intonation contour, more research in this area will necessarily involve close attention to prosody.

\(^7\) Although Cheng (1997) describes these types of questions in Indonesian, it is not yet clear to me how widely accepted they are, as my own consultants reported that they found them rather odd. More data needs to be gathered.

\(^8\) Or, more precisely, the Indonesian lexicon contains at least two types of Foc. Whether or not pun ‘even, also, too’ in Indonesian is also an instance of Foc, as is claimed by Kader (1981) for Malay, requires additional research.

\(^9\) Under a perhaps more explanatory hypothesis, the optional appearance of –kah in questions containing a fronted wh-phrase is accounted for by the interrogative Foc being optionally lexicalized only if a wh-phrase is fronted. Given that wh-phrases are inherently focused, this hypothesis is plausible, although the details remain to be worked out. This alternative hypothesis is desirable for a second reason: the lexically-based analysis proposed in the text is in danger of overgenerating questions in which a non-wh-phrase, in combination with the phonologically null interrogative Foc with the EPP feature, fronts. Such a combination does not appear to be attested in Indonesian (i) (cf. 9a), although, according to Kader (1981), this combination is possible in Malay.

(i) *[Sudah pulang] Siti?
    already go.home Siti
    ‘Did Siti already go home?’

(23) Apa+lah yang mampu aku beri?
what+FOC COMP can ISG give
‘Whatever can I afford to give?’ (Cheng 1997)
It does appear to be the case, however, that Indonesian lacks a non-interrogative Foc with an EPP feature; non-wh-constituents that are focus-fronted are obligatorily marked with –lah.

I’ve thus far accounted for non-interrogative focus constructions, wh-questions (both wh-in-situ and wh-movement varieties), and the two of the three yes-no question formation strategies described above. The third strategy, which involves focus fronting and –kah marking of a constituent (8), resulting in a pragmatically marked question, of course makes use of the same –kah seen in wh-questions. Similarly, the first strategy, which involves a question intonation only (7a), makes use of the phonologically null interrogative Foc which lacks an EPP; this is the same Foc seen in wh-in-situ questions. As noted earlier, these yes-no questions are pragmatically neutral, which means that no sub-constituent is focused; instead, the IP itself presumably bears the focus feature, with the result that the entire clause, in a sense, is focused. Only the second yes-no question strategy (7b) remains to be accommodated, and this is easily done: the yes-no particles, apa/apakah, are also interrogative Foc which bear no EPP feature. In these pragmatically neutral questions, too, Foc checks its focus feature against the entire IP.

To conclude this section, I’ve proposed that Indonesian’s wide variety of question formation processes follows from the wide array of Foc in its lexicon; these are listed in Table 2. I’ve also noted that there appear to be gaps in the paradigm, and that certain logically possible Foc – for example, a phonologically null non-interrogative Foc with an EPP feature – do not appear to be attested in Indonesian.

### Table 2: The Five Flavors of Indonesian Foc

<table>
<thead>
<tr>
<th>head</th>
<th>semantics</th>
<th>EPP?</th>
<th>construction(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-kah</td>
<td>interrogative</td>
<td>+EPP</td>
<td>wh-questions with a fronted wh-phrase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>yes-no questions with a focused non-wh-phrase</td>
</tr>
<tr>
<td>-lah</td>
<td>non-interrogative</td>
<td>+EPP</td>
<td>focus constructions with a fronted non-wh-phrase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rhetorical questions with a fronted wh-phrase</td>
</tr>
<tr>
<td>apa/apakah</td>
<td>interrogative</td>
<td>-EPP</td>
<td>pragmatically neutral yes-no questions</td>
</tr>
<tr>
<td>Ø₁</td>
<td>interrogative</td>
<td>+EPP</td>
<td>wh-questions with a fronted wh-phrase</td>
</tr>
<tr>
<td>Ø₂</td>
<td>interrogative</td>
<td>-EPP</td>
<td>wh-questions with an in-situ wh-phrase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pragmatically neutral yes-no questions</td>
</tr>
</tbody>
</table>

#### 3.4.1. On yang

It is also necessary to consider where the ‘complementizer’, yang, fits into the Rizzi framework. (One possibility is that yang is an instance of Fin, similar to English that. In Rizzi’s framework, Fin (short for Finite) is the head within the CP field that is ‘closest’ to the IP field, and which, in a sense, ‘agrees’ in finiteness with the IP.) In order to account for the limited distribution of yang – in particular, yang only appears when a DP has been focused and fronted – it is logical that yang has some type of strong feature which needs to be checked against the focused constituent, and an EPP feature which is sensitive to the syntactic category of
this constituent. There is also a second, phonologically null, instance of Fin, also with an EPP feature sensitive to syntactic category, which is only compatible with non-DP focused phrases.10

(24)  The Indonesian (Partial) CP

\[
\begin{array}{c}
\text{FocP} \\
\text{XP Foc'} \\
\text{-kah/-lah [EPP, Foc]} \\
\text{FinP} \\
\text{Fin'} \\
\text{yang [EPP]} \\
\text{IP} \\
\text{Ø [EPP]} \\
\text{t . . . . . . . .} \\
\end{array}
\]

Treating yang as Fin is not, however, a crucial component of the proposal advanced here. For example, a Cole et al. (to appear) style analysis of Malay yang constructions as a type of cleft construction appears to be compatible with the key portions of my proposal sketched out in the previous section. Adopting such an analysis for Indonesian yang constructions would have the additional benefit of capturing the apparent similarities between these constructions and relative clauses that were alluded to in Section 3.1, and consequently merits further exploration.

3.4.2. Support from ellipsis

Returning to the central threads of my proposal, I now present further support, from sluicing, for treating –kah as a unique functional head, merged into the derivation separately from the fronted phrase it cliticizes to. As is well known, sluicing cross-linguistically requires C to be phonologically null (e.g. Chung, Ladusaw, and McCloskey 1995; Merchant 2001). Sluicing in Indonesian sluicing requires a ‘bare’ wh-phrase; -kah, surprisingly, is prohibited. (Also, and less surprisingly, yang is blocked as well.)

(25) Ada seseorang yang membuat kue-kue ini,
exist someone COMP meng-make cake-REDUP DEM
tapi saya tidak tahu siapa(*+kah) (*yang).
but 1SG NEG know who(*+QUEST) (*COMP)
‘Someone made these cakes, but I don’t know who.’

10 Under this analysis, other phonologically null Fin of course exist as well; for example, the Fin in regular declarative sentences, in which no constituents are focused or fronted, would not have an EPP feature.
If –kah were affiliated with the wh-phrase prior to its movement to the left periphery – for example, if siapakah were a lexical item – the prohibition illustrated above would be unexplained. On the other hand, if –kah is, as I’ve argued, a type of C (in particular, Foc), this prohibition is entirely expected.

4. Malay v. Indonesian

Although Indonesian and Malay have much in common lexically as well as structurally, to the extent that the two are to a large extent mutually intelligible, there are non-trivial differences in the structure of interrogatives in the ‘near-standard’ varieties of the two languages. I do not at the moment claim that the proposal described here necessarily extends to Malay, given that the Malay facts (as described by Kader (1981), for ‘Standard Malay’, and Cole, et al. (to appear), for ‘Educated Informal Malay’) are, in some critical domains, quite different from the Indonesian facts. Several of the most dramatic differences are briefly described below, although the necessary fine-grained comparative consideration of focus structures in the two languages remains to be undertaken.

First, as shown above, Indonesian -kah and –lah are restricted to marking clause-initial constituents. In Malay, on the other hand, their placement is much freer: -kah and –lah are able to mark most predicate-internal constituents in situ.\(^{11}\) This suggests that Malay –kah/–lah are not themselves located in Foc, although they are nonetheless licensed by Foc.\(^{12}\)

(26) Fatimah kata Siti membeli [buku itukah] semalam? (Malay)
  Fatimah say Siti meng-buy book DEM+QUES yesterday
  ‘Did Fatimah say that Siti bought THAT BOOK yesterday?’
 (=Cole et al. to appear: 17, ex. 37)

The ‘unique focus’ restriction I’ve identified for Indonesian seems to hold of Malay, as well, but in a slightly different fashion. Kader (1981) reports that certain multiple wh-questions are somewhat acceptable in ‘colloquial’ Malay, although not in the ‘Standard’ Malay he is concerned with. Nonetheless, these questions, however, allow one instance of –kah at most:

(27) Siapa(kah) yang pukul siapa(*kah)?
  who(QUES) REL hit who(*QUES)
  ‘Who hit who?’

\(^{11}\) Kader (1981) notes a handful of idiosyncratic restrictions preventing –kah from marking certain in-situ constituents within the predicate. Additionally, -kah and –lah cannot mark an in situ subject in Malay. Kader (1981) suggests that this restriction is not at all syntactic, but instead has its roots in information packaging preferences: subjects in Malay are normally ‘old’ information, and old information cannot be focused.

\(^{12}\) Kader (1981), working within a transformational framework, proposes a set of phrase structure rules that permit –kah to optionally follow any head. For Kader, focus-fronting a constituent is not syntactically driven; rather, it is an optional process that is ‘pragmatically motivated… by the desire of the speaker to put the focus of his utterance as near as possible to sentence-initial position so that the hearer will pay more attention to it’ (p. 97-98).
The ramifications of these (and other) differences between Malay and Indonesian for the theoretical treatment of questions and focus constructions in these languages deserve to be closely examined, and I plan to undertake this in future research.

5. Conclusions

In this paper, I’ve argued that a unified analysis of Indonesian questions – both constituent and yes-no questions – that does not appeal to the problematic Clausal Typing Hypothesis (Cheng 1997) is possible. I’ve sketched out the broad outlines of this analysis, which centers on the hypothesis that Indonesian wh-movement is focus movement, and that the wide range of question types attested in Indonesian follows directly from the generous array of Foc heads, each with unique properties, in its lexicon.

The analysis outlined here raises a large number of questions which merit further research. Some of these I’ve already noted, and others I’ll briefly address here. For one, I’ve suggested that Indonesian, like Italian, has ‘Unique Focus’ (Stoyanova 2008), since it permits at most one focused wh- or non-wh-phrase per clause. Interestingly, Indonesian is unlike the other Unique Focus languages that Stoyanova describes in that it displays a wh-in-situ strategy; the ramifications of this for the typology of wh-questions-as-focus-constructions also await future exploration.

Second, an additional asymmetry with respect to wh-movement that requires further research involves a split in behavior between AdvPs, on the one hand, and PPs and DPs, on the other. AdvPs, such as kapan ‘when’, cannot appear in situ (28a), and are required to front (28b). On the other hand, both DPs (1) and PPs (28) are able to appear in situ as well as in a left-peripheral position.

\[(28) \quad \text{a. } \text{Ali membeli mobil baru dari siapa/*kapan?} \]
\[\quad \text{Ali meng-buy car new from who/*when} \]
\[\quad \text{‘Ali bought a new car from who/*when?’} \]

\[\quad \text{b. Kapan/Dari siapa Ali membeli mobil baru?} \]
\[\quad \text{when/ from who Ali meng-buy car new} \]
\[\quad \text{‘When/From who did Ali buy a new car?’} \]

While adverbials and PPs, both adjuncts, would reasonably be expected to pattern together, they do not. Instead, ‘nominal adjuncts’ (in the terminology of Cole and Hermon 1998) – that is, PPs which contain a wh-DP – pattern with wh-DPs. The underlying cause of this asymmetry is presently poorly understood, and is complicated by the fact that the restriction in (28a) is not observed by all speakers.

Finally, how does topicalization fit into the overall picture? Rizzi’s framework, adopted here, predicts that topics and focus co-occur. More data needs to be collected, but preliminary exploration indicates this prediction is realized (29).
(29)  
a. Kemarin, buku ini, siapa yang baca?
yesterday, book DEM, who COMP read
‘Yesterday, this book, who read?’

b. ? Siapa, buku ini, yang baca?
who, book DEM, who COMP read
‘Who, this book, read?’

References