

OPTIONAL ELEMENTS IN INDONESIAN MORPHOSYNTAX

Helen N Jeoung

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

Julie Anne Legate, Professor of Linguistics

Graduate Group Chairperson

Eugene Buckley, Associate Professor of Linguistics

Dissertation Committee:

David Embick, Professor of Linguistics

Florian Schwarz, Associate Professor of Linguistics

OPTIONAL ELEMENTS IN INDONESIAN MORPHOSYNTAX

COPYRIGHT

2018

Helen N Jeoung

This work is licensed under the
Creative Commons Attribution-
NonCommercial-ShareAlike 3.0
License

To view a copy of this license, visit

<https://creativecommons.org/licenses/by-nc-sa/3.0/us/>

For my family

ACKNOWLEDGMENT

I have many people to thank for their support, not only during the time I was researching and writing this dissertation, but also during the span of my graduate studies at Penn. The mentions here are brief, but they are heartfelt.

Many thanks to my committee, without whom this dissertation would not be possible. Julie Legate has been a supportive advisor who was always ready to look at my data, read through drafts and return comments. I learned syntax from Julie, but more importantly, she brought out my best work throughout my training at Penn; every page of this dissertation has directly or indirectly benefited from her critical eye. David Embick and Florian Schwarz have also contributed in significant ways to my development as a linguist. Dave always had a comment or suggestion that led me to think about the larger context and implications of my work, and I always left his office or his class with new ideas. Florian introduced me to semantics, and has been a patient reader for various drafts of papers; he always provided timely input just when I needed it.

I have been fortunate to be taught by faculty members at Penn who are engaging and passionate about what they do: Alison Biggs, Gene Buckley, Jami Fisher, Bill Labov, Tony Kroch, Mark Liberman, Rolf Noyer, Don Ringe, Gillian Sankoff, Beatrice Santorini and Meredith Tamminga. Thanks to Meredith for talking to me about syntax and variation, and thanks to Alison for collaborating with me on prepositions. In the main office, Amy Forsyth always had time to answer a question or fix an administrative issue. Additionally, I am grateful for support from the University of Pennsylvania, through the GAPSA-Provost Award for Interdisciplinary Innovation, and a Teece Fellowship for dissertation research.

Thanks to my cohort, Taylor Jones, Karen Wen-Hsin Tseng and David Wilson, with whom I shared my first year grad school. The students, postdocs, and visitors who overlapped with me at Penn contributed to an overall atmosphere of hard work, stimulating conversation and mutual encouragement, and they made the journey more enjoyable: Luke Adamson, Faruk Akkuş, Akiva Bacovcin, Ryan Budnick, Pleng Chanchaochai, Sunghye Cho, Ava Creemers, Kajsa Djärv, Sabriya Fisher, Aaron Freeman, Amy Goodwin Davies, Duna Gylfadottir, Anton Ingason, Tricia Irwin, Ava Irani, Soohyun Kwon, Nari Rhee, Milena Šereikaitė, Kobey Shwayder, Einar Sigurdsson, Betsy Sneller, Robert Wilder and Jérémy Zehr.

Conversations with linguists at conferences and talks have contributed to my work in significant ways; I particularly thank Edith Aldridge, I Wayan Arka, Abby Cohn, Peter Cole, Mitcho Erlewine, Idan Landau, Bradley McDonnell, Norvin Richards, Anna Szabolcsi and Ekarina Winarto. I also acknowledge Bill Davies, whose grammar of Madurese has been a valuable resource; I regret that I did not have the opportunity to meet him in person.

My research trips to Indonesia were fruitful and interesting because of the following people. Bapak Isya Mahfud has contributed more data and insights to this dissertation than anyone else, and his facility with multiple languages is invaluable. I also thank Ibu Maimuna for checking data with me, for helping me get around when I was traveling alone, and for her longtime friendship. *Mator sakalangkong*. I have been fortunate to collaborate with university students, who worked as my interns during the summer of 2015 and also cheerfully became language consultants: Iline Megale, Gabriella Melody, Ilkham Hidayat, M Khoirun Najib, and Nadhilah H Semendawai. The following

consultants also provided data for Indonesian and other languages: Amelia Dwi Imanda, Ayunda Puty Sekavani, Hainona Izza Golia, Muhamad Saiful, Risalarus Sa'adah, Wilyun Hamim, and in Philadelphia, Dinah Maria. A number of people provided logistical support for me to conduct research in East Java: Professor Lilis Lestari Wilujeng at Ma Chung University connected me to undergraduate students; Ibu Mira connected me to students at UNISMA; Bapak Billy Mbiliyora and the staff at KPN graciously allowed me to use classroom space and volunteered as practice consultants. Esther, Sarah, Mira and other friends in town always made time for a warm welcome and a meal. *Terima kasih banyak semuanya.*

I am grateful for friends in Philadelphia and beyond, especially those who provided support while I was dissertating. Sharon Song went out of her way to send me notes and gifts, just because. Thanks to Elizabeth McGee Hassrick for chats about academia, and to Jenn Lee for chats about life. Paula Yoo was a lifesaver, driving across town for pickups and offering help when it was most needed. Lindsay Kijewski was my biggest fan, and always had a smile and a hug for me. Thanks to Lindsay, Jonathan Wood, Umi Howard, my Lipman cohort, Anne Ferola and the team at CHIP for meaningful experiences outside of linguistics.

Finally, my family. Young and Yeuh Lee always had an open door for me, without exception. My mother has been a rock of unconditional love and faith, and a role model for a full-time working mother holding it all together. My father, Jun Imatome, had an extensive library at home, and spoke three languages fluently: I am sure he planted the original seed that grew into my interest in language, and I wish I could thank him now. Thanks to my in-laws for their encouragement and support, and to Tracy for coming down to Philly to help out. Thanks also to all my siblings. I especially owe a lot to Naomi: look, I made it! Amelia and Noah have been with me since day one, through the valleys as well as the peaks, and they have been my anchor throughout. And more than anyone else, my husband has had the up-close view of this journey, and has made sacrifices so that I could pursue linguistics.

Thank you.

ABSTRACT

OPTIONAL ELEMENTS IN INDONESIAN MORPHOSYNTAX

Helen Jeoung

Julie Anne Legate

This dissertation investigates the syntax and morphology of several functional morphemes that display surface optionality in Indonesian. Three case studies consider how syntactic environments constrain optional realization. Chapter 2 investigates the declarative complementizers *bahwa* and *kalau*, which are disallowed in case of A-bar movement; I show that *bahwa* is also disallowed in *wh*-in situ questions that do not involve movement. These facts are developed into an analysis of *wh* phrases and the structure of *wh* questions in Indonesian. I also propose that the morpheme *yang*, as well as the null form of the complementizer, constitute a pattern of morphological *wh*-agreement on C. Chapter 3 discusses the verbal prefixes *meN-* and *ber-*, which have received varied analyses in the literature. I argue that *meN-* and *ber-* participate in *wh*-agreement resulting from A-bar movement, and argue against previous analyses that assume that A-movement results in a similar deletion. In addition, I differentiate between deterministic properties that are relevant in the syntax, and non-deterministic properties of *meN-* and *ber-* that are extra-syntactic. This distinction accounts for a number of puzzling properties that have been observed for these prefixes. Chapter 4 discusses possessor sub-extraction in Indonesian, with additional data from similar constructions in Javanese and Madurese. I pursue a novel analysis of the nominal suffix *-nya*, which is optional in possessive DPs: in possessor extraction, this suffix is a pronunciation of the head D. The analysis of *wh*-agreement is extended to the DP domain, where *-nya* marks A-bar movement on phase heads; the consequence is that DP is a phase for syntactic movement. One language-specific finding in this dissertation is that morphological *wh*-agreement applies across three domains: complementizers, verbs and possessive nominals. This has cross-linguistic implications for the phasehood of DP and *wh*-agreement patterns. More broadly, the dissertation contributes a syntactic approach to the analysis of variable morphemes, revealing how multiple factors constrain surface optionality.

TABLE OF CONTENTS

| | |
|---|----|
| ACKNOWLEDGMENT | iv |
| ABSTRACT | vi |
| LIST OF TABLES | x |
| CHAPTER 1 Introduction | 1 |
| 1.1 Standard and colloquial varieties of Indonesian | 3 |
| 1.2 Data and methods | 4 |
| 1.3 Outline of the dissertation | 5 |
| CHAPTER 2 Complementizers | 7 |
| 2.1 Introduction | 7 |
| 2.1.1 Surface alternations in C | 7 |
| 2.1.2 Background: Complementizer-trace effects in other languages | 9 |
| 2.1.3 Background: Wh-agreement in other languages | 11 |
| 2.1.4 Outline of chapter | 15 |
| 2.2 Embedded declarative clauses | 15 |
| 2.2.1 Declarative complement clauses (finite) | 15 |
| 2.2.2 Sentence-initial C | 19 |
| 2.3 Loss of optionality: A-bar movement and questions | 20 |
| 2.3.1 Properties of A-bar movement in Indonesian | 20 |
| 2.3.2 Pseudo-cleft structure | 24 |
| 2.3.3 Nominal movement vs. non-nominal movement | 31 |
| 2.3.4 Three strategies for constituent wh questions | 32 |
| 2.3.5 Strategies for forming polar questions | 35 |
| 2.3.6 Complementizers and A-bar movement | 37 |

| | | |
|---|--|-----|
| 2.3.7 | Complementizers in questions | 44 |
| 2.4 | Deterministic and non-deterministic conditions on Indonesian C | 63 |
| 2.5 | Chapter summary | 66 |
| CHAPTER 3 Verbal prefixes <i>meN-</i> and <i>ber-</i> | | 68 |
| 3.1 | Introduction | 68 |
| 3.1.1 | An overview of the proposal | 70 |
| 3.1.2 | Organization of chapter | 71 |
| 3.2 | Background | 72 |
| 3.2.1 | Verbal prefixes in Indonesian | 73 |
| 3.2.2 | Invariant prefixes | 74 |
| 3.2.3 | On the variability of <i>meN-</i> and <i>ber-</i> | 77 |
| 3.3 | Previous analyses of <i>meN-</i> and <i>ber-</i> | 79 |
| 3.4 | Deterministic factors affecting the realization of <i>meN-</i> and <i>ber-</i> | 91 |
| 3.4.1 | Nominal A-bar movement | 91 |
| Excursus: | Is <i>meN-</i> sensitive to A-movement? | 97 |
| 3.4.2 | Other syntactic and semantic constraints on optionality | 101 |
| 3.5 | Variable contexts and non-deterministic factors | 105 |
| 3.5.1 | Variable deletion rule and its effects | 105 |
| 3.5.2 | Other consequences of variable rule application | 109 |
| Excursus: | Wh-agreement applies to all verbs | 111 |
| 3.6 | Chapter summary | 117 |
| CHAPTER 4 Possessor Extraction | | 118 |
| 4.1 | Introduction | 118 |
| 4.1.1 | Background | 118 |
| 4.1.2 | Internal and external possession | 119 |
| 4.1.3 | Addressing variability in <i>-nya</i> | 121 |
| 4.1.4 | Outline of chapter | 122 |

| | | |
|--------------|--|-----|
| 4.2 | Patterns of possessor extraction in Indonesian | 123 |
| 4.2.1 | Sub-extraction of subject possessors | 123 |
| 4.2.2 | Sub-extraction of object possessors | 126 |
| 4.2.3 | Long distance sub-extraction | 127 |
| 4.2.4 | Sub-extraction from other positions | 129 |
| 4.2.5 | Interim summary | 131 |
| 4.2.6 | Possessor extraction in related languages | 132 |
| 4.3 | The suffix <i>-nya</i> | 138 |
| 4.3.1 | <i>-Nya</i> as possessor agreement | 138 |
| 4.3.2 | <i>-Nya</i> in non-possessive contexts | 145 |
| 4.3.3 | Definiteness effects and <i>-nya</i> | 147 |
| 4.4 | A derivational analysis for possessor extraction | 150 |
| 4.4.1 | Movement through SpecDP | 151 |
| 4.5 | Implications of possessor extraction in Indonesian | 158 |
| 4.5.1 | Phase head D and <i>wh</i> -agreement | 158 |
| 4.5.2 | EPP and subject extraction | 159 |
| 4.5.3 | Revisiting A-bar extraction in Indonesian-type languages | 163 |
| 4.5.4 | Implications for the left periphery in Indonesian | 165 |
| 4.6 | Chapter summary | 167 |
| CHAPTER 5 | Conclusion | 169 |
| APPENDIX | | 171 |
| BIBLIOGRAPHY | | 186 |

LIST OF TABLES

| | | |
|---------|--|-----|
| TABLE 1 | Cole and Hermon's (1998) analysis of WH-words in Malay | 49 |
| TABLE 2 | Semantico-syntactic features on C | 64 |
| TABLE 3 | Verbal prefixes in Indonesian | 73 |
| TABLE 4 | Possessive DPs with possessum <i>rumah</i> 'house' | 141 |

CHAPTER 1 Introduction

In this dissertation I investigate the syntax and morphology of functional elements that display optionality. Three case studies are presented in this dissertation, each of which discusses one or more optional morphemes in Indonesian. The first case study examines the declarative complementizers *bahwa* and *kalau* and their alternation with other forms of C; the second discusses the verbal prefixes *meN-* and *ber-*; the third observes the effect of possessor extraction on the nominal suffix *-nya* that occurs in possessive DPs. These morphemes display surface optionality, as illustrated in the examples below:

(1) Optional complementizers

Susan ber-pikir { *bahwa* / *kalau* / \emptyset } Tina mem-beli se-buah tas kemarin.
Susan MV-think C Tina AV-buy one-Class bag yesterday
'Susan thinks that Tina bought a bag yesterday.'

(2) Optional verbal prefix *meN-*

Aku { *dapat* / *men-dapat* } hadiah.
1s receive AV-receive gift
'I received a gift.'

(3) Optional verbal prefix *ber-*

Dia { *bahasa* / *ber-bahasa* } Indonesia, aku { *bahasa* / *ber-bahasa* } Inggris.
3s language MV-language Indonesian 1s language MV-language English
'She spoke Indonesian, I spoke English.'

(4) Optional possessive suffix *nya-*

{ *Rumah* *Adi* / *Rumah-nya* *Ali* } di-rata-kan kemarin.
house *Adi* house-D *Ali* PV-flat-Appl yesterday
'Adi's house was destroyed yesterday.'

An individual speaker produces both the overt form and the non-overt form; the optional morpheme may be present or omitted without any semantic consequence. This dissertation, then, is concerned with the grammar of individual speakers, and the (intraspeaker) variability displayed by particular morphemes, rather than variability among speakers or variability among different languages/varieties. Additionally, this dissertation is primarily concerned with syntactic structure and

operations that are relevant to the optional morphemes.

Two overarching questions drive this research. These are concerned with the interaction between syntax and optionality:

- (a) Which syntactic conditions constrain the surface optionality of the morpheme(s)?
- (b) How does syntax interact with extra-syntactic factors to produce the overall distribution of the optional forms?

Each case study addresses the questions in both (a) and (b). Because Indonesian has a number of different functional morphemes that display surface optionality, this makes it possible to investigate several different domains within one language (complementizers, verbal structure, possessive nominals), while also observing interactions among these three elements within a single clause.

The approach to Question (a) begins with a descriptive account of the full range of variability displayed by the morpheme(s), with a particular focus on contexts in which the morpheme is *not* variable. This dissertation contributes novel data from colloquial spoken Indonesian (and the related languages Javanese and Madurese), with the goal of presenting observations and patterns that have not been previously noted in the literature on these languages. Question (a) is concerned with taking optionality seriously as part of syntactic analysis.

Question (b) takes an integrative approach to optionality by examining several factors that affect optionality. This line of research is partially motivated by incomplete treatments of variable morphosyntactic phenomena which do not include variability as a formal component of the analysis. For example, the optional nature of *meN-* in clauses such as (2) has been long observed (Wouk 2004; Sneddon 1996; Sneddon et al. 2012). However, optional *meN-* is usually assumed to be “optional pronunciation” or “optional dropping” of the morpheme. This assumes that the optionality is a PF phenomenon that has nothing to do with morphosyntax; cursory mentions of this optional property omit any obligatory contexts in which the morpheme is required or disallowed. I assume that a comprehensive account of the optional morphemes under study requires understanding of the semantic and syntactic structure, as well as post-syntactic (morphophonological) factors

and extra-grammatical (pragmatic, discourse-related) factors. The overall picture that emerges from this approach is that the distribution of these functional morphemes in Indonesian can only be accounted for by integrating syntactic structure/operations with extra-syntactic factors that influence post-syntactic variability.

While the scope of this investigation is a language-internal account of variable elements in Indonesian grammar, this research has cross-linguistic implications for our understanding of clause structure more generally. I present findings that have consequences for our general understanding of complementizer systems and *wh*-in situ questions; patterns of *wh*-agreement across several domains; the phasehood of CP, VoiceP and DP; and possessor extraction. This dissertation also contributes a syntactic treatment of morphemes that display optionality, with consequences for analyses of surface variation cross-linguistically.

1.1. Standard and colloquial varieties of Indonesian

It is well-known among researchers studying Indonesian that Standard Indonesian (or *Bahasa Indonesia*, ISO: ind), the national language of Indonesia, is a prescriptive variety that few speakers claim as their native language. Standard Indonesian is used for official activities such as education, business and government (see Sneddon 2003 for an overview and history). Colloquial Indonesian, on the other hand, is not a monolithic variety but rather a diverse set of varieties of Indonesian spoken by communities throughout Indonesia. While Standard Indonesian is well-documented in grammars (Macdonald and Dardjowidjojo 1967; Dardjowidjojo 1978; Sneddon 1996; Sneddon et al. 2012; among others), colloquial varieties of Indonesian are generally under-described or undocumented (with some exceptions, such as Sneddon 2006, which describes colloquial Jakartan Indonesian).

This dissertation is primarily concerned with the variety of colloquial Indonesian that is spoken on the island of Java (outside of Jakarta), particularly East Java. Colloquial Indonesian differs from Standard Indonesian in a number of ways, but the two varieties are not completely distinct; many clauses are well-formed in both varieties. The distinction between Standard Indonesian and colloquial Indonesian should not be understood as the difference between formal and informal

speech, which is an over-simplification. Although Standard Indonesian is associated with formal situations and considered stiff if used in daily conversation, colloquial Indonesian is not strictly informal. By referring to colloquial Indonesian, I mean the variety of language that is used spontaneously by speakers in everyday situations, and about which speakers have intuitions concerning grammaticality and acceptability. Colloquial Indonesian, then, incorporates both informal, familiar speech as well as polite speech (as when speaking to an important person or an elder). Throughout this dissertation, data is generally from colloquial Indonesian, except where data is drawn from written sources; all data from written sources has also been further checked by consultants.

One final note about Malay. Indonesian and Malay are closely related varieties, although their syntax, morphology and phonology deviate in a number of significant ways. I cite literature on Malay where it is significant to the analysis, but I do not assume that generalizations from Malay will always carry over to Indonesian. I have indicated wherever other sources are based on Malay data.

1.2. Data and methods

The speakers providing data for this research are from the island of Java, and were educated in Standard Indonesian (speakers have at least a high school education; most have university education as well). All speakers speak some colloquial variety of Indonesian on a daily basis with friends, other students, colleagues, and strangers. In addition, all speakers, except one, are fluent in at least one other language that is used in the home; these languages include Javanese and Madurese, which are also prominent on the island of Java.

Research with Indonesian-speaking consultants took place over several periods, between 2015 and 2018. The research was conducted in several locations, including East Java, Indonesia and Philadelphia, Pennsylvania. In addition, part of this research was also conducted by communicating (typing) over the internet with Indonesian speakers with whom I had established a previous working relationship during in-person research. Data were elicited directly from consultants, taken from instances of spontaneous speech, or drawn from written sources. Because optionality is a central concern of this dissertation, all original data were re-checked with consultants for the avail-

ability of multiple realizations of a particular morpheme. All examples have been checked by at least one speaker other than the speaker who provided the initial data.

1.3. Outline of the dissertation

In Chapter 2, I investigate the complementizers *bahwa* and *kalau*, which introduce embedded declarative clauses; *kalau* is an informal variant of C, while *bahwa* is a more formal variant. I claim that *yang* and null C are part of the complementizer system in Indonesian; this proposal is the first to identify four forms of declarative C. *Bahwa* and *kalau* are optional in many embedded clauses. However, they are not completely in free variation with a null form: *bahwa* and *kalau* are disallowed in case of A-bar movement, which was first observed by Saddy (1991). I argue that this pattern is a type of wh-agreement on C, and that *yang* and null C also participate in wh-agreement; Indonesian complementizers instantiate a new pattern of wh-agreement that has not previously been attested (cf. Reintges et al. 2006). Another novel finding in this chapter is that *bahwa* is disallowed in wh-in situ questions that do not involve movement. These facts are developed into an analysis of wh-phrases and the structure of wh questions in Indonesian.

In Chapter 3 I consider the verbal prefixes *meN-* and *ber-*, which have received varied analyses in the literature. Taking *meN-* and *ber-* to occupy the position of the functional head Voice, I propose that *meN-* and *ber-* participate in morphological wh-agreement: A-bar movement through SpecVoiceP requires a null prefix. I argue against previous analyses that assume that A-movement also has the same morphological effect. In addition, I differentiate between deterministic properties that are relevant in the syntax, and non-deterministic properties of *meN-* and *ber-* that are extra-syntactic. I propose that this distinction accounts for a number of puzzling properties that have been observed for these prefixes.

Chapter 4 discusses possessor sub-extraction in Indonesian, with additional data from similar constructions in Javanese and Madurese. I show that a clefted possessor occurring at the left periphery of the clause is not merged in this high position, but rather that it is extracted from its possessive DP, leaving the possessum in situ. The possessor undergoes successive-cyclic A-bar movement through the clause, resulting in wh-agreement in the form of the suffix *-nya*. One con-

sequence of this investigation is a novel analysis for *-nya*, which is optional in possessive DPs. I argue that in possessor extraction, this suffix is a pronunciation of the head D. I extend the analysis of *wh*-agreement to the DP domain, arguing that *-nya* marks A-bar movement on phase heads, with the consequence that DP is a phase for syntactic movement.

In Chapter 5 I summarize the main findings of the dissertation and present overall conclusions.

CHAPTER 2 Complementizers

2.1. Introduction

2.1.1. Surface alternations in C

In this chapter I discuss the Indonesian complementizers *bahwa*, *kalau* and *yang*, as well as the absence of an overt complementizer, which I refer to as null C. In some contexts, realization of C is deterministic, requiring a particular surface form. In other contexts however, the form of C is optional, in the sense that semantic differences do not arise from the choice of form. The overall goal of this chapter is to demonstrate that accounting for the distribution of complementizers in Indonesian requires an analysis of the syntax and semantics of C; an explanation of surface variability among *bahwa*, *kalau*, *yang* and null C must incorporate grammatical factors.

As a brief overview of the data, these complementizers occur in finite clauses, including both declarative and interrogative sentences. In sentences such as (1), the complementizer that introduces an embedded declarative clause may occur as *bahwa*, *kalau* or null C (represented as \emptyset). This is an optional alternation among the three realizations of C. (2) illustrates a different pattern. The wh word *siapa* ‘who’ has moved from an embedded clause, and occurs in a sentence-initial position, where it must be followed by *yang*. I propose that *yang* does not mark focus on the nominal that it follows, as has been suggested in the literature, but rather that it is a form of C that alternates with *bahwa*, *kalau* and null C.

- (1) Susan ber-pikir { *bahwa* / *kalau* / \emptyset / **yang* } Tina mem-beli se-buah tas kemarin.
Susan MV-think C Tina AV-buy one-Class bag yesterday
‘Susan thinks that Tina bought a bag yesterday.’
- (2) Siapa { **bahwa* / **kalau* / * \emptyset / *yang* } Susan pikir { **bahwa* / **kalau* / \emptyset / **yang* } —
who C Suan think C
mem-beli se-buah tas kemarin?
AV-buy one-Class bag yesterday
‘Who does Susan think bought a bag yesterday?’

In long-distance extraction, for the embedded clause from which the wh word has been extracted,

all overt forms are disallowed and null C is required (2). I discuss the implications of this data for our understanding of wh-agreement in complementizer systems, both for Indonesian and cross-linguistically.

Long-distance movement over a complementizer, as in (2), is not the only environment that disallows an overt form of C. In-situ questions such as (4) also disallow the overt complementizer *bahwa*, even without overt movement.

- (3) Pak Dadang meng-harap { *bahwa* / *kalau* / \emptyset } mereka akan men-diskusi-kan masalah
 Mr Dadang AV-expect C 3p will AV-discuss-Appl problem
 pertanian.
 agriculture
 ‘Mr Dadang expects that they will discuss the issue of agriculture.’
- (4) Pak Dadang meng-harap { **bahwa* / *kalau* / \emptyset } mereka akan men-diskusi-kan masalah
 Mr Dadang AV-expect C 3p will AV-discuss-Appl problem
 apa?
 what
 ‘What issue does Mr Dadang expect that they will discuss?’

The pattern in (3-4) is interesting because a wh phrase that remains in situ places constraints on a higher complementizer. In addition, *kalau* has been omitted from previous analyses of complementizers in Indonesian, or mentioned only as a conditional morpheme. By treating *kalau* as a complementizer on par with *bahwa* in this chapter, I show that the two morphemes pattern together in nearly every syntactic environment. Wh-in situ questions are the exception: in (4) *bahwa* is disallowed while *kalau* is possible, a pattern that has not previously been observed.

This chapter addresses both obligatory realizations of C as in (2), as well as optional realizations of C as in (1). I treat *bahwa*, *kalau*, *yang* and the null form as belonging to the category C whose realization is determined at various points in a derivation, rather than by one syntactic or morphological process. The surface realization of C is first determined derivationally, that is, by the syntactic and semantic requirements of structure building, including A-bar movement and the formation of questions; post-syntax, realization of C is also subject to variable operations that apply after insertion of morphophonological forms. The two extra-syntactic mechanisms that give rise to

variable surface realization of C are variable deletion, and pragmatic selection between formal and informal vocabulary items. This approach relies on both features in the syntax as well as extra-syntactic processes to derive the obligatory and optional alternations seen in (1-4). More broadly, I demonstrate that an understanding of both internal (grammatical) and external (extra-grammatical) factors is necessary to account for the distribution of forms that alternate in a single surface position.

2.1.2. Background: Complementizer-trace effects in other languages

Indonesian is not unique in having multiple forms of C, nor in allowing optional realization of C. The Indonesian complementizer system shares a number of properties already attested in other languages. In this section I discuss similarities between the Indonesian pattern and other complementizer systems.

Many languages have optional complementizers for introducing embedded clauses. For example, the Swedish complementizer *att* is optional in (5). In cases of subject extraction however, it must be null (6). This contrasts with object extraction, where the complementizer retains its optional realization (7).¹

- (5) Jag tycker (att) studenterna borde komma i tid.
 I think that students.Def should come in time.
 'I think that the students should come on time.'
- (6) Vem tror du (*att) ___ skulle komma i tid?
 who believe you that should come in time
 'Who do you believe should come on time?'
- (7) Vad tror du (att) John köpte ___?
 what think you that John bought
 'What do you think that John bought?'

The same pattern is also well-known in English. The English complementizer *that* can be null, except in case of subject extraction or relativization:

- (8) the guy (that) you claim (that) Carlos hired ___ yesterday

¹Thanks to Kajsa Djärv for these examples.

- (9) the guy (that) you claim (*that) ___ hired Bill yesterday

This pattern is generally known as the Comp(lementizer)-trace effect, which has been observed in a number of languages and has been much discussed (see e.g. Perlmutter 1968, 1971; Chomsky and Lasnik 1977; Bresnan 1977; Rizzi 1982; Pesetsky 1982; Zaenen 1983; Lasnik and Saito 1984; Culicover 1993; Pesetsky and Torrego 2000; Ishii 2004; among others.) Comp-trace effects are observed both in languages that allow complementizers to be optionally null (like Swedish and English), as well as languages in which the complementizer must be pronounced (such as French).

However, the Indonesian pattern differs from Comp-trace effects in two ways. One difference is that in Comp-trace effects, the C that is local to extraction is affected, but not a higher C. Consider long-distance subject extraction in English:

- (10) Laura claims (that) Mary thinks (that) Bill will love the new film.
 (11) Who does Laura claim (that) Mary thinks (*that) ___ will love the new film?

In (11) the lower C is the first that is crossed by A-bar movement, and consequently must be null, while the higher C may be overt. This contrasts with Indonesian long distance extraction, in which multiple complementizers along the path of movement are affected:

- (12) ?makanan yang kamu pikir \emptyset Susan bilang \emptyset Tina ambil ____.
 food C.Foc 2s think C Susan say C Tina take
 ‘the food that you think that Susan said that Tina took’

In Indonesian, every C that is crossed by nominal A-bar movement must be null. Note that for many speakers, the extraction across multiple clause boundaries with active verbs in (12) is somewhat degraded in acceptability. However, there is a strong contrast between (12) and the ungrammaticality of (13), which shows that extraction over any overt C is impossible.

- (13) makanan yang kamu pikir { *bahwa / *kalau } Susan bilang \emptyset Tina ambil ____.
 food C.Foc 2s think C Susan say C Tina take
 ‘the food that you think that Susan said that Tina took’

A second difference between Comp-trace effects and Indonesian C is that the alternations in Indonesian are not sensitive to a distinction between subjects and non-subjects. Comp-trace effects surface in different ways cross-linguistically, but the range of effects shares the generalization that special morphology on C is required when the (subject) argument immediately to the right of the complementizer undergoes A-bar movement. Other arguments do not trigger the special morphology, which means that various approaches to explaining Comp-trace effects must explain why subjects, but not other arguments, require special marking on C. (A recent overview of analyses for Comp-trace effects is summarized in Pesetsky 2017.)

Explanations for Comp-trace effects, then, will have something to say about subject extraction, but will not account for all the Indonesian facts, since null C is required with both subject and object extraction. Instead, in this chapter I discuss this particular alternation in Indonesian C as a type of wh-agreement.

2.1.3. Background: Wh-agreement in other languages

Wh-agreement is a type of morphological marking that occurs as a reflex of A-bar movement (Chung 1982; Zaenen 1983; Georgopoulos 1985; Tuller 1986; Chung and Georgopoulos 1988; Haik 1990; Watanabe 1996; among others).² In a variety of languages, an A-bar dependency triggers special morphology, which can surface on complementizers, agreement morphology on V or special aspect marking. The appearance of the special morphology does not have any semantic content, and appears to only reflect syntactic movement.

For example, McCloskey (2001, 2002) shows that Irish complementizers take a special form if an A-bar dependency crosses the clause boundary. Finite declarative clauses are introduced by the complementizer *go* (which is also marked for tense) (14). When there is A-bar movement that crosses the clause boundary, illustrated by the relative in (15), the complementizer must occur as one of the allomorphs of *aL*.³ A third complementizer, *aN*, is used when a resumptive pronoun

²Or wh-movement in the sense of Chomsky 1977, 1995; I use the term A-bar movement to avoid confusion with discussion that is specific to interrogative wh words and the movement of wh phrases.

³McCloskey notes that *aL* is not considered a complementizer in traditional Irish grammars, but rather is called a “direct relative particle.”

occurs instead of a gap (16).

- (14) Deir siad gur ghoid na síogaí í.
say they C-[past] stole the fairies her
'They say that the fairies stole her away.'
- (15) an ghirseach a ghoid na síogaí
the girl aL stole the fairies
'the girl that the fairies stole away'
- (16) an ghirseach ar ghoid na síogaí í
the girl aN stole the fairies her
'the girl that the fairies stole away' (McCloskey 2001:67)

The difference between *aL* and *aN* is the type of element that occurs in the embedded clause. In McCloskey's analysis, the complementizer is realized as *aL* when it binds a trace that is left behind by A-bar movement. *aL* introduces every clause from which extraction has occurred, so that it occurs recursively in long distance extraction:

- (17) rud a gheall tú a dhéanfá
thing aL promised you aL do[COND-S2]
'something that you promised that you would do' (McCloskey 2001:68)

In contrast, the complementizer *aN* binds a resumptive pronoun (*í* 'her' in 16). Interestingly, McCloskey argues that in cases of resumption, no movement has occurred; rather, the head of the relative is generated in its high surface position. This means that *aN* does not mark A-bar movement, but only the existence of an A-bar dependency.

Another language that exhibits wh-agreement is Chamorro (Chung 1998; Reintges et al. 2006). Chamorro exhibits two types of wh-agreement: with special inflection on the verb, and also with different forms of C. The verbal type is illustrated below with the verb *fa'gasi* 'wash.'

- (18) Ha-fa'gasi si Juan i kareta.
agr-wash Juan the car
'Juan washed the car.'
- (19) Hayi *fuma'gasi* t i kareta?
who? WH[nom].wash the car

‘Who washed the car?’

(20) Hafa *fi*na’gasése-nña si Henry t pära hagu?
what? WH[obj].wash.Prog-agr Henry for you
‘What is Henry washing for you?’

(21) Hafa pära fa’gase-mmu ni kareta t ?
what? Fut WH[obl].wash-agr Obl car
‘What are you going to wash the car with?’ (Chung 1998:236)

In (19)-(21), A-bar movement of the question word occurs with special morphology on the verb that indicates that the moved argument has nominative (19), objective (20) or oblique (21) Case. This type of wh-agreement is recursive, occurring on every verb along the path of movement. The second type of wh-agreement in Chamorro, which Chung calls “Operator-C agreement,” requires special forms of the complementizer. A special form of C marks both Case and the type of argument that has moved (wh phrase, relativized DP or focused element). As illustrated below, a moved wh-phrase occurs with the complementizer *na* (22), while a locative phrase that has undergone focus movement occurs with *nai* (23) and a relativized argument occurs with *änai* (24):

(22) Manu *na* sumásaga hao?
where? C.Q AGR.live.PROG you
‘Where are you living?’

(23) Gi gima’-mami *nai* sumásaga si Jess
LOC house-AGR C_{EMPH} AGR.live.PROG Jess
‘Jess is living in *our house*.’

(24) Manggi i gima’ [*änai* ma-sangani i päli’ [*na* pära u-saga]]?
where?.is the house C_{REL} WH[obj2].AGR-tell the priest C FUT AGR-live
‘Where is the house where they told the priest that he should live?’

Unlike wh-agreement on Chamorro verbs, which is required on every verb along the path of movement, Operator-C agreement occurs only on the highest complementizer. In the long distance relative in (24), wh-agreement is marked on the highest complementizer, *änai*, while intermediate complementizer occurs as *na*, which is the form that occurs when there is no movement.

In this chapter I show that like Irish and Chamorro, Indonesian exhibits wh-agreement in its

complementizer system. I analyze *bahwa*, *kalau*, *yang* and null *C* as complementizers whose form is determined (in part) by overt A-bar movement. The optionality among *bahwa*, *kalau* and null *C* is lost when a nominal is extracted across *C*, and obligatory forms of *C* occur instead.

Wh-agreement has a number of properties that are shared cross-linguistically. Reintges, LeSourd and Chung (2006) characterize the “typological profile” of wh-agreement as follows:

- (i) Wh-agreement occurs in constructions typically derived by wh-movement, that is, constituent questions, focus constructions and relatives. This is true of the Indonesian pattern, which occurs in relatives, clefted constructions and moved-wh questions.
- (ii) Morphological wh-agreement surfaces on V and C/T.
- (iii) The morphology of wh-agreement is not sensitive to DP-internal features (i.e. person, number, gender), even in languages that have a rich agreement system for phi features.
- (iv) Wh-agreement can register other features of the moved argument, such as Case, grammatical function or category (as already seen in Chamorro).
- (v) Wh-agreement comes in two “flavors”: “Either (1) the agreement is recursive, in the sense that it surfaces in *every* designated head on the path of *wh*-movement; or else (2) the agreement is nonrecursive, meaning that it surfaces only on the *highest* designated head on the path of *wh*-movement. As far as we know, no other patterns are attested.” (Reintges, LeSourd and Chung 2006:167)

Indonesian contributes a new perspective on the nature of wh-agreement, because generalizations (ii) and (v) do not cover the full range of Indonesian data. I show that the morphology of Indonesian wh-agreement does not fall into either of two patterns described in (v). In long-distance A-bar movement, all complementizers crossed by movement must show wh-agreement, but the highest *C* occurs in a different form from intermediate *C* (discussed in Section 2.3.6). Indonesian A-bar movement therefore constitutes a third “flavor” of wh-agreement.

Regarding point (ii), several authors have noted that morphological wh-agreement appears

to occur in only two domains. Zaenen argues that morphological wh-agreement is only marked on complementizers and verbs (Zaenen 1983; Watanabe 1996). This generalization finds some support in Indonesian: Saddy (1991) first noted that wh-movement is not only marked in the complementizer system, but also on active verbs. However, Indonesian also marks wh-agreement within the DP, in cases of possessor sub-extraction (see Chapter 4). This extends the possible range of wh-agreement to nominal morphology. Furthermore, the pattern exhibited by Indonesian wh-agreement suggests that Zaenen's generalization is better framed with reference to syntactic phases.

This chapter, then, is the first part of a three-part investigation into optional morphemes that also participate in wh-agreement. In this chapter I focus on complementizers; Chapter 3 discusses wh-agreement on verbs; Chapter 4 discusses wh-agreement within the DP. However, wh-agreement is not the only factor that determines the distribution of complementizers in Indonesian; I show that questions also constrain possible forms of the complementizer. I show that this is a mechanism separate from wh-agreement.

2.1.4. Outline of chapter

The organization of this chapter is as follows. Section 2.2 shows that the declarative complementizers *bahwa* and *kalau* may introduce embedded declarative clauses in Indonesian, but that they are optional. Section 2.3 shows that in certain syntactic contexts, this optionality is lost, i.e. in cases of A-bar movement and in question formation. Section 2.4 presents an analysis that derives the morphological forms of the complementizer. Section 2.5 concludes.

2.2. Embedded declarative clauses

2.2.1. Declarative complement clauses (finite)

I begin with finite declarative sentences that do not involve overt A-bar movement. Embedded declarative clauses can be introduced by the formal complementizer *bahwa* or the informal complementizer *kalau*.

Polite/formal Indonesian and familiar/informal Indonesian are differentiated in a number of ways, most notably in the choice of personal pronouns, but many lexical items are not specified for a particular register. However, the complementizer *bahwa* is clearly associated with formal, standard Indonesian while *kalau* is associated with informal *bahasa sehari-hari* ‘everyday language.’ Kushartanti (2014), citing work on Standard Indonesian by Kridalaksana (1975), identifies the use of *bahwa* as an indicator of children’s acquisition of formal Indonesian.⁴ This means that when the complementizer *kalau* appears in written Indonesian, it is judged to be inappropriate or non-standard. (When the complementizer *kalau* does appear in informal texts, it is often spelled *kalo*). The examples given throughout this chapter that occur with *kalau* should be taken as illustrations of colloquial, spoken Indonesian. Examples with *bahwa* are possible in both writing and speech.

Throughout much of this discussion, I show that *bahwa* and *kalau* can occur in the same position. Internal grammatical factors that affect one of these forms will affect the other as well, and both forms are disallowed in the same contexts. The one exception is wh-in situ questions, where *bahwa* is disallowed but *kalau* is possible (see Section 2.3.7). For now, I set aside this exceptional case, and treat the two forms as having the same syntactic distribution as declarative complementizers.⁵ Since internal factors do not differentiate *bahwa* from *kalau*, the difference between the two forms is a pragmatic one, i.e. attributed to style and situational formality. *Bahwa* is more common in formal or careful speech (as well as formal writing), while *kalau* is prescriptively avoided in written Indonesian but occurs in informal speech.

In addition, *bahwa* and *kalau* may be optionally absent, which I refer to as a null complementizer and represent as \emptyset . Null C is neutral with respect to register, and is possible in either formal or informal speech.

The result is in an optional three-way alternation in C:

- (25) Kita tahu { *bahwa* / *kalau* / \emptyset } uang itu harus di-alir-kan, bukan di-timbun.
 1p.Incl know C money that must PV-flow-Appl Neg PV-hoarded

⁴However, the complementizer *kalau* is not discussed either of these studies.

⁵The form *kalau* also occurs as the conditional morpheme. As a non-conditional complementizer, it is bleached of conditional semantics.

‘We know that the money must be dispensed, not hoarded.’

- (26) Sara ber-pikir { bahwa / kalau / ∅ } Melly meng-harap { bahwa / kalau / ∅ } Susan akan
 Sara MV-think C Melly AV-hope C Susan will
 menang.
 win
 ‘Sara knows that Melly hopes that Susan will win.’

Here, the alternation among *bahwa*, *kalau* and null C is optional in the sense that all three realizations are possible in the same position. Below, I show that no semantic differences arise from the choice of one form over another; truth-conditions, presuppositions and entailments are not affected.

The three-way optionality is possible with verbs from a range of semantic classes. Verbs of cognition such as *tahu* ‘know’ and *berpikir* ‘think’ are illustrated in (25-26) above. The same alternation is possible for verbs of communication, such as *melapor* ‘report’ and *mengaku* ‘confess,’ as well as verbs of response such as *menolak* ‘deny.’

- (27) Saya belum me-lapor { bahwa / kalau / ∅ } saya akan meng-ikut-i kegiatan itu.
 1s not.yet AV-report C 1s will AV-join-Appl activity that
 ‘I did not yet report that I would join the activities.’
- (28) Dia meng-aku { bahwa / kalau / ∅ } ajaran yang mereka anut ber-diri sendiri tanpa
 3s AV-claim C lesson C.Foc 3p follow MV-stand alone without
 ber-sandar dari enam agama yang telah di-akui negara.
 MV-lean from six religion C.Foc Perf PV-recognize country
 ‘He claims that the teachings that they follow stand on their own without depending on the six religions recognized by the state.’ (“Berkembang di Mojokerto Tahun 1980, Kini Punya 200 Penganut,” Jawa Pos Online, November 10, 2017)
- (29) Zidane men-olak { bahwa / kalau / ∅ } dirinya sengaja meny-ingkir-kan James.
 Zidane AV-reject C self intentional AV-discard-Appl James
 ‘Zidane denied that he intentionally got rid of James.’ (“James Rodriguez Tak Diizinkan Hengkang Oleh Zidane,” Jawa Pos Online, August 2, 2016)

(28-29) are from written texts originally occurring with the formal complementizer *bahwa*. In prescriptive written language, the complementizer *kalau* is avoided, but these sentences are accepted with *kalau* or null C when spoken. Clauses embedded under factives also allow the three surface

realizations of the complementizer:

- (30) Saya meny-esal { bahwa / kalau / \emptyset } saya tidak datang kemarin.
 1s AV-regret C 1s Neg come yesterday
 ‘I regret that I didn’t come yesterday.’
- (31) Ibu Melly ingat { bahwa / kalau / \emptyset } Ali lulus ujian.
 Mrs Melly remember C Ali pass exam
 ‘Mrs Melly remembered that Ali passed the exam.’
- (32) Mereka tidak ingat { bahwa / kalau / \emptyset } di negara orang itu ada undang-undang
 3p Neg remember C at country person that exist laws
 atau-pun suatu peraturan yang harus di-ketahui.
 or-PUN some rule C.Foc must PV-know
 ‘They didn’t remember that in a foreign country there are laws or rules that should be
 known.’ (“Hendak membiayai anak, seorang TKI justru diperdagangkan,” BBC Indonesia
 Online, March 7, 2017.)

These examples show that the presupposition required by factive verbs is not affected by the form of C that it embeds. I conclude that the realization of C is not affected by the semantic class of verb.

Besides active verbs, other constructions that embed declarative clauses allow optionality among *bahwa*, *kalau* and null C. Passive verbs behave as their active counterparts do:

- (33) Sudah di-ketahui { bahwa / kalau / \emptyset } Tina suka kue coklat.
 already PV-discover C Tina like cake chocolate
 ‘It was already discovered that Tina likes chocolate cake.’
- (34) Jangan di-kira { bahwa / kalau / \emptyset } politik tidak di-kait-kan dengan sepakbola.
 Neg.Imper PV-think C politics Neg PV-tie-Appl with football
 ‘Don’t think that politics is not connected to football.’

Nominals and adjectival predicates can also embed declarative clauses with optional C:

- (35) Susan dapat kesimpulan { bahwa / kalau / \emptyset } Tina suka kue coklat.
 Susan get conclusion C Tina like cake chocolate
 ‘Susan drew the conclusion that Tina likes chocolate cake.’
- (36) Sudah jelas { bahwa / kalau / \emptyset } Tina suka kue coklat.
 already clear C Tina like cake chocolate

‘It is clear that Tina likes chocolate cake.’

Finally, examples (27, 32, 34) show that negation does not affect the embedded complementizer. In all cases, the optional alternation among *bahwa*, *kalau* and null C remains.

2.2.2. Sentence-initial C

Declarative root clauses must be introduced by null C:

(37) { *Bahwa / *Kalau / \emptyset } Tina suka kue coklat.
C Tina like cake chocolate
‘Tina likes cake.’

(38) { *Bahwa / *Kalau / \emptyset } saya meny-esal Ali menang.
C 1s AV-regret Ali win
‘I regret Ali won.’

Sentential subjects such as (39) are possible for some, but not all, Indonesian speakers.⁶ For speakers who allow sentential subjects, the embedded clause must begin with *bahwa* or *kalau*:

(39) { Bahwa / Kalau / * \emptyset } Tina suka kue coklat, sudah jelas.
C Tina like cake chocolate already clear
‘That Tina likes cake, is clear.’

However, the embedded clause in subject position may not be extracted with *yang*:

(40) *{ Bahwa / Kalau / \emptyset } Tina suka kue yang sudah jelas.
C Tina like cake C.Foc already clear
(‘It is (the fact) that Tina likes cake that is clear.’)

In general, clauses cannot be extracted or clefted as in (40). This falls under the generalization that in Indonesian, only DPs may be extracted or clefted.

⁶Consultants from East Java did not accept sentential subjects in either standard or colloquial Indonesian, whereas speakers of Jakartan Indonesian found (39) to be grammatical.

2.3. Loss of optionality: A-bar movement and questions

The alternation among *bahwa*, *kalau* and null C is not completely free. We have already seen that in root (matrix) clauses, C must be null. Two further environments place constraints on the realization of C, and disallow the optionality that is usually observed in embedded declaratives. The two environments are: (1) nominal A-bar movement over C, which results in an obligatory form of C; and (2) *wh* in situ questions, which are incompatible with *bahwa* even though there is no overt movement of the *wh* phrase over C.

I begin with background discussion that introduces Indonesian A-bar movement (Sections 2.3.1 – 2.3.3) and strategies for question formation (Sections 2.3.4 – 2.3.5). I then show that A-bar movement results in *wh*-agreement on C (Section 2.3.6) and that interrogative clauses also affect the form of C (Section 2.3.7).

2.3.1. *Properties of A-bar movement in Indonesian*

A-bar movement is a well-studied topic in Indonesian (and its regional varieties, as well as in Malay). In this section I summarize patterns that have previously been discussed in Saddy 1991; Cole and Hermon 1998; Soh 1998; Voskuil 2000; Cole and Hermon 2005; Fortin 2006, 2009; Aldridge 2008; Arka and Manning 2008; Cole et al. 2008b; Sato 2008b, 2012; Yanti 2010.

Indonesian exhibits properties that are typical of A-bar movement cross-linguistically, including leftward displacement accompanied by a gap; long distance dependencies; sensitivity to syntactic islands; and crossover effects. A nominal that has undergone A-bar movement occurs at the left edge of the clause and leaves a gap. This is illustrated below with both constituent *wh* questions and relative clauses.

- (41) Tina suka kue.
Tina like cake
'Tina likes cake.'
- (42) Siapa yang ___ suka kue?
who C.Foc like cake

'Who likes cake?'

- (43) orang yang ___ suka kue
person C.Foc like cake
'the person that likes cake'
- (44) Lani bel-ajar matematika.
Lani MV-study mathematics
'Lani studies mathematics.'
- (45) Apa yang Lani bel-ajar ___ ?
what C.Foc Lani MV-study
'What does Lani study?'
- (46) mata pelajaran yang Lani bel-ajar ___
subject C.Foc Lani MV-study
'the subject that Lani studies'

This movement can be long distance, i.e. can cross more than one clause boundary.

- (47) Susan tahu Tina suka kue coklat.
Susan know Tina like cake chocolate
'Susan knows Tina likes chocolate cake.'
- (48) Apa yang Susan tahu Tina suka ___ ?
what C.Foc Susan know Tina like
'What does Susan know that Tina likes?'
- (49) Kamu pikir Billy beli sepeda motor baru.
2s think Billy buy motorbike new
'You thought Billy bought a new motorbike.'
- (50) Sepeda motor yang kamu pikir Billy beli ___ .
motorbike C.Foc 2s think Billy buy
'the motorbike that you thought Billy bought.'

While A-bar extraction over multiple clause boundaries is possible, it rarely occurs in natural speech. While extraction out of one embedded clause is unexceptional, extraction over multiple active verbs is usually judged less than natural, though not impossible. Long distance extraction seems more likely to be accepted if only one clause boundary is crossed; if the matrix subject is a 1 or 2 personal pronoun; or if the matrix verb is a high-frequency verb such as *tahu* 'know' or *bilang*

‘say’ (informal). Where examples of long distance extraction are given (especially to show relevant morphology on complementizers and/or verbs), my consultants have accepted them as possible, but less preferred than alternate constructions, such as *wh*-in situ questions.⁷

The displacement of nominals is sensitive to syntactic islands, from which movement is not possible. Complex NPs and relatives are islands for movement:

- (51) Susan dapat kesimpulan { bahwa / kalau / \emptyset } Tina suka kue.
 Susan get conclusion C Tina like cake
 ‘Susan drew the conclusion that Tina likes cake.’
- (52) *Siapa yang Susan dapat kesimpulan { bahwa / kalau / \emptyset } ___ suka kue?
 who C.Foc Susan get conclusion C like cake?
 (‘Who did Susan draw the conclusion likes cake?’)
- (53) *Apa yang Susan dapat kesimpulan { bahwa / kalau / \emptyset } Tina suka ___ ?
 what C.Foc Susan get conclusion C Tina like
 (‘What did Susan draw the conclusion that Tina likes?’)
- (54) Kamu suka cerita yang meng-eritik siapa itu?
 2s like story C.Foc AV-criticize who that
 ‘You like stories that criticize who?’ (modified from Saddy 1991:190)
- (55) *Siapa yang kamu suka cerita yang keritik ___ ?
 who C.Foc 2s like story C.Foc criticize
 (‘Who do you like stories that criticize?’) (modified from Saddy 1991:190)

A-bar movement out of adjunct clauses and PPs is likewise impossible:

- (56) Ayah senang ketika me-lihat film itu.
 Father happy when AV-see film that
 ‘Father was happy when he saw that film.’
- (57) *Apa yang Ayah senang ketika lihat ___ ?
 what C.Foc Father happy when see
 (‘What was Father happy when he saw?’)
- (58) Aku marah karena dia me-rusak proyek kami.
 1s angry because 3s AV-ruin project 1s.Excl
 ‘I was angry because he ruined our project.’

⁷Many of the long-distance extraction data presented in Saddy 1991 are impossible for my consultants, for the reasons mentioned here. I have attempted to replace these with alternate examples accepted by my consultants.

- (59) *proyek yang aku marah karena dia rusak ____
 project C.Foc 1s angry because 3s ruin
 ('the project that I was angry because he ruined')
- (60) Aku mau ber-temu dengan Ibu guru.
 1s want MV-meet with Mrs teacher
 'I will meet with the teacher.'
- (61) *Siapa yang aku mau ber-temu dengan ____ ?
 who C.Foc 1s want MV-meet with
 ('Who will I meet with?')

Crossover effects occur when a nominal undergoes A-bar movement over a coindexed pronoun. Both strong crossover (62-63) and weak crossover (64-65) are observed in Indonesian:

- (62) *Siapa₁ yang dia₁ kira akan ____ di-pecat?
 who C.Foc 3s think will PV-fire
 ('Who₁ does he₁ think will be fired?')
- (63) *Siapa₁ yang mereka₁ diskusi-kan ____ ?
 who C.Foc 3p discuss-Appl
 ('Who₁ are they₁ discussing?')
- (64) *Siapa₁ yang ibu-nya sayangi ____ ?
 who C.Foc mother-D love
 'Who₁ does his₁ mother love?'
- (65) *Siapa₁ yang dosen mereka hargai ____ ?
 who C.Foc professor 3p appreciate
 'Who₁ does their₁ professor appreciate?'

In addition to exhibiting properties expected of general A-bar movement, two additional properties are specific to Indonesian (and similar languages of the area). First, nominal A-bar movement over an active verb has consequences for verbal morphology. The basic generalization is that movement from the complement of a transitive active verb (i.e. a verb that can take the prefix *meN-*) is only possible when the verb does not bear the voice prefix:

- (66) Aku meny-impan rahasia.
 1s AV-keep secret
 'I'm keeping a secret.'

- (67) *rahasia yang aku meny-impan ____ .
secret C.Foc 1s AV-keep
'the secret I'm keeping'
- (68) rahasia yang aku simpan ____
secret C.Foc 1s keep
'the secret I'm keeping'
- (69) Siapa yang Bill kira Tom harap Fred cintai ____ ?
who C.Foc Bill think Tom expect Fred love
'Who does Bill think Tom expects Fred loves?' (adapted from Saddy 1991:187)
- (70) *Siapa yang Bill meng-ira Tom meng-harap Fred men-cintai ____ ?
who C.Foc Bill AV-think Tom AV-expect Fred AV-love
'Who does Bill think Tom expects Fred loves?' (adapted from Saddy 1991:187)

The movement of the object over the active verb requires a “bare” verb, i.e. without voice morphology. This pattern has already been much discussed in the literature, and is summarized as follows:

(71) Cole and Hermon's Generalization

The obligatory omission of *meng-* with verbs that would otherwise permit *meng-* indicates the movement of an NP argument over the *meng-* + verb. (Sato 2008b, citing Cole and Hermon 1998:231)

This formulation does not specify whether both A and A-bar movement are implicated; I take the generalization to apply only to A-bar movement.⁸

The second language-specific property of A-bar movement is that extraction always results in a relative or pseudo-cleft structure; this is discussed in the next section.

2.3.2. Pseudo-cleft structure

In this section I look at the structure of pseudo-clefts in Indonesian. All cases of nominal extraction (including constituent wh questions, relatives and declarative clauses with a clefted nominal) result in a pseudo-cleft structure.⁹ Relatives and pseudo-clefts employ the same surface structure; for

⁸Other authors, however, believe that A movement also requires omission of the prefix *meN-*. I argue against this view in Chapter 3.

⁹Pseudo-clefts have a clefted interpretation, i.e. “It is X that...” but are not formed with an overt expletive and copula.

example, (68) can be interpreted as ‘the secret I’m keeping’ or ‘it is a secret that I’m keeping’ (even though the distinction is usually made clear by context and intonation).

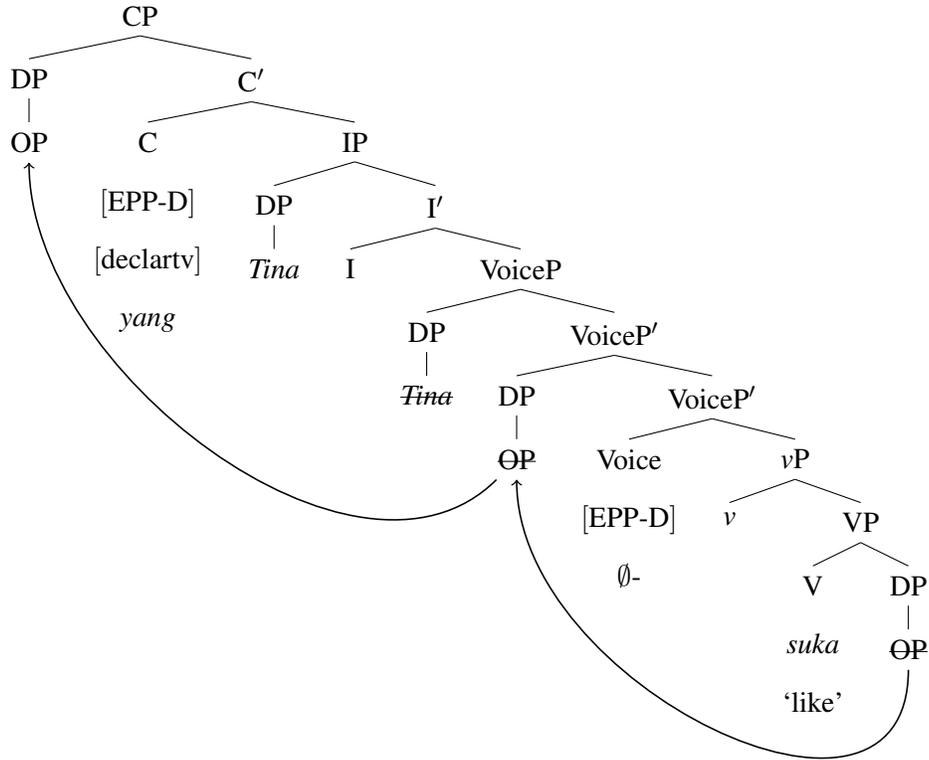
I illustrate pseudo-cleft structure in Indonesian using the object question in (72):

- (72) Apa [yang Tina suka] ?
what C.Foc Tina like
‘What does Tina like?’

The head of the relative/pseudo-cleft occurs at the left periphery, separated from the rest of the clause by the morpheme *yang*. The position of the head nominal is clearly an A-bar position, higher than the grammatical subject of the clause. Clefted arguments may either be definite or indefinite, whereas subjects must be definite/specific in Indonesian.

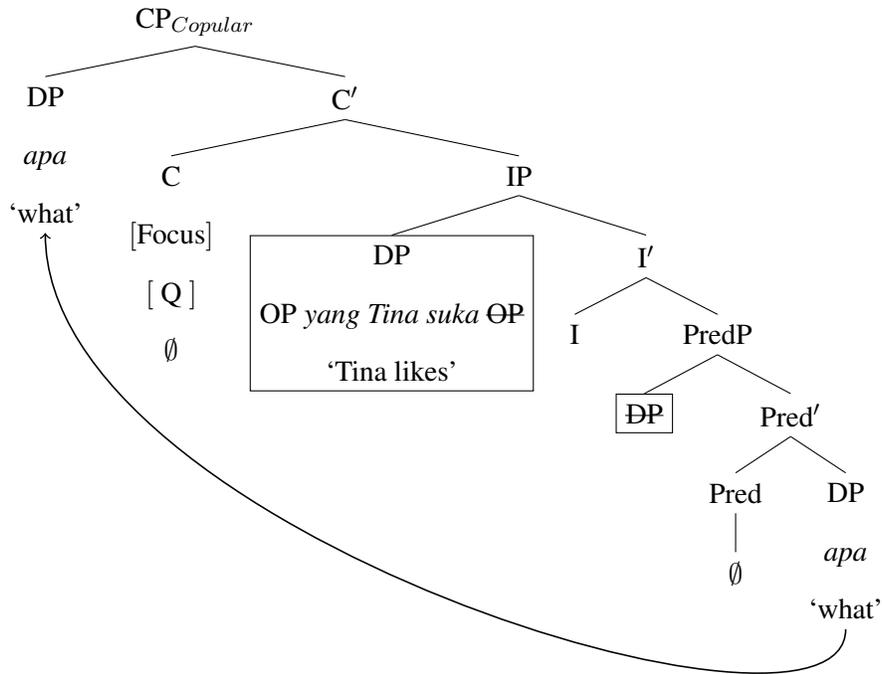
Previous authors have concluded that the nominal head of the cleft is a predicate that has raised to a focused position (see Kader 1976; Cole and Hermon 1998; Cole et al. 1999; Cole et al. 2005; Kroeger 2009 for Malay and Indonesian; also see Paul 2001; Pearson 2001; Massam 2003 for discussion on pseudo-clefts in other Austronesian languages). First, a null Operator undergoes relative-internal movement to form a headless relative:

- (73) Relative-internal Operator movement



Relative-internal movement of OP is motivated by a D-feature on phase heads C and Voice, which targets only DPs for movement. This feature on a phase head forces movement of the nominal to its specifier, so I have labeled the feature [EPP-D]. This headless relative is interpreted as ‘(the thing/the one) that Tina likes.’ This headless relative, *yang Tina suka*, is the subject of the copular clause in (74), and *yang* is embedded in the copular clause (cf. Cole et al. 1999). I assume the headless relative is an IP embedded under a DP projection when it is generated as the subject, since the headless relative is always interpreted as a definite; furthermore, sentential subjects are not possible for my consultants. The headless relative subject begins as the external argument of the predicate in (74), and is raised to subject position in SpecIP.

(74) Matrix copular clause of pseudo-cleft



The nominal that surfaces as the head of the relative or cleft is generated low, as predicate of the copular clause. This nominal is raised to a focused position at the left edge of the clause, which derives the word order in (72). The structure in (74) shows the predicate nominal in both positions (base position and raised to the left periphery). Note that in a cleft structure, matrix C is null, while *yang* realizes the embedded C in the subject. This analysis is similar to that proposed by Cole, Hermon and Tjung (2005), who also argue that (Standard) Indonesian clefts have the form [NP NP] rather than [NP VP].

This structure in (74) is supported by several pieces of evidence in Indonesian. First, the head of the cleft can occur both in sentence-initial position and in post-verbal position, illustrated in the following examples:

- (75) Siapa yang suka kue coklat?
 who C.Foc like cake chocolate
 'Who is it that likes chocolate cake?'

- (76) Yang suka kue coklat siapa?
C.Foc like cake chocolate who
'(The one) that likes chocolate cake, is who?'
- (77) Akibat kesalahan kamu sendiri yang kamu rasakan.
consequence mistake 2s self C.Foc 2s feel
'It is the consequence of your own mistake that you feel.'
- (78) Yang kamu rasakan akibat kesalahan kamu sendiri.
C.Foc 2s feel consequence mistake 2s self
'(The thing) that you feel, is the consequence of your own mistake.'

In (76) and (78) the clefted nominal may occur low, which is its base position as predicate of the matrix copular clause. Another argument for the predicative status of the clefted nominal comes from the distribution of the question particle *-kah* that focuses an element in polar questions. This particle may occur within the predicate, but cannot occur on subjects (this is true for both Malay and Indonesian; see Kader 1976; Cole et al. 1999; Musgrave 2001; Paul 2001). *-Kah* occurs on verbal predicates (79) as well as copular predicates (81). In both cases, *-kah* may not attach to the subject of the clause.

- (79) Mariam me-mukul dokter itu **-kah** tadi?
Mariam AV-hit doctor that Q.Foc just.now
'Was it *the doctor* that Mariam beat just now?'
- (80) *Mariam **-kah** me-mukul dokter itu tadi?
Mariam Q.Foc AV-hit doctor that just.now
'(Was it *Mariam* that beat the doctor just now?') (modified from Musgrave 2001:75)
- (81) Ali rasa Rahman se-orang yang baik **-kah**?
Ali feel Rahman one-person C.Foc good-Q.Foc
'Does Ali feel that Rahman is a *good* person?'
- (82) *Ali rasa Rahman **-kah** se-orang yang baik]?
Ali feel Rahman-Q.Foc one-person C.Foc good
'(Does Ali feel that *Rahman* is a good person?') (modified from Cole, Hermon and Aman 1999:18)

I apply this diagnostic to (76) and (78). If the clefted nominal begins as predicate, it should be able to occur with the question particle *-kah*, in either position. This is the case:

- (83) Tina-**kah** [yang suka kue coklat]?
 Tina-Q.Foc C.Foc like cake chocolate
 ‘Is it *Tina* that likes chocolate cake?’
- (84) [Yang suka kue coklat] Tina-**kah**?
 C.Foc like cake chocolate Tina-Q.Foc
 ‘The one that likes chocolate cake, is it *Tina*?’
- (85) Akibat kesalahan kamu sendiri-**kah** [yang kamu rasakan]?
 consequence mistake 2s self-Q.Foc C.Foc 2s feel
 ‘Is it *the consequence of your own mistake* that you feel?’
- (86) [Yang kamu rasakan] akibat kesalahan kamu sendiri-**kah**?
 C.Foc 2s feel consequence mistake 2s self-Q.Foc
 ‘(The thing) that you feel, is it *the consequence of your own mistake*?’

The head of the cleft, then, patterns as a predicate that can be focused with *-kah*. The headless relative cannot take *-kah*, patterning like a subject. A similar pattern also holds for the emphatic particle *-lah*, which occurs within predicates but not subjects (Sneddon 1996; Kroeger 2009). In the declarative sentences below, *-lah* can occur on the clefted nominal but not on the relative:

- (87) [Yang suka kue coklat] orang itu -**lah**.
 C.Foc like cake chocolate person that Emph
 ‘The one that likes chocolate cake, is *that person*.’
- (88) *[Yang suka kue coklat -**lah**] orang itu.
 C.Foc like cake chocolate Emph person that
 ‘(The one that likes chocolate cake is *that person*.)’

Negation also supports the status of the clefted nominal as a predicate. Verbal and adjectival predicates are negated by the morpheme *tidak*, while nominal predicates are negated by *bukan* (Sneddon 1996; Kroeger 2009; Sneddon et al. 2012).¹⁰

- (89) Ibu tidak sakit hari ini.
 Mother Neg sick day this
 ‘Mother is not sick today.’

¹⁰This is a simplification of the distinction between *tidak* and *bukan*; see Kroeger 2014 for further discussion of the two negation morphemes, which does not bear on the current analysis.

- (90) Ali tidak lulus ujian.
 Ali Neg pass exam
 ‘Ali did not pass the exam.’
- (91) Dia bukan pemenang sebenarnya.
 3s Neg winner really
 ‘He is not the winner really.’

Consistent with other predicate nominals, the clefted nominal can be negated with *bukan*, but not with *tidak*:

- (92) { Bukan / *tidak } dia yang ku cinta ...
 Neg 3s C.Foc 1s love
 ‘It’s not her that I love...’
- (93) { Bukan / *tidak } kau yang patut minta maaf.
 Neg 2s C.Foc should request forgive
 ‘It’s not you who should apologize.’ (modified from Kroeger 2009:822)

I hereafter assume the structure in (74) for clefts, including all questions in which the wh phrase has undergone movement. I later return to this structure in Section 2.3.7 to derive moved-wh questions.

I also assume that null operator movement is the strategy used for all cases of extraction, including wh constituent questions, relatives and pseudo-clefts. In other words, the overt nominal that occurs as head of the relative/cleft does not undergo movement within the relative clause. It is a null operator that undergoes cyclic A-bar movement within a relative clause, rather than the DP that is pronounced. Other authors have also treated null operator movement as the strategy for wh movement or A-bar movement in various languages, assuming that surface displacement obtains through movement of the operator, e.g. Chung 1998; Reintges et al. 2006; McCloskey 2001, 2002. All cases of nominal A-bar movement or wh movement, including relative formation, are cases of null operator movement. For ease of exposition, where the distinction does not bear upon the discussion, I sometimes refer to the syntactic operation simply as nominal movement or A-bar extraction, abstracting away from the movement of the operator versus the interpretation of the overt nominal in the place of the moved operator.

2.3.3. Nominal movement vs. non-nominal movement

The discussion that follows is specifically focused on nominal A-bar movement. Indonesian and related languages differentiate nominal extraction from the movement of other categories, as noted by many authors (see e.g. Cole and Hermon 1998; Sato 2008a,b). For example, PPs and other adjuncts may occur in various positions, but cannot occur with the morpheme *yang*, which I take to be required in all cases of A-bar movement:

- (94) Piala Dunia akan di-ada-kan pada tahun 2018 di Rusia.
cup world will PV-exist-Appl at year 2018 at Russia
'The World Cup will be held in 2018 in Russia.'
- (95) Pada tahun 2018 (*yang) Piala Dunia akan di-ada-kan di Rusia.
at year 2018 C.Foc cup world will PV-exist-Appl at Russia
'In 2018 the World Cup will be held in Russia.'
- (96) Tono meng-erja-kan tugasnya dengan semangat.
Tono AV-work-Appl task with enthusiastic
'Tono worked on the task enthusiastically.'
- (97) Dengan semangat (*yang) Tono meng-erja-kan tugas-nya.
with enthusiastic C.Foc Tono AV-work-Appl task-D.
'Enthusiastically Tono worked on the task.'

Unlike DP movement, when non-nominal constituents cross over the verb, the movement does not affect active verbal morphology, e.g. in (97) the verb retains its active prefix (*meN-*). Wh phrases also show a distinction between nominals and non-nominals. Adverbial and PP wh phrases occurring in a sentence-initial position can occur with the active verbal prefix as in (98) and (99):

- (98) Bagaimana Ali me-mandu kereta?
how Ali AV-drive car
'How did Ali drive the car?' (Cole and Hermon 1998:226)
- (99) Di mana Ali mem-beli pangsapuri?
at where Ali AV-buy condominium
'Where did Ali buy a condominium?' (Cole and Hermon 1998:226)

I do not attempt to explain the puzzling distinction between nominal vs. non-nominal extraction

here. The discussion focuses on nominal movement throughout this chapter, excluding the movement of other types of constituents such as PPs and adverbs.

2.3.4. *Three strategies for constituent wh questions*

Wh constituent questions may be formed in one of three ways in Indonesian: in situ wh questions, sentence-initial wh questions and partial wh questions (partial wh movement) (Saddy 1991; see also Cole and Hermon 1998). In Indonesian, wh-in situ can be used as a clarification question, but wh-in situ constructions are unremarkable as matrix questions and occur very frequently. Note however that when subjects are questioned, wh in situ is generally not possible. This is due to a requirement that grammatical subjects in Indonesian be definite or specific, which rules out wh phrases as subjects (Sneddon et al. 2012).^{11,12}

Wh-in situ questions and sentence-initial wh questions are illustrated below, with the wh originating in different argument positions, including grammatical subject, object, ditransitive object and possessor. A sentence-initial wh phrase occurs in a scopal position, i.e. the matrix clause is interpreted as an interrogative. The interpretation of a wh-in situ question is equivalent to a sentence-initial wh question: the sentence has interrogative force and the expected response is that of a constituent wh question.

(100) Sentence-initial wh (subject)

Siapa yang ___ meny-uka-i kue coklat?
who C.Foc AV-like-Appl cake chocolate
'Who likes chocolate cake?'

(101) Wh-in situ (object)

Billy beli makanan apa?
Billy buy food what
'What food did Billy buy?'

(102) Sentence-initial wh (object)

¹¹There appears to be variation in this requirement in some varieties of Indonesian, e.g. see Sneddon 2006. For some speakers, 'who' subject questions are possible just in case the questioned subject is among the addressees, i.e. 'who among you..?' I leave this issue aside as an area that requires further research.

¹²This is true of the grammatical (surface) subject, as well as the logical subject that remains in its thematic position in Object voice clauses. See Guilfoyle et al. 1992; Cole et al. 2008b.

Makanan apa yang Billy beli ___?
food what C.Foc Billy buy
'What food did Billy buy?'

- (103) Wh-in situ (ditransitive, direct object)

Kamu mem-beli-kan ibu apa?
2s AV-buy-Appl mother what
'What did you buy mother?'

- (104) Sentence-initial wh (ditransitive, direct object)¹³

Apa yang kamu beli-kan ___ untuk ibu?
what C.Foc 2s buy-Appl for mother
'What did you buy for mother?'

- (105) Wh-in situ (ditransitive, indirect object)

Kamu mem-beli-kan siapa bunga?
2s AV-buy-Appl who flower
'Who did you buy flowers (for)?'

- (106) Sentence-initial wh (ditransitive, indirect object)

Siapa yang kamu beli-kan ___ bunga?
who C.Foc 2s buy-Appl flower
'Who did you buy flowers (for)?'

- (107) Wh-in situ (possessor)

Adik mem-baca buku siapa?
younger.sibling AV-read book who
'Who is it that little brother is reading (their) book?'

- (108) Sentence-initial wh (possessor)

Siapa yang adik baca buku-nya ___?
who C.Foc younger.sibling read book-D
'Who is it that little brother read (his) book?'

Wh-in situ questions are also possible for questioning a prepositional object or adjunct clause:

- (109) Wh-in situ (prepositional object)

¹³Note that when the direct object is questioned in a moved-wh sentence, the indirect object must occur in a PP. Some authors (e.g. Sato 2012; see examples 18b, 19b) report that it is possible to extract the direct object over an indirect object occurring without a preposition. This is not possible in the variety of colloquial Indonesian spoken by my consultants from East Java or Jakarta.

Kamu marah sama siapa?
2s angry with who
'Who are you angry with?'

(110) Wh-in situ (prepositional object, possessor)

Mereka pulang dari rumah siapa?
3s go.home from house who
'Whose house did they go home from?'

(111) Wh-in situ (adjunct clause)

Ayah senang ketika me-lihat apa?
Father happy when AV-see film what
'What was Father happy when he saw?'

(112) Wh-in situ (adjunct clause)

Dia marah karena siapa yang me-rusak proyek kami?
3s angry because who C.Foc AV-ruin project 1s.Excl
'Who was he was angry because (he) ruined our project?'

However, A-bar movement out of PPs and adjuncts is disallowed (as discussed in Section 2.3.1), so sentence-initial wh questions and partial wh questions cannot be formed from these in situ questions.

Moved wh questions have a slightly different interpretation than the wh-in situ questions, since the moved wh results in a pseudo-cleft. The information in the relative structure is presupposed, while the head of the cleft is new information (see discussion of pseudo-clefts in 2.3.1). The wh-in situ question in (101) does not carry a presupposition, while in the clefted question in (102), the information within the relative clause is presupposed as old information, 'Tono knows that Billy bought x.'

In partial wh movement, the wh word moves to an initial position within an embedded clause. As originally observed by Saddy (1991), partial wh questions such as (113) and (114) have two interpretations: a matrix question reading, and an embedded reading with an indirect question. In embedded readings, the matrix clause does not have interrogative force.

(113) Partial wh (subject)

Susan tahu siapa yang — meny-uka-i kue coklat
 Susan know who C.Foc AV-like-Appl cake chocolate
 Matrix: ‘Who does Susan know likes chocolate cake?’
 Indirect: ‘Susan knows who it is that likes chocolate cake.’

(114) Partial wh (object)

Tono tahu apa yang Billy beli —
 Tono know what C.Foc Billy buy
 Matrix: ‘What does Tono know that Billy bought?’
 Indirect: ‘Tono knows what it is that Billy bought.’

(115) Partial wh (ditransitive)

Lani tahu siapa yang kamu beli-kan — bunga.
 Lani know who C.Foc 2s buy-Appl flower
 Matrix: ‘Who does Lani know you bought flowers (for)?’
 Indirect: ‘Lani knows who it is you bought flowers (for).’

As previously discussed, when the wh word is moved it is followed by the relative morpheme *yang*. Partial wh movement shows that *yang* marks the surface position of the clefted nominal, which means that *yang* does not always mark question scope: partial wh questions can be interpreted as matrix questions, in which the wh phrase is interpreted as having widest scope over the sentence.

2.3.5. Strategies for forming polar questions

Polar questions may be formed in several ways. Question intonation may be used with a declarative sentence to form a yes-no question as in (116) and (117).

(116) Ibu mau pulang sekarang? (with question intonation)

Mrs want go.home now
 ‘Are you going home now?’

(117) Dia sakit? (with question intonation)

3s sick
 ‘Is she sick?’

A clause with declarative word order may optionally begin with the question word *apa*, optionally cliticized with the question particle *-kah*. No semantic difference arises between (117) and (119).

- (118) Apa(-kah) anak saya perlu latihan lagi?
 what-Q child 1s need practice again
 ‘Does my child need more practice?’
- (119) Apa(-kah) dia sakit?
 what-Q 3s sick
 ‘Is she sick?’

Another strategy for forming a polar question is to focus the questioned element, which occurs in a sentence-initial position, cliticized with the question particle *-kah* (Fortin 2007). *-Kah* may attach to different categories:

- (120) Sakit-kah dia?
 sick-Q.Foc 3s
 ‘Is she sick?’ (modified from Fortin 2007:54)
- (121) Terlalu gemuk-kah dia?
 too chubby-Q.Foc 3s
 ‘Is he too chubby?’ (modified from Fortin 2007:54)
- (122) Cukup minum-kah kita?
 enough drink-Q.Foc 1p.Incl
 ‘Have we drunk enough?’
- (123) Untuk itu-kah aku datang?
 for that-Q.Foc 1s come
 ‘For that, I came?’

The clitic *-kah* can also attach to nominal arguments. Unlike other categories cliticized with *-kah*, nominal constituents are always followed by *yang*, consistent with previous examples of A-bar movement that we have already seen.

- (124) Pohon ini-kah yang ___ akan di-tebang?
 tree this-Q.Foc C.Foc will PV-cut.down
 ‘Is it this tree that will be cut down?’ (modified from Fortin 2007:55)
- (125) (Apa-kah) orang itu-kah yang sesungguhnya ___ me-mimpin perusahaan?
 what-Q person that-Q.Foc C.Foc truly AV-lead company
 ‘Is it that person who truly leads the company?’

Note that when the clitic *-kah* occurs on a focused element, it may co-occur with the question word *apa(-kah)*, as in (125). However, *apakah* is compatible with polar questions only, and cannot be combined with in situ constituent wh questions:

- (126) (*Apa-kah) Kamu sedang baca apa?
 what-Q 2s Prog read what
 ‘What are you reading?’
- (127) (*Apa-kah) Pertunjukkan bunga api akan di-ada-kan di mana?
 what-Q show fireworks will PV-exist-Appl at which
 ‘Where will the fireworks show be held?’

When a polar question is embedded under a predicate that selects an interrogative clause, the question is either introduced by *apa(-kah)* or by the complementizer *kalau*.

- (128) Dia tanya { apa-kah / kalau } anak-nya perlu latihan lagi.
 3s ask what-Q C child-D need practice again
 ‘She asked whether/if her child needs more practice.’
- (129) Kami ragu-ragu { apa / kalau } dia bisa lulus ujian.
 1p.Excl doubtful what-Q C 3s can pass exam
 ‘We are doubtful whether/if he can pass the exam.’ (modified from Sneddon 1996:323)

In Section 2.3.7, I return to the fact that *kalau* can introduce embedded polar questions as in (129), as embedded or indirect questions are compatible with *kalau*, but not *bahwa*.

2.3.6. Complementizers and A-bar movement

Thus far we have seen examples illustrating nominal A-bar movement, and different strategies for question formation in Indonesian; these place distinct constraints on the realization of complementizers. Since A-bar movement subsumes some types of questions (i.e. questions in which a nominal wh phrase undergoes A-bar movement to an intermediate or sentence-initial position), I begin by looking at the effects of general A-bar movement on complementizers. In Section 2.3.7, I continue on to a discussion of complementizers in questions.

Saddy (1991) first showed that long distance nominal movement in Indonesian affected not only verbal morphology, but also complementizers. Saddy's generalizations were also taken up by Cole and Hermon (1998) in research on wh questions in the related variety, Malay; other authors who discuss Indonesian complementizers include Sneddon (1996); Sato (2008b); Fortin (2009); Yanti (2010); Sneddon et al. (2012).

Constructions that involve nominal A-bar movement include wh constituent questions as well as the formation of relatives and clefts. Recall that embedded clauses without nominal movement allow for variable realization of the complementizer as *bahwa*, *kalau* or null C. Saddy (1991) observed that the overt complementizer *bahwa* is ruled out whenever long distance A-bar movement crosses C, and the complementizer must be null instead; I extend this generalization to include *bahwa*. I argue below that whenever A-bar movement occurs, two obligatory forms of C are required: *yang* is the form required for C whenever the moved nominal (i.e. a null Operator) lands in its specifier, whereas intermediate CPs, through which cyclic movement passes, require null C.

As mentioned in Section 2.3.1, when a nominal undergoes A-bar movement over an active verb (i.e. a verb which bears the prefix *meN-*), the verb must occur without the active prefix. I take this to be an instance of wh-agreement, and I briefly review this pattern since I use verbal wh-agreement to diagnose A-bar movement throughout this discussion. In (130), without A-bar movement, the active prefix *meN-* may occur on both matrix and embedded verbs.¹⁴ In (131) *apa* has moved to an intermediate position, at the left edge of the embedded clause, and has only crossed the embedded verb. The embedded verb must be bare, while the matrix verb may bear the active prefix *meN-*:^{15,16}

- (130) Kamu { kira / meng-ira } Lani sudah { beli / mem-beli } apa?
 2s think AV-think Lani already buy AV-buy what
 'What do you think Lani bought?'

¹⁴Here I do not address the optional realization of the active prefix in examples such as (130); see Chapter 3 for discussion of optional Voice morphology.

¹⁵The verb *kira* is closer in meaning to English *guess*. Declarative counterparts to these clauses often occur with *pikir*, 'think' instead of *kira*.

¹⁶The verb *mengira* is considered formal with the active prefix. My consultants preferred the form *kira* for these examples, noting that *mengira* is grammatical but pragmatically unusual in spontaneous speech. I use *mengira* to illustrate the interaction of A-bar movement and the active voice prefix.

- (131) Kamu { kira / meng-ira } apa yang Lani sudah { beli / *mem-beli } ___?
 2s think AV-think what C.Foc Lani already buy AV-buy
 ‘What do you think Lani bought?’
- (132) Apa yang kamu { kira / *meng-ira } Lani sudah { beli / *mem-beli ___ }?
 what C.Foc 2s think AV-think Lani already buy AV-buy
 ‘What do you think Lani bought?’

When *apa* has crossed both verbs in (132), both verbs must be bare, without *meN-*. (Also see similar examples in 66-70.) This is not specific to *wh* elements or questions, but a general property of nominal A-bar movement. Similar effects are seen in relatives and clefts:

- (133) sepeda motor yang kamu { kira / *meng-ira } Lani sudah { beli / *mem-beli } ___
 motorbike C.Foc 2s think AV-think Lani already buy AV-buy
 ‘the motorbike that you think Lani bought’
- (134) wanita yang kamu { kira / *meng-ira } ___ sudah { beli / mem-beli } sepeda motor
 woman C.Foc 2s think AV-think already buy AV-buy motorbike
 ‘the woman that you think bought a motorbike’

Extraction of the embedded object in (133) requires that both verbs be bare, but extraction from embedded subject position (134) allows the embedded verb to bear the active prefix *meN-*, since A-bar movement of *wanita* ‘woman’ has not crossed this verb.

Nominal A-bar movement has similar consequences for the realization of complementizers, a parallel pattern first noted by Saddy (1991). In (135) the alternation among *bahwa*, *kalau* and null C is optional. When the embedded object undergoes A-bar movement in (136) however, the null complementizer is required:¹⁷

- (135) Aku kira { kalau / bahwa / ∅ } Lani sudah mem-beli beras.
 1s think C Lani already AV-buy rice
 ‘I think Lani bought rice.’
- (136) bahan yang kamu kira { *kalau / *bahwa / ∅ } Lani sudah beli ___
 ingredient C.Foc 2s think C Lani already buy
 ‘the ingredient that you think that Lani bought’

¹⁷Note that beginning with example (135), I do not detail optional alternations in verbal prefixes.

A similar pattern holds for A-bar movement from subject position. If a nominal crosses CP, an overt form of the complementizer is ruled out. In (138), A-bar movement of the matrix subject does not affect the lower complementizer. On the other hand, movement of the embedded subject over the complementizer in (139) requires null C.

- (137) Susan ber-pikir { bahwa / kalau / \emptyset } Tina meny-uka-i kue coklat
 Susan MV-think C Tina AV-like-Appl cake chocolate
 ‘Susan thinks that Tina likes chocolate cake.’
- (138) orang yang — ber-pikir { bahwa / kalau / \emptyset } Tina meny-uka-i kue coklat
 person C.Foc MV-think C Tina AV-like-Appl cake chocolate
 ‘the person that thinks that Tina likes chocolate cake’
- (139) orang yang Susan pikir { *bahwa / *kalau / \emptyset } — meny-uka-i kue coklat
 who C.Foc Susan think C AV-like-i cake chocolate
 ‘the person that Susan thinks likes chocolate cake’

I have previously noted that *yang* must appear immediately following a nominal that has undergone A-bar movement. *Yang* has frequently been identified as a relative morpheme or focus marker (Sneddon 1996; Saddy 1991; Arka 2000; Sneddon et al. 2012). The issue of focus is relevant since the head of a relative or pseudo-cleft is always focused, and *yang* must occur in these constructions. However, I argue that *yang* does not mark focus on the nominal that it follows; rather, *yang* is part of the complementizer system in Indonesian, on par with *bahwa* and *kalau* (this is also briefly suggested in Fortin 2009). First, recall that *yang* does not form a constituent with the nominal that it follows. Rather, it must precede the headless relative clause (refer to the tree structures in 73, 74), even when the clefted nominal remains in a low position (see example 78). Second, the linear position of *yang* is that expected of a complementizer, at the left periphery of a clause (matrix or embedded). *Yang* is always in complementary distribution with other complementizers (*bahwa/kalau/null C*):¹⁸

¹⁸Saddy (1991) reports that *bahwa* and *yang* can co-occur as in the example below; this must be interpreted as an embedded question.

- (1) Bill tahu **bahwa** siapa **yang** Tom cintai.
 Bill knows that who Tom loves
 ‘Bill knows who Tom loves.’ (Saddy 1991:188, example 12)

I have not found consultants that accept this construction. However, assuming that speakers of some variety of Indonesian

- (140) Susan pikir { *bahwa* / *kalau* / \emptyset / **yang* } Tina suka kue coklat.
 Susan think C Tina like cake chocolate
 ‘Susan thinks that Tina likes chocolate cake.’
- (141) Susan tahu siapa { **bahwa* / **kalau* / * \emptyset / *yang* } — suka kue coklat.
 Susan know who C like cake chocolate
 ‘Susan knows who it is that likes chocolate cake.’

Furthermore, *yang* precedes the position of in-situ grammatical subjects (see 139, and other examples throughout), which occupy the specifier of IP (Chung 1976; Guilfoyle et al. 1992; Cole and Hermon 2005; Cole et al. 2008b). Assuming that A-bar moved nominals land in a specifier of CP, it is reasonable to conclude that *yang* is the head C.¹⁹

Interestingly, in cases of long distance movement such as (139), *yang* cannot occur recursively, but only surfaces once after the position of the moved nominal. Consequently, *yang* always marks the highest CP affected by overt A-bar movement; *yang* does not necessarily mark scopal position (see Section 2.3.7 for further discussion of covert movement of wh phrases and their scopal interaction with quantifiers.) We have already seen that in Chamorro, wh-agreement is marked only on the highest C crossed by A-bar movement, whereas intermediate C retains its usual form. (see discussion surrounding example 24). Indonesian also marks the highest C with *yang*, but I argue that (137-139) instantiate a pattern that is distinct from Chamorro, because intermediate C also shows morphological wh-agreement. Note that in (137), the complementizers that occur without movement are *bahwa*, *kalau* or null C. In contrast, when this complementizer is an intermediate clause boundary crossed by long-distance movement in (137), it must be realized as null C. I take both *yang* in the matrix clause and obligatory null C in the embedded clause to be morphological

do allow *bahwa* and *yang* to co-occur, this pattern is amenable to a split-CP analysis (cf. Rizzi 1997 and subseq.) I assume that *bahwa* always appears to the left of the moved nominal, and *yang* to its right. This suggests that *bahwa* occurs in ForceP and *yang* in FocusP.

¹⁹Malay and Indonesian clearly deviate with respect to the distribution of *yang*. In Malay, *yang* can occur in the same position as *bahawa*, the Malay variant of *bahwa*. The following example is grammatical in Malay, but not in Indonesian:

- (1) Ali memberitahu saya tadi **yang** Fatima sakit semalam.
 Ali told me before that Fatimah sick yesterday
 ‘Ali told me earlier that Fatimah was sick yesterday.’ (Cole, Hermon and Aman 1999:14)

Cole et al. (1999) find that in Malay, “either the complementizer *bahawa* or the complementizer *yang* can be used to introduce complement clauses, but only *yang* can be used to introduce relative clauses” (1999:14). This is an interesting point of cross-dialectal variation that I leave for further study.

wh-agreement. This means that wh-agreement is marked on all C that are in the path of movement in Indonesian, but the highest C is morphologically distinct from intermediate C. If only one C is crossed by A-bar movement, it counts as highest C, and is realized as *yang*.

The Indonesian complementizer system, then, exhibits a pattern that extends the possible wh-agreement patterns identified by Reintges et al. 2006. The realization of C as *yang* or null C in case of A-bar movement is neither the fully recursive pattern, nor ‘the highest-C only’ pattern. Indonesian wh-agreement on highest C is morphologically distinct from intermediate C.

Complementizers in Igbo (Nigeria) also exhibit a morphological distinction between highest C and intermediate C (Amaechi and Georgi 2017).²⁰ Igbo exhibits a subject-object asymmetry that is typical of Comp-trace effects. Questioned objects and adjuncts can be extracted in Igbo, and occur with the focus marker *kà* as in (142). Questioned subjects, on the other hand, remain in situ, and do not occur with *kà* (143).

(142) Ònyé *kà* ‘Obí hùrù n’-áhíá
 who FOC Obi see P-market
 ‘Who did Obi see at the market?’

(143) Ònyé hùrù Àdà n’-áhíá
 who saw Ada P-market
 ‘Who saw Ada at the market?’ (Amaechi and Georgi 2017, ex. 6b, 6c)

However, in long distance questions in Igbo, both subjects and objects can undergo extraction, and both occur with *kà*:

(144) Úchè chère *nà* Òbí hùrù Àdà n’-áhíá
 Uche thinks that Obi saw Ada P-market
 ‘Uche thinks that Obi saw Ada at the market.’

(145) Ònyé *kà* Úchè chère *nà* Òbí hùrù ___ n’-áhíá
 who FOC Uche thinks that Obi saw P-market
 ‘Who does Uche think that Obi saw at the market?’

²⁰My presentation of the data differs from that in Amaechi and Georgi 2017; errors are my own. Thank you to Doreen Georgi for discussion about the Igbo pattern.

- (146) Ònyé kà Úchè chère (*nà) — hùrù Àdá n’-áhía
 who FOC Uche thinks that saw Ada P-market
 ‘Who does Uche think saw Ada at the market?’ (modified from Amaechi and Georgi 2017,
 ex. 35a, 35b, 35c)

The embedded clauses in (145) and (146) are marked differently. In case of long distance object extraction, a special form of the complementizer introducing the embedded clause is not required (compare 144 and 145). However, in long-distance subject extraction (146), the complementizer must be null.

Focus marking on the highest C in Igbo parallels the Indonesian complementizer system. Amaechi and Georgi (2017) analyze *kà* as the realization of C when a moved argument lands in its specifier. If this analysis for Igbo is correct, then Igbo appears to mark extraction on the highest C only. However, I suggest that Igbo complementizers can be understood as displaying three different forms in case of long distance extraction of a subject. Amaechi and Georgi show that extraction of subjects, but not objects, require that the lowest complementizer be null. This results in a Comp-trace effect, in which there is an asymmetry in marking subject extraction vs. object extraction. In cases of long distance subject extraction, then, three different forms of C are required:

- (147) Ònyé kà Helen si nà Úchè chère (*nà) — hùrù Àdá n’-áhía
 who FOC Helen say that Uche think that saw Ada P-market
 ‘Who did Helen say that Uche thinks saw Ada at the market?’ (Amaechi and Georgi, p.c.)

If a subject is extracted from the most deeply embedded clause, the Comp-trace effect requires a null form; an intermediate C is realized as the default complementizer *ná*; and the highest C is realized as *ká*. I have suggested that the ‘flavor’ of Indonesian wh-agreement on C is not predicted by Reintges et al. 2006. Additionally, Igbo displays another possible pattern of wh-agreement, in which Comp-trace effects interact with wh-agreement to yield three different forms of C in (147).

In Indonesian, movement over the verb and C are taken to be movement through the edge of syntactic phases, i.e. the specifiers of VoiceP and CP, respectively. The fact that A-bar movement has morphological consequences for the verb and the complementizer in Indonesian provides support

for long distance movement that occurs in shorter steps, via successive-cyclic movement (Chomsky 2000, 2001). Phase-cyclic analyses of nominal extraction have been implemented for Indonesian and various related languages (e.g. see Aldridge 2008; Cole et al. 2008b; Sato 2008b; Legate 2014; Yanti 2010). I take up a discussion of syntactic phases again in Chapter 4, where I argue that D is also a phase head in Indonesian.

2.3.7. Complementizers in questions

A-bar movement is not the only environment in which C cannot be optionally realized as *bahwa* or *kalau*. In this section I show that all constituent questions are incompatible with the overt complementizer *bahwa*, including in situ questions with no overt movement over C.

As already discussed, A-bar movement of a wh word over a complementizer requires null C; the overt forms *bahwa* and *kalau* cannot occur. This is illustrated with the questioned subjects below. (149) shows partial wh movement, while (150-151) show the wh subject in sentence-initial position.

- (148) Ali pikir { *bahwa* / *kalau* / \emptyset } Fatima akan di-pecat minggu depan.
 Ali think C Fatima will PV-fire week ahead
 ‘Ali thinks that Fatima will be fired next week.’
- (149) Ali pikir siapa yang ___ akan di-pecat minggu depan?
 Ali think who C.Foc will PV-fire week ahead
 ‘Who does Ali think will be fired next week?’
- (150) Siapa yang Ali pikir { **bahwa* / **kalau* / \emptyset } ___ akan di-pecat minggu depan?
 who C.Foc Ali think C will PV-fire week ahead
 ‘Who does Ali think will be fired next week?’
- (151) Siapa yang ___ pikir { *bahwa* / *kalau* / \emptyset } Fatima akan di-pecat minggu depan?
 who C.Foc think C Fatima will PV-fire week ahead
 ‘Who thinks that Fatima will be fired next week?’

Note especially that in (151), the wh phrase has been extracted from the matrix clause, but this movement has not crossed the embedded complementizer, which is optionally realized as *bahwa*/*kalau*/ \emptyset . A similar pattern holds for questioned objects in partial wh movement (153) and sentence-initial wh

(154).

- (152) Pak Dadang meng-harap { bahwa / kalau / \emptyset } mereka akan men-diskusi-kan masalah
Mr Dadang AV-expect C 3p will AV-discuss-Appl problem
pertanian.
agriculture
'Mr Dadang expects that they will discuss the issue of agriculture.'
- (153) Pak Dadang meng-harap masalah apa yang mereka akan diskusi-kan ___?
Mr Dadang AV-expect problem what C.Foc 3p will discuss-Appl
'What issue does Mr Dadang expect that they will discuss?'
- (154) Masalah apa yang Pak Dadang harap { *bahwa / *kalau / \emptyset } mereka akan
problem what C.Foc Mr Dadang expect C 3p will
diskusi-kan ___?
discuss-Appl
'What issue does Mr Dadang expect that they will discuss?'

Thus far these facts are consistent with the previous pattern that I have shown for A-bar movement. When a wh phrase is A-bar moved, morphological wh-agreement is marked on the highest C by the form *yang*, while intermediate C must be null. Any complementizer that is not crossed by A-bar movement retains its optional realization as *bahwa/kalau*/null C.

Unexpectedly however, wh-in situ questions place an additional constraint on the complementizer *bahwa*, even if no movement has crossed C. (Recall that wh phrases generally may not occur in subject position, but must be extracted. Consequently, I use non-subject questions to illustrate wh in situ.)

- (155) Mereka akan men-diskusi-kan masalah apa?
3p will AV-discuss-Appl problem what
'What issue will they discuss?'
- (156) Pak Dadang meng-harap { *bahwa / kalau / \emptyset } mereka akan men-diskusi-kan masalah
Mr Dadang AV-expect C 3p will AV-discuss-Appl problem
apa?
what
'What issue does Mr Dadang expect that they will discuss?'

(156) illustrates two points that are important to the discussion here. First, wh-in situ object ques-

tions such as (156) do not involve any overt movement over C, yet they place a requirement on the form of a complementizer in a position higher than the *wh* phrase. Second, this is the only context we have seen in which *bahwa* is ruled out, while *kalau* is possible. Thus far their syntactic distribution has been identical, with both forms possible in the same position or disallowed in the same position (excepting considerations of formality and register).

In addition to object questions such as (156), other types of *wh*-in situ also disallow *bahwa* in a higher position. If a *wh* phrase remains in situ within an adjunct clause (157) or PP (158), these are also incompatible with *bahwa*:

- (157) Melly ber-pikir { *bahwa / ?kalau / \emptyset } aku marah karena dia me-lakukan apa?
 Melly MV-think C 1s angry because 3s AV-do what
 ‘What did Melly think I was angry because he did?’
- (158) Ali ber-kata { *bahwa / ?kalau / \emptyset } Tono tinggal dengan siapa?
 Ali MV-say C Tono live with who
 ‘Who did Ali say that Tono lives with?’

These examples show that in questions (with a matrix question interpretation), *bahwa* cannot introduce any clause that contains a *wh* phrase, even if the *wh* word occurs in an adjunct.²¹ *Kalau*, on the other hand, is compatible with *wh*-in situ in (156) and some speakers also accept *kalau* in (157-158). Other speakers are uncertain or noted that *kalau* was degraded in (157-158). Despite the varying judgments about *kalau*, I note that there is a strong contrast between *bahwa* and *kalau* in (157-158), and that all speakers reject *bahwa* in *wh*-in situ questions.

We have already seen that adjunct clauses and PPs are islands for syntactic movement, so that overtly-moved *wh* questions cannot be extracted from them; *wh*-in situ questions contrast with moved-*wh* questions because *wh* in situ does not show island sensitivity in Indonesian.²²

²¹*Bahwa* is possible however, with indirect interpretations of partial *wh*, as discussed below.

²²These *wh*-words questions do not appear to require D-linking (Pesetsky 1987) in order to appear in situ; they can occur without prior discourse context, although this is clearer in simple *wh*-in situ questions (i.e. not embedded). Pesetsky notes that some *wh*-in situ questions that appear island-insensitive do exhibit Subjacency effects, citing observations in Choe 1984 and Nishigauchi 1984: for Japanese *wh*-in situ, the set of felicitous answers to *wh*-in situ questions correlates with islands. This distinction is not relevant for Indonesian. For example, felicitous answers to (158) include ‘Tono lives with Billy’ as well as ‘with Billy’ and ‘Billy.’

Additionally, (156) and (157) show a contrast between verbal morphology and complementizer morphology. The active verbs in the embedded clauses do not show *wh*-agreement (i.e. they retain the active prefix *meN-*), which indicates that no A-bar movement has occurred. Therefore, the unavailability of *bahwa* is yet unexplained. To my knowledge the fact that *bahwa* is incompatible with *wh*-in situ questions has not been discussed with regard to Indonesian or Malay.²³ In the discussion that follows, I consider possible approaches to understanding how *wh*-in situ questions in Indonesian are derived without movement.

Pesetsky (1987), Cheng and Rooryck (2000), Cheng (2003, 2009) and Bayer (2006) suggest that *wh*-in situ varies across languages, and that more than one mechanism for licensing and interpreting *wh*-in situ is available. I review three types of analyses in light of the Indonesian data, without claiming that all *wh*-in situ questions are derived in the same way. For Indonesian, I argue that the data do not support a derivation of *wh*-in situ questions in the narrow syntax. Two other possible analyses – a covert movement approach at LF, and an approach that does not involve any movement – are both compatible with the lack of *wh*-agreement on active verbs. Based on independent evidence from *wh*-indefinites and verbal morphology, I pursue an analysis for Indonesian *wh*-in situ questions that does not involve any movement, overt or covert. I suggest that the reason why C must be null in *wh*-in situ questions has to do with the composition of *wh* words, and the requirement that *wh* variables must be bound by a question Operator. I propose that the complementizer *bahwa* blocks *wh*-variable binding. This operation is associated with the CP domain rather than the verbal domain.

The first approach to *wh*-in situ questions appeals to a copy theory of movement (cf. Chomsky 1995), in which any position in a movement chain may be pronounced. If we assume that movement occurs in the syntax and proceeds in cyclic fashion through the edge of phases, then in Indonesian long distance questions the *wh* phrase could be pronounced in multiple positions along its path of movement: its base position (resulting in *wh*-in situ), intermediate positions (partial *wh*), or highest scopal position (sentence-initial *wh*). This type of analysis is pursued by Reintges et al.

²³Saddy (1991) gives several examples in which *bahwa* introduces an embedded clause with a *wh* phrase (see Saddy's examples 37 and 39; *bahwa* is in parentheses). I note that these examples are ungrammatical for my Indonesian consultants, both with and without *bahwa*.

(2006) for Passamaquoddy and Coptic Egyptian (following proposals in Groat and O’Neil 1996; Bobaljik 2002; and others). The evidence comes from morphological wh-agreement, which they take to mark A-bar movement in overt syntax before spellout to PF/LF. This movement is followed by later pronunciation of the lowest copy (at PF).

Under this type of “apparent wh-in situ” analysis, A-bar movement of the wh-phrase obtains in the syntax, but can only be observed through wh-agreement on C. A variant of this approach is to maintain that the wh phrase remains in situ, while a silent element separates from it and undergoes movement (as suggested for Japanese in Watanabe 1992). The silent element is a question operator, or alternatively, only a wh-feature (cf. Aoun and Li 1993; Watanabe 2001) that moves to scopal position. For Indonesian, both of these approaches are challenged by the fact that verbs do not show wh-agreement as they do in all other cases of A-bar movement. Since pronunciation of copies occurs at PF, the choice of which copy to pronounce should have no effect on wh-agreement.

A second approach is to assume that in wh-in situ questions, the wh element does not move in the syntax, but rather that there is covert movement for semantic computation at the level of Logical Form (LF). For example, Huang (1982) proposes that for Chinese wh-in situ questions, the wh-element covertly moves to a scopal position at LF (generally undergoing wh movement as outlined in Chomsky 1977). The LF of wh-in situ questions, then, results in a structure that is the same as sentence-initial wh questions. If the null morphology on active verbs is a reflex of overt movement in syntax, then covert movement does not affect spellout of the prefix *meN-*. By the same logic however, the complementizer in (156) should not be prevented from being spelled out as *bahwa*, since only covert movement obtains at LF. Note that this is a problem whether covert movement at LF is assumed to be subject to island effects or immune to island effects.

The deviation between complementizers and verbal morphology, which otherwise pattern together for wh-agreement in Indonesian, motivates an analysis that does not involve movement. I therefore pursue an approach in which the interpretation of wh-in situ questions requires neither covert movement at LF nor movement in the narrow syntax. Rather, wh-in situ questions derive their matrix question interpretation through binding by a silent Operator that is merged high in the

structure.²⁴ This question Operator was first proposed in Cole and Hermon 1998 for questions in Malay, which is closely related to Indonesian.²⁵ Cole and Hermon propose that wh words are not uniformly generated across languages. Overt wh words such as *apa* ‘what’ and *siapa* ‘who’ minimally consist of a wh variable (represented as VAR). In languages like Chinese, the wh word is stored (or “lexicalized”) as VAR only, and only receives a question interpretation if it is bound by a separate question Operator (OP). In languages like English, the wh variable is lexicalized together with the question Operator (represented as OP+VAR). In this case, the wh variable and the Operator always undergo syntactic movement together so that the surface position of *who*, *what*, etc. always marks the scope of the question Operator. According to Cole and Hermon’s analysis, languages like Malay employ more than one option for wh words. The wh element and the Operator may be generated together (OP+VAR); they can also be generated separately in different syntactic positions (represented as OP...VAR). These two options are shown in Table 1.

| wh-word | type of question | |
|---|----------------------|---|
| [OP+VAR] <i>apa</i> ‘what,’ <i>siapa</i> ‘who’ | moved-wh questions | sensitive to islands; wh-agreement on verb |
| [OP...VAR] <i>apa</i> ‘what,’ <i>siapa</i> ‘who’ | wh-in situ questions | does not obey islands; no wh-agreement on verb |

Table 1: Cole and Hermon’s (1998) analysis of WH-words in Malay

For moved-wh questions in Malay, Cole and Hermon propose that the language employs the wh word that contains both a variable and Operator (OP+VAR), and this wh-word moves to its surface position via syntactic movement. As in Indonesian, A-bar movement in Malay is blocked by islands, and also requires the bare verb form when an active verb is crossed by movement. In partial wh questions, the wh word (OP+VAR) moves to an intermediate position: if OP remains in this position, it yields an indirect (embedded) question reading; if OP undergoes covert movement to matrix CP, it yields a matrix question interpretation. (Note that this analysis does not account for

²⁴Note that the question Operator is distinct from the Operator discussed in Section 2.3.2, which undergoes relative-internal A-bar movement. The question Operator binds a wh-variable in its scope.

²⁵Cole and Hermon 1998 report that their data is from educated Malay speakers living in Singapore. Malay and Indonesian are widely considered to be different standardized varieties of the same language, although colloquial and regional varieties of Malay and Indonesian may differ in significant ways.

the clefted nature of moved-wh questions; I return to this point shortly.)

Under Cole and Hermon's account, wh-in situ questions in Malay employ a separate variable and question Operator, represented as (OP..VAR). Neither the question Operator nor the wh-word undergo movement. The wh-word (VAR) is generated in its surface position, and the silent OP is generated in matrix CP. Additional support for this idea comes from the fact that wh-words occur independently as variables (or wh-indefinites) in Malay and Indonesian. The examples below illustrate this briefly for *apa* and *siapa*, although other wh-words can also occur as wh variables (see Sneddon 1996; Cole and Hermon 1998; Sneddon et al. 2012). The following illustrate non-interrogative uses of the wh variables:

- (159) Dia tidak mem-beli apa-apa untuk saya.
3s Neg AV-buy what-Redup for 1s
'He did not buy *anything* for me.' (modified from Cole and Hermon 1998:239)
- (160) Saya tidak kenal siapa-pun di universitas itu.
1s Neg know who-PUN at university that
'I don't know *anyone* at that university.' (modified from Cole and Hermon 1998:239)
- (161) Kamu boleh minum apa saja.
2s may drink what SAJA
'You can drink *anything*.' (Sneddon 1996:171)
- (162) Apa-pun yang terjadi, jangan mundur.
what-PUN C.Foc happen Neg.Imper retreat
'*Whatever* happens, don't back down.'
- (163) sesuai untuk siapa-siapa yang suka makanan Jepang
appropriate for who-Redup C.Foc like food Japan
'appropriate for *whoever/anyone* that likes Japanese food'
- (164) Siapa saja boleh ikut.
who SAJA may join
'*Anyone* can come along.' (Sneddon 1996:171)

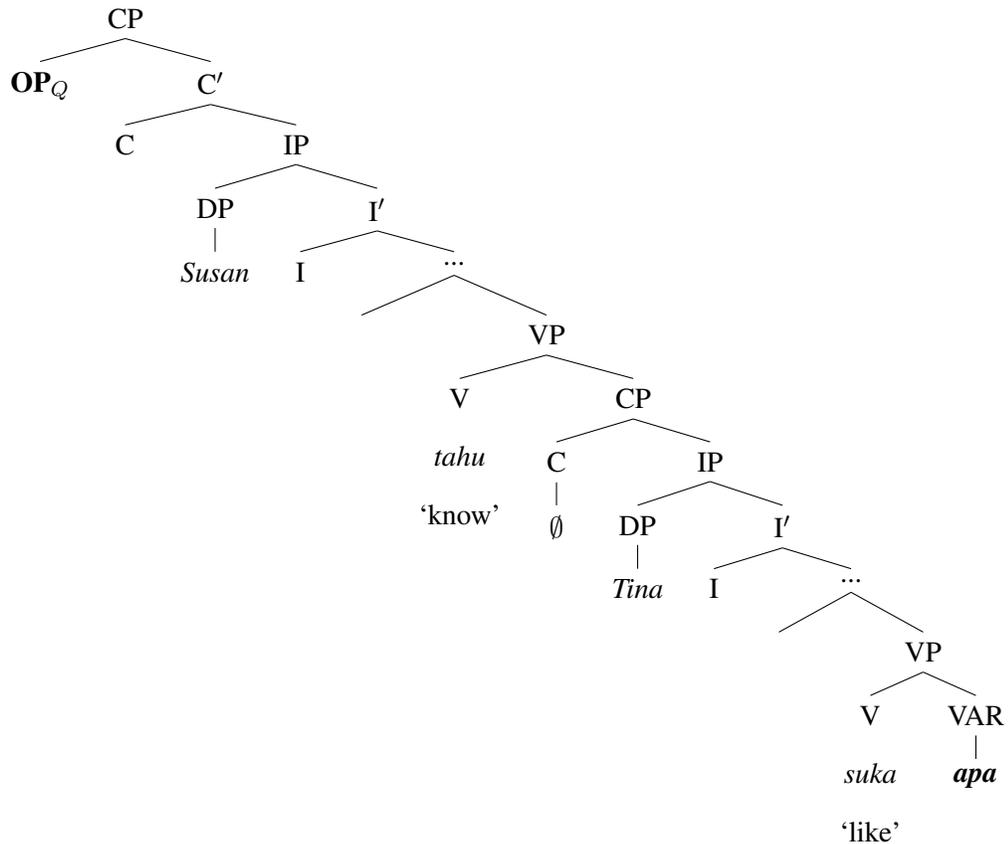
In these examples, the wh-word occurs in a reduplicated form (*apa-apa*, *siapa-siapa*) or with a particle (*-pun*, *saja*). These are non-interrogative contexts, demonstrating that *siapa* and *apa* can occur as a non-interrogative element, i.e. a wh variable bound by a non-interrogative Operator.

Returning to the derivation of questions under Cole and Hermon's analysis, the same wh-variables discussed above occur as in situ wh elements. For example:

- (165) Susan tahu Tina suka apa?
 Susan know Tina like what
 'What does Susan know that Tina likes?'

In (165) the overt wh-word *apa* is only a wh-variable (VAR). The matrix question interpretation is derived from a silent question Operator (OP), which is merged in matrix CP, from which position it unselectively binds the wh variable in the embedded clause. I illustrate this with the tree in (166):

- (166) Wh-in situ question with unselective binding of *apa*



I adopt the general treatment of wh-in situ questions proposed by Cole and Hermon (1998) for Malay: the interpretation of wh-in situ questions is derived by unselective binding of a wh-

variable by a separately generated question Operator, as shown in (166). This analysis for wh-in situ questions does not require any movement, which is welcome for Indonesian since wh-in situ can occur inside islands, and since no wh-agreement is triggered on active verbs.

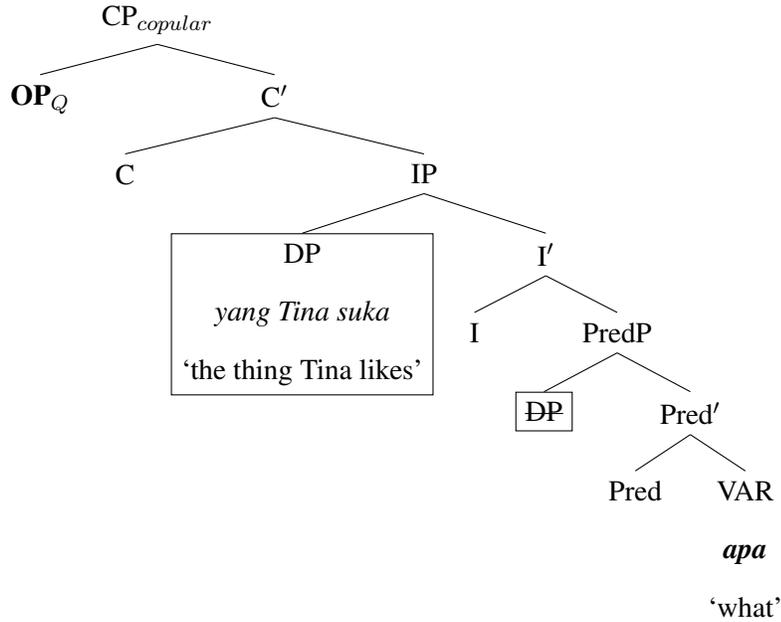
I propose, however, that a single analysis can extend to all types of wh constituent questions in Indonesian, not just wh-in situ. In other words, wh-in situ questions, partially-moved wh questions and sentence-initial wh questions all contain a wh variable (VAR) and a question Operator (OP), which are merged separately in the syntax. The basic proposal is as follows:

- Indonesian wh words are merged as a wh-variable (VAR);
- A question Operator unselectively binds VAR, resulting in interrogative force;
- Matrix question readings of partial wh questions arise when a question OP is merged in matrix CP: [OP_Q You know [what Tina likes]];
- Indirect readings of partial wh questions arise when VAR is bound by a question OP in the embedded clause: ‘You know [OP_Q what Tina likes].’

My proposal simplifies Cole and Hermon’s analysis by requiring only one type of stored wh-word (VAR), and only one type of operation (binding by a question Operator) to derive the range of wh questions in Indonesian. Cole and Hermon’s analysis requires that two different sets of Malay wh words are stored in the speaker’s grammar: a word that fuses the question Operator and wh variable (OP+VAR) as well as an overt wh variable that is generated separately from the Operator (OP...VAR). Cole and Hermon also require two different mechanisms for deriving moved-wh questions: for sentence-initial wh questions, the wh word (OP+VAR) moves to its surface position; for (matrix readings of) partial wh movement, the wh word (OP+VAR) moves, followed by covert movement of this wh word (OP+VAR) after spell-out. Under the new analysis, it is unnecessary to make this distinction between sentence-initial wh and partial wh questions in Indonesian. All wh-words in Indonesian are merged in the syntax simply as a wh-variable (VAR), whether in situ or moved. This variable is bound by the question Operator which is merged high in the structure This preserves the identity of wh words, and simplifies the analysis of question syntax in Indonesian.

I re-examine the structure of moved-wh questions in light of this proposal. Recall that all moved-wh questions result in a pseudo-cleft, while wh-in situ questions are not clefts. For a moved-wh question, a cleft structure begins as shown in (167), with the wh-phrase low in its base position:

(167) Cleft structure with wh-variable and question Operator merged separately

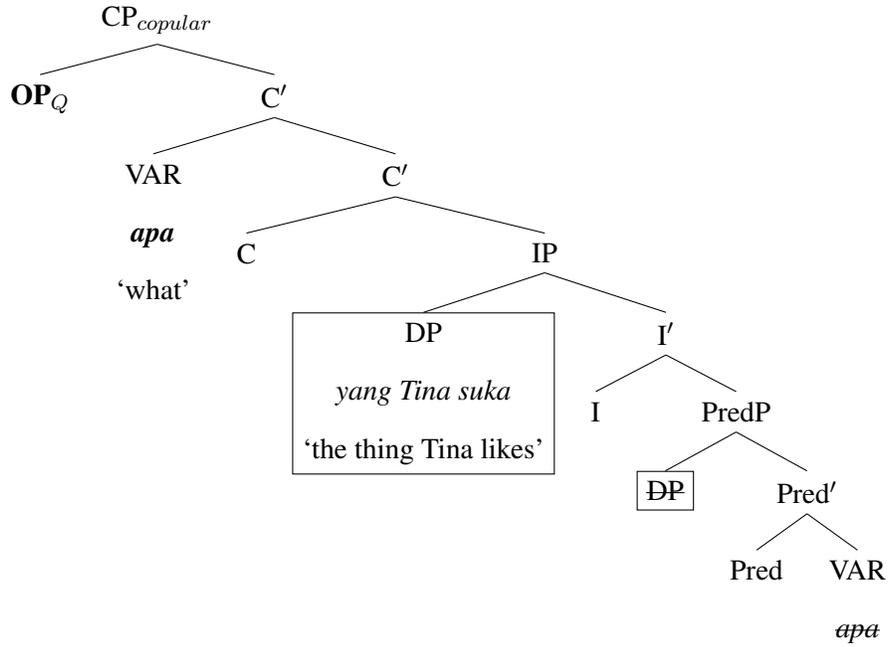


I have proposed that *apa* is a wh-variable that does not contain an interrogative element. To be interpreted as a matrix question, the wh-variable must be bound by a question Operator, which I have represented as OP_Q in the specifier of matrix CP. Note that the wh-phrase *apa* may remain in this low position and occur after the headless relative:

(168) Yang Tina suka ____, apa?
 C.Foc Tina like what
 'The thing that Tina likes, is what?'

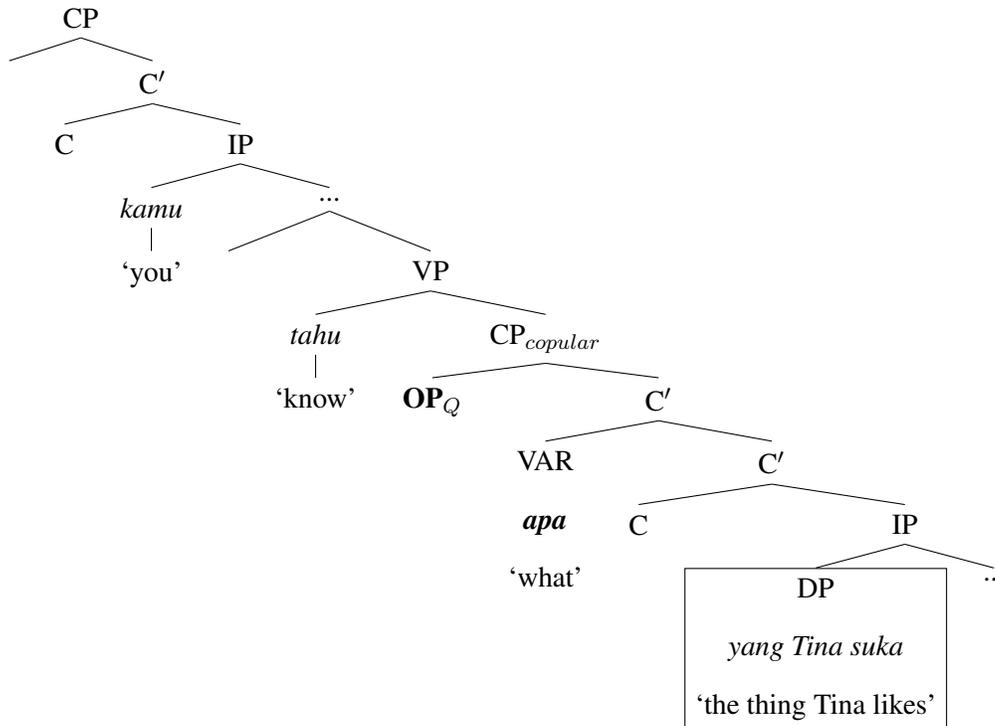
When the wh-phrase *apa* moves to matrix CP, it results in a sentence-initial wh question. The variable is still bound by the question Operator:

(169) Cleft structure with question Operator and wh-variable occurring in sentence-initial position



Let us look at partial wh movement based on (169). In a partial wh question, the clefted structure in (169), i.e. a copular CP, is embedded within another clause:

(170) Partial wh question with question Operator and wh-variable occurring separately



The resulting partial-wh question has two readings:

- (171) Kamu tahu apa yang Tina suka ____.
 2s know what C.Foc Tina like
 Matrix: 'What do you know Tina likes?'
 Indirect: 'You know what Tina likes.'

The indirect reading is derived directly from the structure shown in (170): the question Operator occurs in the embedded clause, so the matrix clause does not have interrogative force. A matrix interpretation can also be derived from (170): if the question Operator occurs in matrix CP, rather than in the embedded clause, then it carries interrogative force into the matrix clause. In both cases, the Operator binds the wh-variable *apa*.

Movement of the (overt) wh-word is strictly local within one CP under this analysis. Long-distance movement obtains via the null relative Operator; the wh-word only undergoes short movement to SpecCP within the copular clause. This is true for both sentence-initial wh questions and partial wh questions.

This analysis also sheds new light on a puzzling pattern that is reported in both Saddy 1991 and Cole and Hermon 1998: partial wh movement is not only constrained by islands, which prevent movement out of the island, but movement *within* a syntactic island does not appear to be possible. For example, Cole and Hermon test the ability of *apa* to move within a relative clause:

- (172) Kamu sayang [perempuan yang Ali fikir [yang telah makan **apa**]]
 you love woman that Ali thinks that already eat what
 ‘You love the woman who Ali thinks ate what?’
- (173) *Kamu sayang [perempuan yang Ali fikir [**apa**_i yang telah makan t_i]]
 you love woman that Ali thinks what that already eat
 (‘You love the woman who Ali thinks ate what?’) (Cole and Hermon 1998:235)

As previously discussed, wh-in situ is not constrained by syntactic islands; the wh word may remain in situ in (172), within a relative island. In contrast, the partial wh question in (173) is not possible, even though no overt movement has crossed the boundary of the relative clause. Both Saddy (1991) and Cole and Hermon (1998) take this pattern to mean that overt movement of the wh element within a syntactic island is not possible. This surprising pattern is used to support Cole and Hermon’s proposal that overt wh movement of *apa* within the relative can be followed by covert movement to scopal position; it is the covert movement that is not possible across an island boundary in (173).

However, the pattern in (172-173) receives an alternate explanation under the present analysis. Consider the most deeply embedded clause in (173): [*apa yang telah makan t_i*] ‘what (the woman) ate.’ This is a clefted structure, which means that relative-internal movement has obtained from object position. Note, however, that the subject of this clause is the extracted argument *perempuan* ‘woman.’ The source of the problem, then, is not that *apa* moves within a syntactic island, but that two extractions have been forced from the same relative clause: both *perempuan* and *apa* have undergone movement within the relative clause. To my knowledge, multiple extraction has not been reported to be possible for either Indonesian or Malay. Thus (173) is ruled out. The grammaticality of (172) is also easily explained: wh-in situ involves no syntactic movement, and does not utilize a cleft strategy. The wh-word *apa* does not move from its base position; the argument *perempuan* is therefore free to undergo A-bar movement from the embedded clause.

The structure of clefts and wh questions naturally yield the restriction on movement of a wh word within an island. There is no longer any reason, then, to appeal to covert movement to account for (173).

Next, I return to the effect of *bahwa* in partial wh movement. In (174-175), null C and *bahwa* deviate with respect to available interpretations for partial wh sentences. A null complementizer is compatible with both readings, but *bahwa* only permits the indirect reading.

- (174) Sudah jelas \emptyset Susan tahu siapa yang ___ suka kue coklat
 already clear Susan know who C.Foc like cake chocolate
 ✓ Matrix: ‘Who is it clear that Susan knows likes chocolate cake?’
 ✓ Indirect: ‘It is clear that Susan knows who it is that likes chocolate cake.’
- (175) Sudah jelas bahwa Susan tahu siapa yang ___ suka kue coklat
 already clear C Susan know who C.Foc like cake chocolate
 X Matrix: ‘Who is it is clear that Susan knows likes chocolate cake?’
 ✓ Indirect: ‘It is clear that Susan knows who it is that likes chocolate cake.’

Previously I showed that *bahwa* cannot introduce a clause with a wh phrase that remains in situ, and I have proposed that the reason is because the question operator cannot bind a variable across this complementizer. Similarly, *bahwa* does not allow a matrix question interpretation in sentences with partial wh movement because the variable in the moved wh phrase must be bound by a question Operator in matrix CP. Binding across *bahwa* is not possible, so the question OP must be in an embedded position in (175), so that interrogativity is restricted to the embedded clause (below *bahwa*), resulting in an indirect question. Another example of an indirect reading forced by *bahwa* is given below with the wh-word *berapa* ‘how many’:

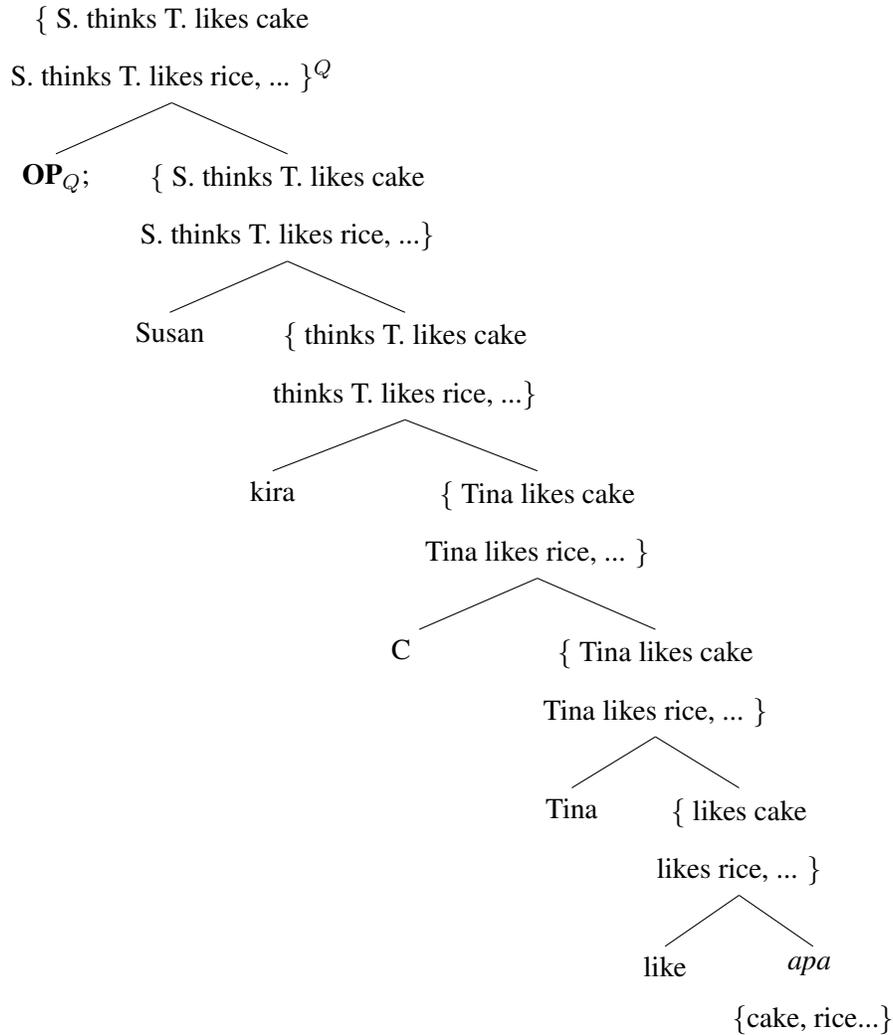
- (176) Aku kira { bahwa / kalau / \emptyset } lebih dari 100 orang ikut dalam rombongan.
 I_s think C more from 100 person join in crowd
 ‘I think more than 100 people were in the crowd.’
- (177) Sudah di-ketahui bahwa berapa orang yang ikut dalam rombongan.
 already PV-discover C how.many person C.Foc join in crowd
 X Matrix: ‘How many people were known to be in the crowd?’
 ✓ Indirect: ‘It was known how many people were in the crowd.’

Berapa behaves like the wh-variables that we already seen throughout this chapter, *apa* and *siapa*. *Bahwa* does not allow a matrix question interpretation, so (177) is interpreted as an indirect question. Under the present analysis, this means that a question Operator binds *berapa* from the CP of the embedded clause, below *bahwa*.

The ability of the overt complementizer *bahwa* to block variable binding is a novel finding, and the reason for this is not completely clear. If this blocking effect applies more broadly than in the interpretation of questions, then *bahwa* is expected to block other types of inter-clausal operations as well. One possibility is that *bahwa* is not compatible with the semantic computation of alternatives. Following Hamblin 1973; Karttunen 1977; among others, I take the denotation of a question to be a set of alternatives, i.e. possible answers to the question. The wh-word *apa* ‘what’ in (178) introduces this set of alternatives, e.g. {cake, rice, soup, bread...}. The semantic computation of this set of alternatives proceeds from the bottom up, passing along the alternatives in stepwise fashion. This is illustrated in a simplified way in (179):

(178) Susan kira { *bahwa / kalau / \emptyset } Tina suka apa.
 Susan think C Tina like what
 ‘What does Susan think Tina likes?’

(179) Alternatives introduced by wh-in situ



Once the set of alternatives reaches the matrix CP, it combines with the question Operator. The sentence then has interrogative force, and the denotation of the matrix sentence is the set of alternatives {Susan thinks Tina likes cake, Susan thinks Tina likes rice, ... }.

Under this view of *wh* in situ questions, then, binding of the *wh* variable in *apa* depends on a set of alternatives being passed along the stepwise computation at LF, until a set of propositions reaches the question Operator. If the embedded complementizer in (179) is null, the computation proceeds without exception. If the intermediate C is *bahwa*, however, the alternatives fail to be passed further up the structure.

If this is on the right track, and the incompatibility of *bahwa* in wh in situ questions arises from the interpretation of this overt form, this makes a prediction that *bahwa* might interfere with other types of semantic computation.²⁶ For example, if a different type of Operator requires binding across a clausal boundary, then the presence of *bahwa* is expected to prevent such binding. Some initial evidence from embedded disjunctions appears to confirm this. Consider the embedded disjunction in the English sentence in (180) for example:

- (180) Lisa said Susan is from New York or New Jersey.
- a. Lisa said “Susan is from New York or New Jersey.”
 - b. Lisa said Susan is from New York, or Lisa said Susan is from New Jersey (but I can’t remember which).

Two possibilities exist for the scope of disjunction (cf. Rooth and Partee 1982; and subsequent). The embedded disjunction can be interpreted locally (180a), in which case the embedded clause contains the disjunction: [Susan is from New York] \vee [Susan is from New Jersey]. A second interpretation is also possible (180b), in which the disjunction scopes higher than the clause boundary: [Lisa said Susan is from New York] \vee [Lisa said Susan is from New Jersey]. In this case, the disjunction is interpreted higher than the clause in which it occurs.

In Indonesian, the same ambiguity is possible when null C introduces the embedded clause:

- (181) Lisa meng-ata-kan Susan ber-asal dari New York atau New Jersey.
 Lisa AV-say-Appl Susan MV-origin from New York or New Jersey
- a. \surd Lisa said “Susan is from New York or New Jersey.”
 - b. \surd Lisa said Susan is from New York, or Lisa said Susan is from New Jersey.

If *bahwa* introduces the embedded clause however, the disjunction cannot scope higher than C:

- (182) Lisa meng-ata-kan bahwa Susan ber-asal dari New York atau New Jersey.
 Lisa AV-say-Appl C Susan MV-origin from New York or New Jersey

²⁶Thanks to Mitcho Erlewine for discussion and suggestions related to this point.

- a. ✓ Lisa said “Susan is from New York or New Jersey.”
- b. X Lisa said Susan is from New York, or Lisa said Susan is from New Jersey.

The effect of *bahwa* in (182) is to limit the disjunction to a local interpretation, i.e. within the clause it embeds. I assume that the interpretation of disjunction requires an abstract Operator (i.e. not the surface form of the disjunction), so the two scope possibilities in (180) and (181) indicate that the disjunction Operator can move higher than its surface position. In (181) the Operator can raise out of the embedded clause and scope over the matrix clause. In (182) however, the Operator is limited to the embedded clause, which suggests that *bahwa* blocks the movement of the disjunction Operator. The effect of *bahwa* is to block Operator binding across the clause boundary, similar to the unavailability of binding by a question Operator in *wh*-in situ questions. Similar blocking or intervention effects have been observed in other languages, with negation and focus Operators; see Larson 1985; Han and Romero 2004; Beck 2006; Beck and Kim 2006.²⁷ This initial evidence suggests that *bahwa* can have semantic effects beyond the interpretation of questions, but may also apply to other Operators as well.

Finally, the ability of *kalau* to occur in *wh*-in situ questions has not yet been discussed. To re-summarize briefly, when occurring as a complementizer, *kalau* is more common in informal speech than in formal or written Indonesian. *Kalau* has the same distribution as *bahwa* when introducing declarative clauses. *Bahwa* is ruled out in both moved-*wh* questions and *wh*-in situ questions, while *kalau* is ruled out only in moved-*wh* questions as a consequence of *wh*-agreement on *C*. *Kalau* is possible in *wh*-in situ questions (although consultant judgments are not completely consistent), where *bahwa* is impossible. I repeat the relevant *wh*-in situ examples below:

- (183) Pak Dadang meng-harap { **bahwa* / *kalau* / \emptyset } mereka akan men-diskusi-kan masalah
 Mr Dadang AV-expect C 3p will AV-discuss-Appl problem
 apa?
 what
 ‘What issue does Mr Dadang expect that they will discuss?’

²⁷To my knowledge, intervention effects are noted to be a property of focus-sensitive operators such as *only*, *even* and negation. I leave open for future study whether *bahwa* is an exception to this generalization, since it is not related to focus.

- (184) Melly ber-pikir { *bahwa / ?kalau / \emptyset } aku marah karena dia me-lakukan apa?
 Melly MV-think C 1s angry because 3s AV-do what
 ‘What did Melly think I was angry because he did?’
- (185) Ali ber-kata { *bahwa / ?kalau / \emptyset } Tono tinggal dengan siapa?
 Ali MV-say C Tono live with who
 ‘Who did Ali say that Tono lives with?’

For what reason does *kalau* differ from *bahwa* in the wh-in situ sentences above? Since wh-in situ requires binding by a question Operator binding across C, these examples show that *kalau* does not block this relation in the same way that *bahwa* does. I have briefly mentioned in Section 2.3.5 that for embedded polar questions, either the question word *apa(kah)* or *kalau* may introduce the embedded question (*kalau* is less formal than *apakah*, which is required in written Indonesian):

- (186) Dia tanya { apa-kah / kalau } kita bisa ber-temu jam sembilan.
 3s ask what-Q C 1p.Incl can MV-meet hour nine
 ‘She asked if/whether we can meet at nine o’clock.’
- (187) Masih di-ragu-kan { apa-kah / kalau } pendapat ini dapat di-terima.
 still PV-doubt-Appl what-Q C opinion this able PV-accept
 ‘It is still doubtful whether/if this opinion can be accepted.’ (modified from Sneddon 1996:323)

These sentences have an indirect question reading, and the matrix clause does not have interrogative force, so I assume that there is no binding by a question Operator across C. However, these examples show that when the matrix verb selects for an interrogative clause, *kalau* may introduce a question. In general then, the complementizer *kalau* is compatible in interrogative contexts, for all speakers. This provides a partial explanation for the well-formedness of the wh-in situ question in (183) and (for some speakers) the acceptability of *kalau* in (184-185). I assume that when *kalau* occurs as C, it may bear an interrogative feature [Q].

As for the inconsistent judgments for *kalau* in (184-185), I suggest that another property of this form is relevant here. Aside from its use as an informal complementizer, *kalau* occurs in conditional clauses. (Other conditional morphemes include *jika*, *jikalau*, which are considered

formal.) As a conditional morpheme, *kalau* is compatible only with non-past orientation. This is illustrated in (188):

- (188) *Kalau* saya kuliah, saya akan ambil jurusan matematika.
KALAU 1s go.to.univ 1s will take major mathematics
✓ ‘If I go to university, I will major in mathematics.’
✓ ‘If I were to go to university, I would major in mathematics.’
✓ ‘When I go to college, I will major in mathematics.’
X ‘When I went to college, I majored in mathematics.’

Kalau introduces the antecedent of a conditional clause. The clause introduced by *kalau* may have several different temporal readings, but is not compatible with past orientation. Returning to the examples in (183-185), only sentences with past orientation are degraded for some speakers. (184) and (185) are not compatible with *kalau* because they are naturally interpreted as past tense. In contrast, the embedded clause in (183) is in the future tense. Furthermore, the semantics of the matrix verb *mengharap* ‘expect, hope’ contrasts with *berpikir*, *berkata* ‘think,’ ‘say’ because it introduces future possibilities.

2.4. Deterministic and non-deterministic conditions on Indonesian C

The discussion thus far has demonstrated that a number of semantic and syntactic conditions contribute to the surface alternations in Indonesian complementizers. The findings that have been discussed thus far are summarized in Table 2, with a list of semantico-syntactic features on C. Here I assume the general framework of Distributed Morphology (Halle and Marantz 1993, 1994) to discuss a post-syntactic phase in which abstract morphemes are replaced with phonological strings that match their featural content (Vocabulary insertion).

The feature [root] results in null C, without exception. This means that all non-embedded clauses cannot occur with an overt complementizer. I have mentioned that some consultants allow sentential subjects introduced by *bahwa*, and also that when a clefted nominal occurs at the left

| | Features on C | Surface form(s) of C | |
|------------------------|----------------|---------------------------------|---|
| Root | [root] | ∅ | |
| Non-root (embedded) | [Focus, EPP-D] | <i>yang</i> | Morphological wh-agreement (These features co-occur with [declarative]) |
| | [EPP-D] | ∅ | |
| | [declarative] | <i>bahwa</i> ~ <i>kalau</i> ~ ∅ | Variable; subject to discourse-pragmatic factors |

Table 2: Semantico-syntactic features on C

edge of the clause it must be immediately followed by the complementizer *yang*. Both these cases, however, are cases of embedded C. In the case of sentential subjects, *bahwa* belongs to the CP constituent embedded in subject position. In pseudo-clefts, recall that a headless relative occurs as subject of a copular clause, and the complementizer *yang* is embedded within the subject (see tree structure in 74).

The feature [EPP-D] occurs on functional heads and drives A-bar movement by attracting a DP (i.e. an Operator) to its specifier. If the nominal does not stay in this position, but continues to a higher position, the result of this A-bar movement is that C must be null. When C bears both [EPP-D] and [Focus], then the extracted argument remains in the specifier of C. We have already seen that in this case, C must be realized as *yang*. Morphological wh-agreement, then, results from the features [EPP-D] and [Focus], which determine the final position of the moved argument as well as its surface form.

The features [root], [EPP-D] and [Focus] result in deterministic conditions on the realization of C. As we have already seen however, not all of the realizations of C result from deterministic conditions. In the absence of [root], [EPP-D] and [Focus], the realization of C is variable, and may occur as *bahwa*, *kalau* or null C. This optional alternation has no semantic or syntactic motivation, and does not depend on syntactic structure or syntactic operations.

Consider a complementizer that bears a [declarative] feature. This feature is compatible with both *bahwa* and *kalau*, either of which may introduce an embedded declarative clause. The semantico-syntactic conditions on C are not deterministic, since three different realizations are possible. When C[declarative] is spelled out in the morphophonology, *bahwa*, *kalau* or null C may occur. One possible way to impose deterministic conditions on *bahwa* and *kalau* is to introduce a formal/informal distinction in the morphosyntax, e.g.: C that bears a feature [formal] is realized as *bahwa*; C that bears [informal] is realized as *kalau*. However, this approach is not adopted because it entails a strict binary opposition between *bahwa* and *kalau*, whereas either form may be used in many contexts; furthermore, there is little reason to introduce these features into the syntax since no grammatical contrasts depend on a formal-informal distinction.

The optionality among *bahwa*, *kalau* and null C is implemented by two separate mechanisms. First, the choice between two overt forms, *bahwa* and *kalau*, is a pragmatic alternation between formal and informal morphemes that compete for usage. Whether *bahwa* or *kalau* is inserted at C is not determined by the featural content of C. Rather, these are exponents of competing vocabulary items whose featural content allow either to be inserted at a C node during vocabulary insertion. Competition between formal and informal roots is not unusual; in this case however, the competition is between formal and informal variants of a functional morpheme.

The second mechanism that derives optional realization of C is a variable deletion rule that applies after spellout, in the morphophonology:

(189) Variable deletion rule

C [declarative] → ∅

Variable rules have been implemented in order to capture non-deterministic phenomena in other languages (cf. Labov 1969 and subsequent). The rule in (189) derives the optional alternation between overt declarative C (*bahwa*, *kalau*) and the null C. The deletion operation targets the feature [declarative]; no distinction is made between the formal and informal complementizer. The rule also assumes that C bearing the feature [Focus] are not deleted; focused elements are semantically and

phonologically prominent, and cannot be elided.

The surface realization of C is first determined derivationally, that is, by the syntactic and semantic requirements of structure building including A-bar movement and formation of questions. External to syntax, realization of C is also subject to two mechanisms that give rise to variable surface realization: variable deletion of C (overt *bahwa/kalau* vs. null C), and pragmatic selection between formal and informal vocabulary items (*bahwa* vs. *kalau*)

2.5. Chapter summary

In this chapter, the investigation of optional declarative C requires syntactic operations as well as extra-syntactic processes to derive the obligatory and optional alternations in Indonesian declarative complementizers.

I have shown that when nominal A-bar extraction occurs, the highest C crossed by movement occurs as *yang*, while intermediate C is null; other forms are ruled out. Furthermore, *bahwa* is ruled out in all questions, whereas *kalau* may occur in wh-in situ questions that have a non-past orientation. These morphosyntactic and semantic patterns are not easily captured by a description of the surface distribution of these forms. Neither does a syntactic account fully explain the variability shown by *bahwa* and *kalau*: without A-bar movement or a wh-in situ question, the overt forms *bahwa/kalau* may optionally be null. I have proposed a variable deletion rule to account for this optionality. The variability between *bahwa* and *kalau*, on the other hand, arises from pragmatic choice between formal and informal vocabulary. An understanding of both internal and external factors is necessary to account for the distribution of forms that alternate in a single surface position.

This study in the complementizer system of Indonesian also has broader implications. I have demonstrated that Indonesian complementizers participate in wh-agreement (along with verbal voice prefixes). Indonesian C shows a pattern of wh-agreement that differentiates intermediate landing sites from the final landing site of an A-bar moved nominal, a pattern which is previously unattested. I also proposed a new analysis for wh-phrases, modifying the analysis in Cole and

Hermon 1998), and discussed the implications of pseudo-cleft structure for moved-wh questions in Indonesian.

CHAPTER 3 Verbal prefixes *meN-* and *ber-*

3.1. Introduction

The topic of this chapter is the optional realization of two verbal prefixes in Indonesian, *meN-* and *ber-*. Unlike other verbal prefixes in Indonesian, these appear to be optional in many environments. In the sentences below, verbs prefixed with *meN-* or *ber-* may also occur without the prefix:

- (1) Aku { dapat / men-dapat } hadiah.
1s receive AV-receive gift
'I received a gift.'
- (2) Rika { ringkas / me-ringkas } pelajaran pagi ini.
Rika summarize AV-summarize lesson morning this
'Rika summarized a lesson this morning.'
- (3) Dia { bahasa / ber-bahasa } Indonesia, aku { bahasa / ber-bahasa } Inggris.
3s language MV-language Indonesian 1s language MV-language English
'She spoke Indonesian, I spoke English.'
- (4) Mereka adalah orang-orang yang { pendidikan / ber-pendidikan } baik.
3p Cop person-Redup C.Foc education / MV-education good
'They are people that are highly educated.'

As in other chapters of this thesis, I describe the availability of both the prefixed and unprefixed forms in the same position as “optionality” or “variability.” These terms are used for the availability of more than one surface form, and are intended to be neutral with regard to the source of the optionality; several different factors affect the realization of *meN-* and *ber-*. These two prefixes display intra-speaker variability (rather than inter-speaker or inter-dialectal variability); the prefixed form and the unprefixed form have the same meaning in these clauses, and an individual speaker produces both forms.

In contrast to (1-4), in some clauses *meN-* and *ber-* are not optional. For example, while *ber-* is optional in (5), the prefix is disallowed when the object is questioned (6). There is likewise a contrast between the two verbs in (7): the prefix *meN-* is optional on the first verb, but required on the second:

- (5) Dia { main / ber-main } game komputer sampai larut.malam.
 3s play MV-play game computer until midnight
 ‘He played computer games until midnight.’
- (6) Apa-kah yang dia { main / *ber-main } sampai larut.malam?
 what-Q C.Foc 3s play MV-play until midnight
 ‘What did he play until midnight?’ (modified from Soh 2013:169)
- (7) Aku { rasa / me-rasa } gatal dan muka-ku { *merah / me-merah }.
 1s feel AV-feel itchy and face-1s red AV-red
 ‘I felt itchy and my face reddened.’

The terms “obligatory realization” and “non-optional realization” are used to describe contexts in which only one form of the verb is possible; either the prefix must occur, or the prefix is disallowed.

This chapter investigates which environments constrain the variability of *meN-* and *ber-*, and proposes an analysis that accounts for the distribution of these prefixes. The variability of *meN-* and/or *ber-* in Indonesian has been observed by previous authors, e.g. Wallace 1979; Anderson 1983; Kaswanti Purwo 1989; Wolff 1990; Sneddon 1996; Voskuil 1996; Gil 2002; Nuriah 2004; Wouk 2004; Fortin 2006; Sato 2008b; Sneddon et al. 2012. The same prefixes occur in other varieties of Malay, and these have also been discussed in the literature, e.g. Soh 1998; Nomoto 2006; Nomoto and Shoho 2007; Donohue 2008; Soh and Nomoto 2009, 2011, 2015; Soh 2010, 2013; Yanti 2010; McKinnon et al. 2011; Saad et al. 2015.

Descriptively, there are three types of environments with regard to variability in *meN-/ber-*: (i) environments that disallow *meN-/ber-*, (ii) environments that require *meN-/ber-*, and (iii) environments in which *meN-/ber-* is optional. Both (i) and (ii) are non-optional contexts that require a prefix or disallow a prefix. These contexts have been reported to include imperatives, modal uses, nominal A-bar movement, change of state verbs, certain transitive or intransitive constructions, and many idiosyncratic roots; these are further discussed below. An overview of these environments has been included in some grammars and previous work on verbal morphosyntax (see references cited above). However, the approach that I take in this chapter is to examine the various factors that have been reported to affect the optionality or realization of *meN-* and *ber-*, and to identify these as either deterministic or non-deterministic. The distinction between deterministic and non-deterministic

factors is an important one for the main argument in this chapter, which is that syntactic factors are deterministic and do not allow optional realization of *meN-/ber-*. I show that many of the reported properties of *meN-* and *ber-* are tendencies, rather than inherent properties that arise from deterministic factors. Optionality arises from post-syntactic deletion of *meN-* or *ber-*, a process which is subject to a number of non-deterministic factors.

3.1.1. An overview of the proposal

I assume a modular and derivational view of grammar in which syntax and (morpho)phonology are separate components of the grammar; syntax is an autonomous system, and precedes the phonological component. Syntactic structures are built from roots and abstract functional heads that project syntactic structure; these heads bear features that also drive syntactic operations. After the syntactic component of the grammar is complete, the structure undergoes spellout; in the post-syntactic part of the grammar, phonological strings are inserted for abstract morphemes and other derivational processes also apply. The approach that I take to variability in *meN-* and *ber-* depends on this distinction between the syntactic and post-syntactic components of the grammar; the variable rules that apply to *meN-* and *ber-* must be implemented in the phonology rather than in the syntax, which is autonomous.

In Section 3.4, I argue that *meN-* and *ber-* occupy the structural position of the functional head Voice; the evidence for this structural position comes from patterns of nominal A-bar movement through the edge of the syntactic phase. However, the head Voice is not always pronounced: *meN-* and *ber-* are subject to a number of factors that affect their (optional) realization. These factors fall into two types, deterministic and non-deterministic. Deterministic factors result in obligatory realization, i.e. a prefixed verb or a bare verb. These deterministic factors are syntactic (most notably, nominal A-bar movement, which results in morphological wh-agreement). If deterministic factors are not present in the syntax, then the prefix *meN-* or *ber-* is subject to variability in the morphophonology. I implement this optionality with the following rule:

- (8) Variable deletion rule

$meN- \rightarrow \emptyset$

$ber- \rightarrow \emptyset$

This variable rule is a post-syntactic operation and its rate of application varies among different speakers; it may also vary within an individual speaker. This predicts that application of the rule is influenced by a variety of factors, both grammatical (internal) and non-grammatical (external).

The proposed rule makes explicit the source of optionality of *meN-/ber-*, by placing the variability in the post-syntactic component of the grammar. It also provides an explanation for certain puzzling properties of *meN-* and *ber-*. Some of the properties attributed to *meN-* or *ber-* appear to be syntactic or semantic, but are not deterministic. For example, previous authors have contrasted overt *meN-* with the bare verb by associating the prefix with aspectual properties or co-occurrence with modals. I argue that these are strong tendencies, but these are not deterministic conditions; I present data that support this view. These non-deterministic properties arise from various factors that are correlated with the rate of variable deletion. There are also many roots that require *meN-* or *ber-*, and many others in which a root never occurs with *meN-* or *ber-*; these are “idiosyncratic” roots in the sense that they are not predictable. Superficially, these idiosyncratic cases appear to be deterministic, because different roots require or disallow the prefix. However, I suggest that these cases are also compatible with a view in which the variable deletion rule for *meN-/ber-* is applied at different rates for different roots. Finally, conflicting or uncertain judgments from different speakers are also accounted for under this proposal.

3.1.2. *Organization of chapter*

This chapter has two overall goals. First, I show that the overt prefixes *meN-* and *ber-* are not completely in free variation with a null prefix. Rather, certain syntactic conditions disallow *meN-* and *ber-*. I take an established generalization about Indonesian syntax, that nominal A-bar movement requires a bare verb, to be an instantiation of morphological wh-agreement. This extends the analysis of wh-agreement in Indonesian to the verbal domain; wh-agreement also occurs in the complementizer system (discussed in Chapter 2) and within the DP (Chapter 4).

Second, this chapter demonstrates that both semantico-syntactic factors and extra-syntactic factors are necessary to account for the distributional and variable patterns in *meN-* and *ber-*. In previous accounts of *meN-*, some authors have argued that the variability between a *meN-* verb and a bare verb (without *meN-*) is explained by syntactic factors. Others argue for a discourse-pragmatic analysis in explaining the function(s) of *meN-*.¹ I show that neither of these approaches alone accounts for the optionality of *meN-* and *ber-*. A syntactic analysis alone provides no explanation for the optionality in (1)-(4). On the other hand, surface variability in *meN-/ber-* is often said to be a case of optional pronunciation or optional “dropping” of the prefix. The discussion in this chapter shows that a description of surface variability fails to capture several syntactic factors that disallow optionality; assuming optional pronunciation in the phonology is also insufficient to account for all the data.

The outline of this chapter is as follows. In Section 3.2 I review background on Indonesian verbal prefixes and give an overview of the proposal. In Section 3.3 I survey previous analyses of *meN-* and *ber-* in Indonesian. Section 3.4 discusses deterministic factors that result in an obligatory realization of *meN-* or *ber-*. Section 3.5 discusses optionality arising from post-syntactic deletion, and non-deterministic factors affecting variable realization. Conclusions are given in Section 3.6.

3.2. Background

Indonesian verbs are well-described in grammars (Macdonald and Dardjowidjojo 1967; Wolff 1990; Wolff et al. 1992; Sneddon 1996; Sneddon et al. 2012). The syntax and morphology of verbal affixation has been studied by many previous authors, both for standard Indonesian as well as colloquial varieties of Indonesian (Wallace 1979; Kana 1983; Kaswanti Purwo 1989; Saddy 1991; Soh 1998; Voskuil 2000; Nuriah 2004; Wouk 2004; Cole and Hermon 2005; Fortin 2006; Donohue 2007; Aldridge 2008; Arka and Manning 2008; Cole et al. 2008b; Sato 2008b, 2012; Nomoto 2013). In addition, there is substantial work on verbal morphology in Malay, which is a variety of the same language (Nomoto 2006; Donohue 2008; Soh and Nomoto 2009, 2011; Soh 2010, 2013; Nomoto

¹The prefix *meN-* has received much more attention in theoretical analyses than the prefix *ber-*, which is mostly described in descriptive grammars.

and Wahab 2014; Soh and Nomoto 2015). Several related languages spoken in Indonesia have also been documented as having a similar pattern of verbal affixation (e.g. Arka 2002, 2003, 2008, 2009; Gil 2002; Himmelmann 2002; Wouk and Ross 2002; Davies 2005; Arka and Manning 2008; Sato 2008a; Yanti 2010; Soh and Nomoto 2009, 2015; Hidajat 2013; Soh 2013; Arka and Wouk 2014; Legate 2014; Nomoto and Wahab 2014; Jeoung 2017, among others.)

In this section I briefly outline verbal prefixes in Indonesian, with a focus on the two verbal prefixes that are variable, *meN-* and *ber-*.

3.2.1. Verbal prefixes in Indonesian

The verbal prefixes are shown in Table 3. Four of the prefixes are listed as voice morphemes: Active voice *meN-*, Middle voice *ber-*, Passive voice *di-* and Object voice \emptyset -. The prefix *ter-* occurs in stative, involitive and abilitative contexts; *per-* is a causative morpheme; and the circumfix *ke-an* occurs on adversative verbs.

| | | | |
|-----------------------------|----------------------------------|-------------------------------------|---|
| Variable (in some contexts) | Active voice (AV) | <i>meN-</i> (<i>N-</i>) bare form | <i>me-lihat</i> ‘see’ <i>lihat</i> ‘see’ |
| | Middle voice (MV) | <i>ber-</i> bare form | <i>ber-setuju</i> ‘agree’ <i>setuju</i> ‘agree’ |
| Non-variable | Passive voice (PV) | <i>di-</i> | <i>di-lihat</i> ‘seen’ |
| | Object voice (OV) | \emptyset - | \emptyset - <i>lihat</i> ‘see’ |
| | Stative/ Involitive/ Abilitative | <i>ter-</i> | <i>ter-lihat</i> ‘seen (accidentally)’ |
| | Adversative | <i>ke- -an</i> | <i>ke-hujan-an</i> ‘caught in the rain’ |
| | Causative | <i>per-</i> | <i>mem-per-kuat</i> ‘strengthen’ (‘make strong’) |

Table 3: Verbal prefixes in Indonesian

There is disagreement in the literature about the function of *meN-* and *ber-*. I survey various analyses in Section 3.3. Throughout this chapter, I identify *meN-* as an Active voice (AV) morpheme and *ber-* as a Middle voice (MV) morpheme; I take these to occur as the functional head Voice; support for this view is given in Section 3.4.

I do not assume that all the affixes listed in Table 3 occupy the same syntactic position in the clause, nor that they occur as the same functional head. The prefixes *meN-* and *di-* can co-occur with the causative prefix *per-*, and *per-* also occurs in Object voice with the null prefix \emptyset -. The same prefixes (*meN-*, *ber-* and the null OV prefix) may also occur with verbal suffixes *-kan*, *-i* and *-an* (see Sneddon et al. 2012 for a comprehensive overview of possible combinations of prefixes and suffixes). With the exception of *per-*, the prefixes in Table 3 are in complementary distribution in Indonesian.²

Two prefixes are variably realized, while the others are invariant. I briefly outline the non-variable prefixes below, followed by a description of variable prefixes *meN-* and *ber-*.

3.2.2. Invariant prefixes

The data below illustrate that verbal morphology for passive voice (PV), object voice (OV), the prefix *ter-* and the circumfix *ke-an* are obligatory, and invariant. In (9), omitting the passive prefix *di-* renders the sentence ungrammatical (or changes its argument structure to an active clause). In contrast, the same root is shown in an active clauses, with the optional prefix *meN-*.

(9) Non-variable *di-* (Passive voice)

²There are a few verbs that occur with both *meN-* and *ber-* (*ber-* occurs closer to the root). This is not a productive pattern, and I have found only a few examples such as *mem-ber-henti-kan*, *mem-ber-daya-kan*. These are best treated as compounds in which the *ber-* form is merged as V; note that it can be passivized. The meaning of the *mem-ber-* form is distinct from the *ber-* form: *memberhentikan* ‘to fire (someone)’ from intransitive *berhenti* ‘stop’; *memberdayakan* ‘to equip’ from intransitive *berdaya* ‘to have power.’

- (1) Mobil kami ber-henti.
car 1p.Excl MV-stop
‘Our car stopped.’
- (2) Kami meng-henti-kan mobil.
1p.Excl AV-stop-Appl car
‘We stopped the car.’
- (3) Pak Fajar mem-berhenti-kan kami.
Mr Fajar AV-fire-Appl 1p.Excl
‘Mr Fajar fired us.’
- (4) Kami di-berhenti-kan.
1p.Excl PV-fire-Appl
‘We were fired.’

- a. Pesan ini akan { *sampai-kan / di-sampai-kan } kepada orang kampung.
message this will reach-Appl PV-reach-Appl to person village
'This message will be conveyed to the people in the neighborhood.'
- b. Aku sudah { sampai-kan / meny-ampai-kan } pesan ini kepada orang kampung.
1s already reach-Appl AV-reach-Appl message ini to person village
'I already conveyed the message to the people in the neighborhood.'

The prefix *ter-* occurs in stative, involitive, or abilitative clauses, and is obligatory. The invariant *ter-* verbs below contrast with active forms which allow variable *meN-*.

(10) Non-variable *ter-* (Stative)

- a. Meja ini { *buat / ter-buat } dari kayu.
table this make Stat-make from wood
'This table is made of wood.'
- b. Se-orang tukang { buat / mem-buat } meja dari kayu.
one-person craftsman make AV-make table from wood
'The carpenter built a table from wood.'

(11) Non-variable *ter-* (Involitive)

- a. Buku-nya { *bawa / ter-bawa } oleh teman-ku.
book-D carry Invol-carry by friend-1s
'The book was taken (accidentally) by my friend.'
- b. Teman-ku { bawa / mem-bawa } buku buat bel-ajar.
friend-1s carry AV-carry book for MV-study
'My friend carried a book for studying.'

(12) Non-variable *ter-* (Abilitative)

- a. Pesawat itu { *lihat / ter-lihat } di langit.
plane that see Abil-see at sky
'The plane can be seen in the sky (despite it being very high).'
- b. Dia bisa { lihat / me-lihat } pesawat di langit.
3s can see AV-see plane at sky
'He could see a plane in the sky.'

Ke-an occurs on adversative verbs, and is likewise obligatory:

(13) Non-variable *ke-an* (Adversative)

- a. Kami { *curi / ke-curi-an } saat meng-inap di hotel.
 1p.Excl steal Advers-steal-Advers moment AV-stay at hotel
 ‘We got robbed (experienced robbery) when we stayed at a hotel.’
- b. Orang itu { curi / men-curi } dompet-ku.
 person that steal AV-steal wallet-1s
 ‘That person stole my wallet.’

Finally, the prefix for Object voice is unusual because it is obligatorily null. In Object voice clauses, the Theme occurs as grammatical subject and the Agent remains in its thematic position adjacent to the verb (for discussion of clause structure in Object voice, see Chung 1976; Guilfoyle et al. 1992; Cole et al. 2008a). (14a) shows that other overt prefixes cannot occur in Object voice.

(14) Non-variable \emptyset - (Object voice)

- a. Rumah ini mau aku { \emptyset -beli / *mem-beli / *ber-beli / *di-beli }.
 house this want 1s OV-buy AV-buy MV-buy PV-buy
 ‘I want to buy this house.’
- b. Aku mau { beli / mem-beli } rumah ini.
 1s want buy AV-buy house this
 ‘I want to buy this house.’

The fact the OV prefix is obligatorily null means that it is identical in form to a bare active verb, i.e. when an active verb does not occur with *meN-*, as in (15). This results in a potential ambiguity when the Theme of a transitive clause appears in sentence-initial position, followed by a preverbal Agent:

(15) Active clause with topicalized Theme

Rumah ini, aku mau \emptyset -beli.
 house this 1s want AV-buy
 ‘This house, I want to buy.’

The active clause (15) resembles an Object voice clause, but it can be identified as an Active clause by a number of syntactic diagnostics, including the relative order of the modal/auxiliary and the Agent argument. In OV the modal/auxiliary is followed by the Agent, but in AV the Agent must be

followed by the modal/auxiliary, as is the case in (15). This distinction is important when considering surface variability of the prefix *meN-*: when a modal/auxiliary does not occur in a clause such as (14a) or (15), the clause is ambiguous between OV and an AV clause with a topicalized Theme.

3.2.3. On the variability of *meN-* and *ber-*

Both *ber-* and *meN-* have several allomorphs. With roots that begin with a consonant, *ber-* is sometimes realized without its last segment, as in *be-serta* ‘be alongside,’ *be-kerja* ‘work’; it is realized as *bel-* with only one root, *bel-ajar* ‘study, learn.’ The final segment of *meN-* is a homorganic nasal that depends on the following segment (see Sneddon et al. 2012 for a description of nasal changes in *meN-*). Additionally, the prefix may be truncated and realized as *N-* instead of *meN-*.³ For both *meN-* and *ber-*, allomorphic alternations are not included in this discussion of variability. However, it is worth noting here that in varieties of Indonesian where both *meN-* and *N-* occur, there is a potential three-way contrast in surface forms among *meN-*, *N-*, and the bare form of the verb (see recent studies about variation among these three forms, including discussion of *N-* resulting from phonological truncation of *meN-*, i.e. Hidajat 2013; Kushartanti 2014; Kurniawan 2015, 2016). In this chapter, I treat both overt variants (*meN-* and *N-*) as the same syntactic object, and consider the optionality between an overt prefix (represented as *meN-*) and a non-overt prefix. The reason for this is that both *meN-* and *N-* are subject to the same syntactic constraints, and conditions that require or disallow one form also require or disallow the other.

By claiming that *meN-* and *ber-* are “optional” or “variable,” I assume that presence or absence of these morphemes does not change truth conditions, affect grammaticality or result in other semantic differences. In other words, the prefixed form and the unprefixed form are semantically equivalent, and yield the same interpretation. This is confirmed by diagnostics for semantic equivalence, some of which are presented here. First, contrastive sentences with the conjunction (*te)tapi* ‘but’ can draw out differences between two semantically similar verbs as in (16):

³In some varieties of Indonesian, particularly in colloquial Jakartan Indonesian, the active voice prefix is *N-* rather than *meN-*. This segment also undergoes phonological alternations depending on the initial segment of the stem; these are mostly the same changes to the nasal segment in *meN-*, with a few differences; see Wallace 1979; Sneddon 2006; Hidajat 2013; Kurniawan 2015.

- (16) Dia mem-andang tapi tidak me-lihat.
 3s AV-look but Neg AV-see
 'He looks but he doesn't see.'

The same verb occurring with a prefixed form and a bare form is not felicitous in this type of clause, showing that there is no contrastive difference in meaning between the prefixed form and the bare form:

- (17) #Dia mem-andang tapi tidak pandang.
 3s AV-look but Neg look
 'He looks but he doesn't look.'
- (18) #Dia bahasa Indonesia, tapi aku ber-bahasa Indonesia.
 3s language Indonesian but 1s MV-language Indonesia
 'She spoke Indonesian, but I spoke Indonesian.'

Since the same root occurs in both forms however, semantic differences between the two may be subtle. The following clauses draw out aspectual distinctions in different forms of a verb (with the same root). The adverb *tetap* 'still, constantly' describes a persistent state or activity, whereas *tiba-tiba* 'suddenly' is compatible with punctiliar or instantaneous achievements. Both prefixed and bare forms are possible with these adverbs:

- (19) Mereka tetap { bohong / ber-bohong }.
 3p still lie MV-lie
 'They are still lying.'
- (20) Mereka tetap { curi / men-curi } perhatian.
 3s still steal AV-steal attention
 'They are still stealing the attention.'
- (21) Tiba-tiba aku { peluk / ber-peluk } tubuh kakak.
 suddenly 1s hug MV-hug body older.sib
 'Suddenly I hugged my older sister.'
- (22) Wanita itu tiba-tiba { lompat / me-lompat } dari kursi.
 woman that suddenly jump AV-jump from chair
 'The woman suddenly jumped up from her chair.'

Where *meN-* and *ber-* are described as optional or variable, the prefixed form and the bare form

pattern together in semantic diagnostics such as these, given a verb root and clause type in which a prefixed form and a bare form are both possible.

It is worth noting that these distributional patterns are found to be more reliable than speaker intuitions about the meanings of functional elements such as verbal prefixes. For example, when asked whether there is any difference between *bohong* and *berbohong* in (19), my consultants offered a variety of responses, including: the bare form is used in anger; the bare form is simple reporting while the prefixed form will be followed by more details; the prefixed form is a description of (the subject's) true nature, while the bare form is neutral; the prefixed form is proper, while the bare form is everyday speech; etc. Anecdotally speaking, these replies sometimes have an ad hoc quality, and consultants often disagreed with one another. Under the account that I develop here, such inconsistent responses are not altogether unexpected. When a clause allows variable *ber-* (or *meN-*), the option between an overt prefix and a non-overt prefix is influenced by a variety of factors, both grammatical and extra-grammatical. See Section 3.5 for further discussion.

3.3. Previous analyses of *meN-* and *ber-*

There is disagreement in the literature about the functions of *meN-* and *ber-* in Indonesian (and Malay). I survey various analyses in this section; each addresses a particular pattern or contrast exhibited by the prefixes, but I show that exceptions exist for every generalization. Furthermore, few have addressed optional realization as part of the analysis. Generally speaking, no analysis completely accounts for all the syntactic properties and distribution of *meN-/ber-* in all contexts.

I begin with the prefix *meN-*, which I have labeled an active voice (AV) marker. Early grammars identified *meN-* as a marker of active voice or agentive voice, in opposition to the passive voice marker *di-*, and this view is held by many contemporary authors as well (Macdonald and Dardjowidjojo 1967; Dardjowidjojo 1978; Sneddon 1996; Arka and Manning 1998; Cole and Hermon 1998; Arka 2000; Voskuil 2000; Aldridge 2008; Cole et al. 2008a; Son and Cole 2008; Sneddon et al. 2012). In active voice, the thematic Agent or the external argument of the verb occurs as the preverbal subject of the clause (23). This is in opposition to passive clauses, in which an internal object or

Theme occurs as subject, and the external argument or Agent is optionally expressed in a PP (24).

- (23) Rika me-ringkas pelajaran pagi ini.
Rika AV-summarize lesson morning this
'Rika summarized a lesson this morning.'
- (24) Pelajaran ini di-ringkas (oleh Rika) pagi ini.
lesson this PV-summarize by Rika morning this
'The lesson was summarized (by Rika) this morning.'

Other authors view these prefixes as markers that indicate which argument is the trigger or focused element (Guilfoyle et al. 1992; Gil 2002; Wouk 2004). Under this view, *meN-* marks Agent/Actor focus or Agent/Actor trigger in (23), while *di-* marks Patient/Theme/Undergoer focus or Patient/Theme/Undergoer trigger in (24). This approach places Indonesian within the larger context of Austronesian languages that mark the most prominent argument in the clause (the trigger) with verbal morphology (see Kroeger 1993; Rackowski 2002; Rackowski and Richards 2005; and others on verbal morphology in Tagalog).

There are exceptions to both views. *MeN-* can also occur on unaccusative verbs that do not have an external argument or Agent/Actor. For example, the intransitive verbs *mencair* 'melt,' *menguning* 'turn yellow,' and *meledak* 'explode' occur with *meN-* in (25-27), yet these clauses have only an internal argument or Undergoer:

- (25) Es di Kutub Utara sedang men-cair.
ice Loc Pole North Pole AV-liquid
'Ice at the North Pole is melting.'
- (26) Ini padi yang mulai meng-uning.
this rice.field C.Foc begin AV-yellow
'This is the rice field that is beginning to turn yellow.'
- (27) Tiba-tiba gunung api me-ledak.
suddenly volcano AV-explode
'Suddenly the volcano exploded.'

Intransitive verbs that are affixed with *meN-* include a wide range of semantic types. The list below shows that intransitive verbs prefixed with *meN-* do not seem to form any pattern:

(28) Intransitive verbs that take *meN-* (list not exhaustive)

me-lapor ‘report’
me-ledak ‘explode’
me-letus ‘burst’
me-luncur ‘slide’
mem-beku ‘freeze’ (also ber-beku ‘freeze’)
men-angis ‘cry’
men-cair ‘melt’
men-darat ‘land’
men-didih ‘boil’
meng-eluh ‘complain’
meng-ering ‘dry’
meng-inap ‘spend the night’
meng-utara ‘go north’
men-ikah ‘be married’
men-jadiah ‘become’
meny-anyi ‘sing’
meny-erah ‘surrender’

For Malay, Soh and Nomoto (2011) argue that all intransitive *meN-* verbs are unergative, and that unaccusatives do not occur with *meN-*. Soh and Nomoto assume that both A-bar movement and A movement over a *meN-* verb requires a bare verb form (or the movement causes *meN-* deletion, in their terminology). According to the Unaccusativity Hypothesis (Perlmutter 1978), the Theme of an unaccusative verb begins as an internal argument; when this argument undergoes A movement to subject position, it crosses the verb; this should cause *meN-* deletion. In contrast, the sole argument of an unergative verb is an external argument, which does not cross the verb when it raises to subject position; this does not cause *meN-* to delete. The implication is that intransitive verbs that occur with *meN-* must only have an external argument, i.e. unergatives. This proposal attempts to unify the various types of predicates that occur with *meN-*: *meN-* always selects an external argument,

whether it occurs on a transitive or intransitive verb.

However, in Section 3.4, I argue that only A-bar movement affects the realization of *meN-*, and A-movement does not; see evidence presented in that discussion. This weakens the strict correlation between *meN-* and unergativity. Additionally, Soh and Nomoto's claim is surprising for a number of Indonesian *meN-* verbs in (28). For example, change of state predicates with external causation, such as *membeku* 'freeze,' *mencair* 'melt,' and other predicates such as *meletus* 'burst,' *menjadi* 'become' are cross-linguistically unaccusative. Furthermore, the syntactic diagnostics used to differentiate unergatives from unaccusatives do not hold for Indonesian. For example, Soh and Nomoto (2011) observe that the single argument of an unaccusative verb may occur in a postverbal position, whereas the single argument of an unergative must occur in a preverbal position. For Indonesian, many verbs, including those with *meN-*, allow subjects to occur postverbally, with the resulting word order V(O)S; I assume that VP-raising accounts for this variation in word order. I proceed without the assumption that *meN-* verbs always select an external argument in Indonesian.

Returning to the intransitive verbs in (28), these cannot take an object. However, many of these verbs allow *meN-* to occur simultaneously with an applicative suffix (*-kan* or *-i*). In this case the verb is transitive.

- (29) Matahari men-cair-kan es batu.
sun AV-liquid-Appl ice rock
'The sun is melting ice cubes.'
- (30) Ibu meng-uning-kan nasi dengan kunyit.
Mother AV-yellow-Appl rice with turmeric
'Mother colors the rice yellow with turmeric.'
- (31) Mereka gagal me-ledak-kan sasaran.
3p fail AV-explode-Appl target
'They failed to explode the target.'

This pattern is not only seen with unaccusatives. Intransitive unergatives can also be made transitive if both *meN-* and *-i* occur:

- (32) Anak-anak me-lompat.
 child-Redup AV-jump
 ‘The children jumped.’
- (33) Anak-anak me-lompat-i pagar sekolah yang tinggi.
 child-Redup AV-jump-Appl fence school C.Foc tall
 ‘The children jumped (over) the tall school fence.’

It is clear that *meN-* occurs on both transitive and intransitive verbs. Yet some authors have claimed that *meN-* marks transitivity, or both (active) voice and transitivity (e.g. Chung 1976; Saddy 1991; Cole and Hermon 1998). This view is based on several observations that correlate *meN-* with (in)transitivity. First, while intransitive verbs take several different prefixes (*meN-*, *ber-*, *ter-*, *ke-an*), almost all transitive verbs occur with *meN-* (Sneddon 1996). There are important exceptions that defy this generalization: see examples of *ber-* verbs that take complements in (60-63). A different observation correlates *meN-* with intransitivity: Sneddon (1996) notes that *meN-* is not optional on intransitive verbs; e.g. the verbs in (28) must occur with *meN-* prefixation. While this is the case in Standard Indonesian, it does not hold in colloquial spoken Indonesian; most intransitive verbs in (28) may occur without *meN-*, e.g.:

- (34) Aku { lapor / me-lapor } jam 9.
 1s report AV-report hour 9
 ‘I reported at nine o’clock.’
- (35) Luka-mu akan { kering / meng-ering } pelan-pelan dari dalam.
 wound-2s will dry AV-dry slowly from interior
 ‘Your wound will dry slowly from the inside.’

Thus it is not the case that all intransitive *meN-* verbs must occur with the prefix.

Another correlation between *meN-* and transitivity is observed by Fortin (2006): a “semantically transitive” verb that allows optional realization of *meN-*, cannot occur in bare form if the object is not expressed.

- (36) Ali sedang (mem)-baca buku itu.
 Ali Prog AV-read book that
 ‘Ali is reading (that book).’

- (37) *Ali sedang baca.
 Ali Prog read
 ('Ali is reading.') (Fortin 2006:51, ex. 13)

According to Fortin, the bare form of the verb is disallowed in (37) because *baca* is a semantically transitive verb without an overt object. This is unexpected if *meN-* only marks voice (or Agent focus/trigger). For my consultants however, sentences such as (37) are unremarkable in colloquial Indonesian, even if they are not acceptable in standard Indonesian. Semantically transitive verbs do occur in bare form, without an overt object, in informal written sources as well. Both of these observations regarding transitivity, then, appear to be tendencies rather than grammatical principles. Given the grammaticality of (37) in colloquial Indonesian, the correlation between an overt object and *meN-* is weakened.

A different pattern, also related to transitivity, is that a transitive verb with an object may occur with or without *meN-*, but if the verb occurs with an object clitic, *meN-* is obligatory (Musgrave 2001; Sneddon et al. 2012; also see Kroeger 2014 on Malay):⁴

- (38) Dia makan kue itu.
 3s eat cake that
 'He ate the cake.'
- (39) Dia meng-ambil kue itu lalu makan.
 3s AV-take cake that then eat
 'He took the cake and ate (it).'
- (40) Dia meng-ambil kue itu lalu { *makan-nya / me-makan-nya }.
 3s AV-take cake that then eat-3s AV-eat-3s
 'He took the cake and ate it.' (modified from Sneddon 1996:165)
- (41) Narti { *tunggu-ku / men-unggu-ku / *tunggu-mu / men-unggu-mu }.
 Narti wait-1s AV-wait-1s wait-2s AV-wait-2s
 'Narti is waiting for me/you.'

The relevant contrast is between (39), which does not occur with a clitic pronoun, and (40), which

⁴In Chapter 4 I argue that the clitic *-nya* is the functional head D in cases of possessor extraction. This example does not involve extraction, so I leave *-nya* glossed as 3s, consistent with assumptions in the literature. However, I note that the status of *-nya* is an open issue.

occurs with the clitic *-nya*. The same pattern is claimed to exist for first and second person clitics (41). Example (40) is particularly striking because *makan* ‘eat’ belongs to the set of verbs that rarely occurs with the prefix *meN-* (see 38). Unlike previous observations about the correlation between *meN-* and transitivity, this pattern appears to hold in colloquial Indonesian as well as Standard Indonesian; it is a deterministic condition on the realization of *meN-*.

Interestingly, the pattern only holds for clitics, not for free pronouns. Compare (40-41) with the following examples in which both the prefixed form and the *meN-* form occur with free pronouns:

- (42) Dia meng-ambil kue itu lalu { makan itu / me-makan itu }.
 3s AV-take cake that then eat that AV-eat that
 ‘He took the cake and ate it.’
- (43) Narti { tunggu aku / men-unggu aku / tunggu kamu / men-unggu kamu }.
 Narti wait 1s AV-wait 1s wait 2s AV-wait 2s
 ‘Narti is waiting for me/you.’

Note that the pronouns in (40-41) are called clitics because of their phonological attachment, but these clitics do not undergo movement at all. They occur immediately postverbally, in the same position as free object pronouns. This suggests that transitivity, case marking or thematic relations are not at issue. The question, then, is what differentiates these forms from free pronouns, and why this should require the prefix *meN-*.

I suggest that the reason has to do with the ability of personal pronouns and clitics to occur in different argument positions, including as possessors. For example, when the clitic *-nya* occurs on a verb without *meN-*, it is identical to a nominalized form of the verb (i.e. a gerundive form), which also occurs with the suffix *-nya*:

- (44) Makan-nya cepat sekali, kayaknya lapar.
 eat-D fast very apparently hungry
 ‘(His) eating was very fast, apparently (he was) hungry.’

This is less natural with 1 singular clitic *-ku* or 2 singular clitic *-mu*:

- (45) ?Makan-ku cepat sekali / ?Makan-mu cepat sekali
 eat-1s fast very eat-2s fast very
 ('My eating is very fast' / 'Your eating is very fast')

However, it is natural to use both a free pronoun or name with the suffix *-nya*, which results in nominalization (a gerundive form):⁵

- (46) Aku makan-nya cepat sekali. / Kamu makan-nya cepat sekali.
 1s eat-D fast very 2s eat-D fast very
 'My eating is very fast.' / 'Your eating is very fast.'
- (47) Tono makan-nya cepat sekali, kayaknya lapar.
 Tono eat-D fast very apparently hungry
 'Tono's eating was very fast, apparently (he was) hungry.'

Free pronouns can occur as possessive pronouns, but they do not form gerundives:

- (48) *Aku makan aku cepat sekali / *Makan aku cepat sekali }
 1s eat 1s fast very eat 1s fast very
 ('My eating is very fast')
- (49) *Makan dia cepat sekali.
 eat 3s fast very
 ('His eating was very fast.')

A bare verb with object clitic and a gerundive verb, then, have the same form. I suggest that *meN-* is required with an object clitic so that its form is distinct from a gerundive form. The same requirement does not hold for a free object pronoun because it is not homophonous with a gerundive form. This particular deterministic condition on *meN-*, therefore, arises from a syntactic condition (transitive verb with an object clitic) as well as an extra-syntactic factor, which is a constraint on homophonous forms: the bare verb with object clitic *makan-nya* is ruled out because it is identical to the gerundive form. The prefix *meN-* is required in order to disambiguate the two. In sum, the pattern does not arise because *meN-* marks transitivity, but because of competition between two homophonous forms.

⁵(46) resembles possessor extraction, in which the possessor moves out of its possessive DP and the possessum is suffixed with *-nya*; see Chapter 4. For the purposes of this discussion of homophony, nothing hinges on an analysis of possessor extraction in these examples.

Returning to analyses of *meN-*, Soh and Nomoto (2009, 2015) claim that *meN-* has aspectual effects in Malay. They note that *meN-*-prefixed verbs, when contrasted with bare verbs, have a progressive interpretation in Malay:

- (50) Harga elektrik turun.
price electricity fall
'The price of electricity fell.'
- (51) Harga elektrik men-urun.
price electricity meN-fall
'The price of electricity is falling.' (Soh and Nomoto 2009:151, ex. 7a-b)

The fall in price is reported to be abrupt in (50), whereas the falling in (51) is gradual. Indonesian speakers share the same intuition for the contrast between (50) and (51); the second sentence is interpreted to mean that the event is in progress. However, I suggest that this interpretation is strongest when the two sentences are contrasted with each other. In isolation, the aspectual meaning of *menurun* is not obligatory in (51). For example, both *turun* and *menurun* may occur with the adverbs *tetap* 'still, constantly,' *tiba-tiba* 'suddenly,' the perfective morpheme *telah*, and the progressive morpheme *sedang*.⁶ The ability of both forms to occur with different aspectual interpretations is illustrated below:

- (52) Harga elektrik tiba-tiba { turun / men-urun }.
price electricity suddenly fall AV-fall
'The price of electricity suddenly fell.'
- (53) Jumlah penduduk miskin telah { turun / men-urun } dengan tajam.
total population poor Perf fall AV-fall with sharp
'The total population of the poor has fallen sharply.'

It is also worth noting that *meN-* occurs on verbs that describe a single (punctual) event, e.g. see example (32) with *me-lompat* 'jump.' These facts show that progressive aspect can be encoded by aspectual morphemes such as adverbs, or else it can arise in context. In other words, the correlation between *meN-* and progressive aspect is taken to be a tendency rather than a property of the prefix.

⁶Soh and Nomoto 2009 note that *sedang* may occur with some verbs that describe transitory states.

Some authors argue that *meN-* has discourse-pragmatic functions (e.g. Wallace 1979; Kana 1983; Kaswanti Purwo 1989; Voskuil 1996; among others). For example, Kaswanti Purwo (1989) argues that *meN-* is used instead of a bare verb in the following discourse-pragmatic contexts: in narration, especially 3 person narration; for backgrounding; in responses; in requests for confirmation; in direct speech to an addressee; in indirect imperatives; in the presence of a modal. Voskuil (1996) argues that *meN-* is a voice marker, but that the bare form, rather than the *meN-*-prefixed form, has discourse-pragmatic properties and emotive associations. Importantly however, these discourse-related properties are non-deterministic factors for the realization of *meN-*. In other words, *meN-* is not strictly *required* by any of the contexts noted above; neither is the bare verb required. For example, a bare verb is also possible in most of the contexts that Kaswanti Purwo identifies as associated with *meN-*: bare verbs can be used in 3 person narration, for responses and requests, with a modal, etc. Likewise, Voskuil's claim that the bare form has emotive associations is also a non-deterministic tendency: a *meN-*-prefixed verb may always occur in emotive contexts instead of a bare verb, although *meN-* verbs are less frequent. These discourse-pragmatic properties are taken to be probabilistic patterns that emerge in usage.

Before turning to *ber-*, there are a number of roots that can take either *meN-* or *ber-*. This property is not predictable based on the semantics of the root:

- (54) Verbs that take either *meN-* or *ber-*
- ber-bekas / mem-bekas 'leave a mark/stain'
 - ber-dengkur / men-dengkur 'snore'
 - ber-derit / men-derit 'scream'
 - ber-harap / meng-harap 'hope, expect'
 - ber-ludah / me-ludah 'spit'
 - ber-nyanyi / meny-anyi 'sing'
 - ber-sebar / meny-ebar 'spread out'
 - ber-tari / men-ari 'dance'
 - ber-teduh / men-eduh 'take shelter'

The distinction between *meN-* and *ber-* (when comparing these affixed to the same root) is often reported to be transitive vs. intransitive, or between eventive vs. stative predicates (Soh and Nomoto 2009, 2011). Attributive uses of *ber-* verbs (as in *meja berbekas*) do fit this distinction, describing only states, e.g. *membekas* ‘leave a mark/stain’ vs. *meja berbekas* ‘table with a stain.’ However, the list of verbs in (54) shows that this distinction does not hold in all cases. I find that the *ber-* and *meN-* forms can have the same semantics, such that the pairs of verbs pattern together according to semantic diagnostics (see Section 3.2.3). For example, *berbekas* and *membekas* have the predicative meaning ‘leave a mark/stain.’ For many verbs, therefore, neither semantic class nor syntactic behavior differentiates between the prefixes *ber-* and *meN-*.

Turning now to *ber-*, which I have labeled a middle voice (MV) marker, authors agree that the functions of *ber-* are numerous, and often unpredictable. For example, Sneddon et al. (2012) list various functions for *ber-*: reflexive predicates (55), forming verbs whose meaning is ‘have [root]’ (56) or ‘use [root]’ (57), reciprocal relationship (58), and other semantically diverse predicates (59):

- (55) *ber-diam* (*diri*)
MV-silent self
‘be silent’
- (56) *ber-isteri*
MV-wife
‘have a wife’
- (57) *ber-topi*
MV-hat
‘wear a hat’
- (58) *ber-sahabat*
MV-close.friend
‘be close friends’
- (59) *ber-ada* ‘exist’

ber-enang ‘swim’

ber-ada ‘exist’ *ber-gembira* ‘be glad’

ber-henti ‘stop’

ber-hitam ‘be black colored’

ber-pikir ‘think’

ber-putar ‘turn, spin’

ber-sedih ‘be sad’

Like intransitive *meN-* verbs, *ber-* verbs include both unaccusatives (*berada* ‘exist’) and unergatives (*berenang* ‘swim’). Many adjectival predicates can also occur with *ber-* (*bergembira* ‘be glad,’ *berhitam* ‘be black colored’).

Ber- is usually claimed to mark intransitive verbs (Macdonald and Dardjowidjojo 1967; Sneddon 1996), even though we have already seen that some intransitives occur with *meN-*, *ter-* or *ke-an*. Hopper and Thompson (1980) identify *ber-* as marking constructions which have “low transitivity”: this is not only associated with the absence of an object, but also with the ability to be nominalized, expressing continuing states, and occurrence in reflexives and reciprocals (1980:278). However, while *ber-* mostly occurs on intransitive verbs, it is possible for a few *ber-* verbs to take complements. This includes nominal complements (60-61) as well as clausal complements (62-63):

- (60) Ayah saya ber-ternak ayam.
Father 1s MV-farm chicken
‘My father breeds chickens.’
- (61) Dia (ber)-main game komputer sampai larut malam.
3s MV-play game computer until late night
‘He played computer games until the middle of the night.’ (modified from Soh 2013:169)
- (62) Aku (ber)-harap bahwa mereka akan pulang hari ini.
1s MV-expect 3p will go.home. day this
‘I expect they will come home today.’
- (63) Kita (ber)-setuju bahwa belajar bahasa asing itu penting.
1p MV-agree C MV-study language foreign that important
‘We agree that studying a foreign language is important.’

In brief, the overall picture that emerges from this survey is that *meN-* and *ber-* both escape simple characterization as markers of voice, focus/trigger or transitivity. The functions of *meN-* and *ber-* can be descriptively generalized, with a number of exceptions which must be memorized, as

they are not predictable. Although *meN-* may be associated with (in)transitivity in Standard Indonesian, the colloquial spoken Indonesian that is discussed here does not place the same constraints on the optionality of *meN-*. I have also shown that while *meN-* is descriptively associated with discourse-pragmatic properties, these are tendencies that arise from usage, not from deterministic factors.

For the purposes of this chapter, which is focused on variable and non-variable realizations of the prefixes, I treat *meN-* and *ber-* as morphemes that occupy a particular structural position, rather than having a single grammatical function. Both *meN-* and *ber-* occur in the position of the functional head Voice: the main evidence for this is that *meN-* and *ber-* are sensitive to A-bar movement that proceeds through the edge of phases, which suggests that both *meN-* and *ber-* are phase heads. I argue for this view in the next section.

3.4. Deterministic factors affecting the realization of *meN-* and *ber-*

In this section I discuss contexts in which *meN-* and *ber-* are not optional. These are deterministic factors, such that either a prefixed form or a bare verb is required in the specified environment.

3.4.1. Nominal A-bar movement

A well-known syntactic pattern in Indonesian is that nominal A-bar movement over an active verb requires a bare verb instead of a prefixed verb. This has been observed by many previous authors for *meN-* (Wallace 1979; Kana 1983; Kaswanti Purwo 1989; Saddy 1991; Cole and Hermon 1998; Soh 1998; Voskuil 2000; Nuriah 2004; Wouk 2004; Cole and Hermon 2005; Fortin 2006, 2009; Donohue 2007; Aldridge 2008; Arka and Manning 2008; Cole et al. 2008b; Sato 2008b, 2012; Yanti 2010; Nomoto 2013). In Chapter 2 I discuss properties of nominal extraction; see that discussion for a summary of general properties of A-bar movement in Indonesian, including syntactic islands and pseudo-cleft structure.

The relationship between A-bar movement and verbal morphology is summarized in Cole and Hermon's Generalization (which was originally formulated for Malay):

(64) Cole and Hermon's Generalization

The obligatory omission of *meng-* with verbs that would otherwise permit *meng-* indicates the movement of an NP argument over the *meng-* + verb. (Sato 2008b, citing Cole and Hermon 1998:231)

As illustrated below, A-bar movement of a DP from the complement of a transitive verb is only possible when the verb does not bear *meN-*. The prefix is usually optional on these verbs:

- (65) Ayah sedang { per-baik-i / mem-per-baik-i } genteng.
Father Prog Caus-good-Appl AV-Caus-good-Appl roof.tiles
'Father is fixing the roof tiles.'
- (66) Apa yang Ayah sedang { per-baik-i / *mem-per-baik-i } ___?
what C.Foc Father Prog Caus-good-Appl AV-Caus-good-Appl
'What is Father fixing?'
- (67) Siapa yang ___ sedang { per-baik-i / mem-per-baik-i } genteng?
who C.Foc Prog Caus-good-Appl AV-Caus-good-Appl roof.tiles
'Who is fixing the roof tiles?'
- (68) Aku { pakai / me-makai } HP Nokia.
1s use AV-use cell.phone Nokia
'I use a Nokia cell phone.'
- (69) HP yang aku { pakai / *me-makai } ___
cell.phone C.Foc 1s use AV-use
'the cell phone that I use'
- (70) orang yang ___ { pakai / me-makai } HP Nokia
person C.Foc use AV-use cell.phone Nokia
'the person that uses a Nokia cell phone'

The movement of an object DP over the verb requires a bare verb in (66) and (69). When DP movement occurs from subject position, as in (67) and (70), it does not affect the realization of *meN-*: both the prefixed form and the bare form are possible. Thus, subject extraction differs from object extraction because the nominal does not cross the verb, so that the movement does not have consequences for verbal morphology.

The same pattern is observed with nominal extraction from a *ber-* clause, as observed by Soh

(2013). Recall that a few *ber-* verbs can take complements; if a DP is extracted over these verbs, *ber-* is disallowed and the verb must appear in bare form:

- (71) Dia { main / ber-main } game komputer sampai larut malam.
 3s play MV-play game computer until late night
 ‘He played computer games until the middle of the night.’ (modified from Soh 2013:169)
- (72) Apa-kah yang dia { main / *ber-main } ___ sampai larut.malam?
 what-Q C.Foc 3s play MV-play until midnight
 ‘What did he play until midnight?’
- (73) Siapa yang ___ { main / ber-main } game komputer sampai larut malam?
 who C.Foc play MV-play game computer until late night
 ‘Who played computer games until the middle of the night?’
- (74) Aku { harap / ber-harap } mereka akan pulang hari ini.
 1s expect MV-expect 3p will go.home day this
 ‘I expect that they will come home today.’
- (75) Siapa yang kamu { harap / *ber-harap } ___ akan pulang hari ini?
 who C.Foc 2s expect MV-expect will go.home day this
 ‘Who do you expect will come home today?’
- (76) Siapa yang ___ { harap / ber-harap } mereka akan pulang hari ini?
 who C.Foc expect MV-expect 3p will go.home day this
 ‘Who expects that they will come home today?’

The prefix *ber-* is disallowed whether the DP moves from object position as in (72), or from an embedded clause in long distance movement (75). Similar to *meN-*, subject extraction does not affect the realization of *ber-*. While the relationship between nominal extraction and *meN-* is well established, the data above show that Cole and Hermon’s Generalization should be extended to *ber-* verbs as well.

The loss of optionality in *meN-/ber-* in case of A-bar movement has two implications. First, the obligatory null morphology (i.e. bare verb) in case of A-bar movement is taken to be an instance of morphological wh-agreement. Cross-linguistically, wh-agreement occurs in structures derived by A-bar movement (wh-movement), such as constituent questions and relatives. The resulting morphology usually occurs on V or C (Reintges et al. 2006; see Chapter 2 for a full summary of

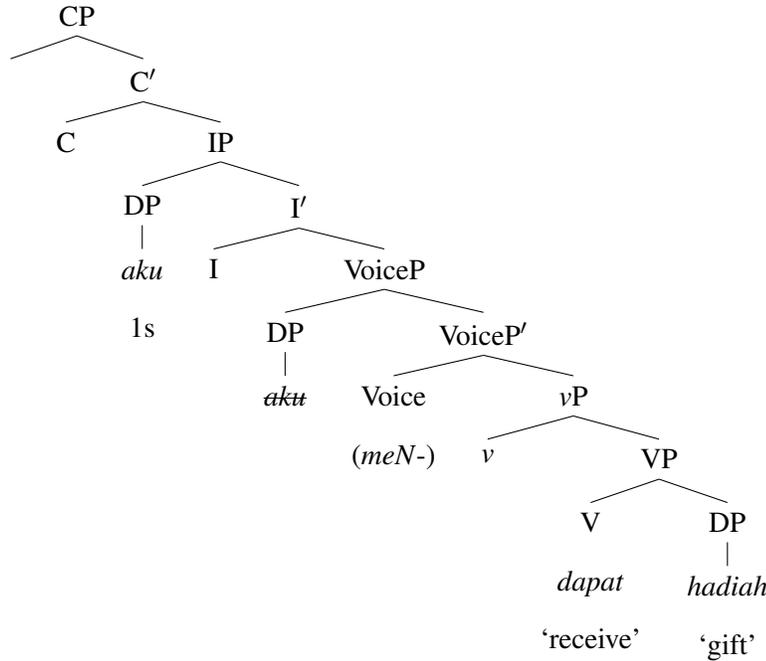
cross-linguistic generalizations about wh-agreement). Throughout this thesis, I argue that A-bar movement has morphological consequences on complementizers, verbs and nominals. Extraction over C requires that the complementizer must be null (for intermediate C) or must occur as *yang* (for highest C). Possessor extraction over D requires that D occurs overtly as the suffix *-nya* must be overt. In the verbal domain, wh-agreement is required when a DP undergoes A-bar movement from the complement of the verb to a high position in the clause. The form of wh-agreement in this case is a null verbal prefix instead of *meN-/ber-*.

The morphemes that are involved in Indonesian wh-agreement share an interesting property: they are optionally realized in many environments, but not in case of nominal extraction. In other words, A-bar movement is marked by a *loss* of optionality. Generally speaking, nominal A-bar movement is a deterministic factor in the realization of C, D and *meN-/ber-*.

The effect of nominal A-bar movement on *meN-/ber-* also implicates syntactic phases. The structure of a clause with a *meN-* prefixed verb (77) is given in (78).

- (77) Aku { dapat / men-dapat } hadiah.
1s receive AV-receive gift
'I received a gift.'

- (78) Clause structure with *meN-*



The extended verbal structure of the clause includes both vP and VoiceP. Verbal prefixes are hosted in a functional head: *v* hosts the causative morpheme *per-*, and Voice hosts the voice morphemes *meN-*, *ber-*, *di-*, and the null Object voice prefix.⁷ Following Pylkkanen 2002, 2008; Harley 2013; Legate 2014, the external argument is introduced in the specifier of VoiceP. This external argument is raised to the preverbal position of grammatical subjects in SpecIP, as shown in (78).

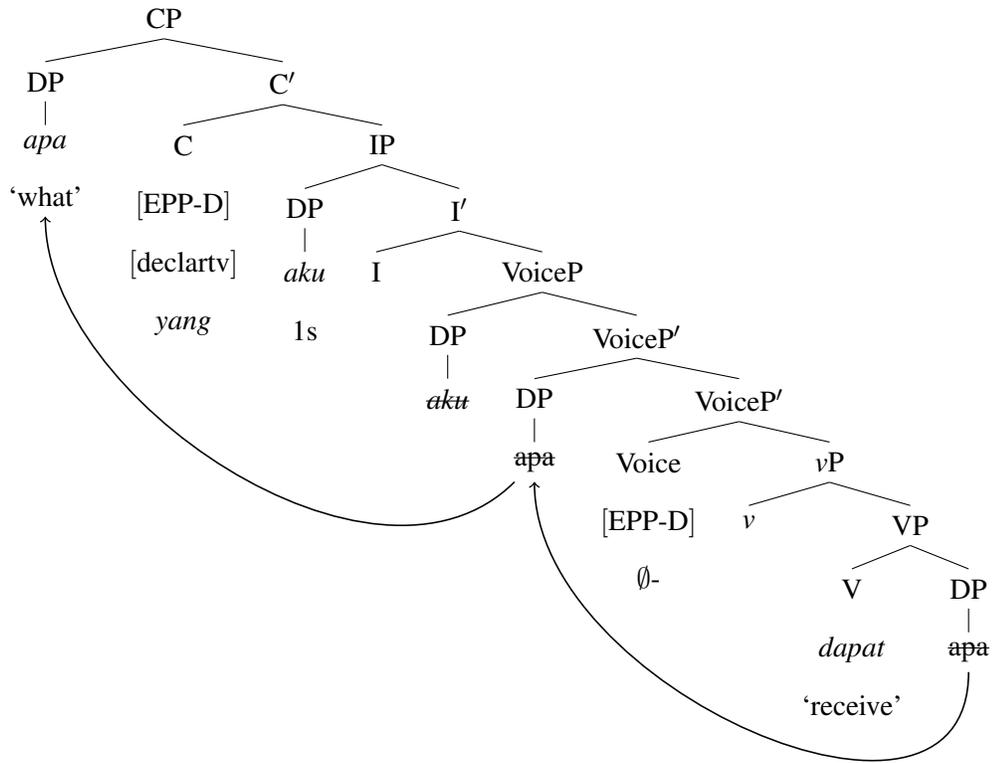
The other morphemes involved in wh-agreement are assumed to be phase heads: complementizers and the prefix *-nya* are taken to be phase heads C and D, respectively. The phase head, which attracts a nominal to its specifier for extraction, is specially marked as a result of this movement. Likewise, I take *meN-/ber-* to occur in the position of a phase head, Voice. Successive-cyclic A-bar movement from the complement of the verb must proceed through the specifier of VoiceP (suggested in Chomsky 2000, 2001, 2008 for vP; also see van Urk and Richards 2015 for recent work showing that successive-cyclic movement proceeds through the edge of every verb phrase and the edge of every clause). This movement results in special (null) morphology on the head Voice.

⁷The prefix *meN-* may co-occur with the causative morpheme *per-*. *Ber-*, on the other hand, cannot co-occur with *per-*, which leaves open the possibility that *ber-* occurs as a combined functional head, VoiceP/vP. This possibility does not affect the analysis here since this combined projection remains the highest verbal projection; successive-cyclic movement would pass through the edge of the VoiceP/vP. I leave this as an open question.

An example of object extraction is illustrated in the tree structure below:

(79) Apa yang aku dapat ___?
 what C.Foc 1s receive
 'What did I receive?'

(80) A-bar extraction of object through Spec, VoiceP⁸



In case of object extraction through the edge of VoiceP, the phase head has an edge feature [EPP-D] which targets DPs for movement through its specifier. The effect of nominal A-bar movement through SpecVoiceP is that the functional head Voice cannot be spelled out as *meN-* or *ber-*. Instead, the verb must occur bare, without a voice prefix. Precisely speaking, wh-agreement results from the [EPP-D] feature on a phase head; [EPP-D] on Voice is a deterministic factor in the realization of *meN-/ber-*. (See also Chapters 2 and 4 for a derivational account of nominal extraction through the

⁸For ease of exposition, movement of the wh-phrase *apa* is shown here instead of Operator movement. See Chapter 2 for discussion on the structure of clefts and Operator movement in nominal extraction.

edge of CP and DP, respectively.)

Excursus: Is *meN*- sensitive to A-movement?

Some authors have taken Cole and Hermon's Generalization to apply to *both* A-bar movement and A-movement, whereas I claim that *meN*- is only sensitive to A-bar movement. I address this issue here, because it has implications for the wh-agreement analysis that I propose. If both A-bar movement and A-movement result in a bare verb, this weakens the case for wh-agreement, which is a phenomenon specific to A-bar (wh-) movement.

The question is whether A movement over a *meN*- verb requires a null prefix (there is general consensus that this is the case for A-bar movement). Early work on the correlation between *meN*- and nominal movement in Saddy 1991 did not differentiate between A and A-bar movement, noting only that NP displacement over an active verb was not possible with *meN*-. Although much of the literature on this correlation examines the formation of questions and relatives (typical instances of A-bar movement), Cole and Hermon (1998) make the explicit claim that *meN*- is sensitive to both A and A-bar movement in Malay: "the loss of *meng-* is not restricted to wh-questions. Rather, it also occurs in object preposing constructions and in relative clause formation and focus movement" (1998:232). "Object preposing" constructions are Object voice clauses, in which the internal argument (Theme) of the transitive verb occurs as preverbal grammatical subject, as in (81) (see also example 14).

- (81) Ali_i saya (**men-*)cubit t_i
Ali 1s meng-pinch
'I pinched Ali.' / 'Ali was pinched by me.' (Cole and Hermon 1998:232)

Both Chung 1976 and Cole and Hermon 2005 provide evidence that the Theme in Object voice occurs in the position of grammatical subjects, an A position. Note that the verb must occur in its bare form (without *meN*-), which Cole and Hermon attribute to A-movement of the object *Ali* across the verb. In later work on Indonesian relatives, Cole and Hermon (2005) further note, "In our analysis it is irrelevant whether a noun phrase moves across a verb by A-movement or A-bar

movement. The fact that both A-movement and A-bar movement result in the same morphological result (the omission of *meN-*) might be viewed as a virtue since the analysis provides a unitary analysis for a wide range of data or as a flaw since, as an anonymous referee points out, it is uncommon in the languages of the world for A and A-bar to result in the same morphological effects on the verb” (2005:85, footnote 17; the relevant movement is shown in their example 34).

The claim that *meN-* is sensitive to both A-bar movement and A movement has been adopted by other authors as well (e.g. Nuriah 2004; Soh and Nomoto 2011; Sato 2012; Soh 2013; Nomoto 2013). Sato (2012) extends Cole and Hermon’s analysis to Object voice clauses in Kendal Javanese, assuming that both A-movement and A-bar movement cause *meN-* to delete (2012:37).

If correct, this view weakens the case for the lack of voice prefix as a type of wh-agreement in Indonesian, since wh-agreement is cross-linguistically noted to be specific to A-bar movement. Furthermore, I have argued that wh-agreement in Indonesian is marked on phase heads when DP movement passes through the edge of the phase; but raising to subject position to satisfy EPP is not usually taken to be cyclic movement through a phase edge (although alternate analyses claim that passive vP, or all types of vP, are phases, e.g. Legate 2003; Bošković 2014).

Evidence from Object voice clauses supports the view that *meN-* is sensitive only to A-bar movement, and that the bare verb form in Object voice has nothing to do with A-movement of the Theme/object. Cole and Hermon assume that Object voice clauses such as (81) would occur with the active voice prefix *meN-*, if not for A-movement of the Theme over the verb. Transitive clauses and Object voice clauses, then, both have the same structure at the beginning of the derivation, at which point *meN-* is not yet ruled out. In an active clause, the external argument becomes the grammatical subject, but in an Object voice clause, it is an internal argument that becomes grammatical subject while the external argument remains in situ.⁹ It is important to note that unlike active clauses, which can be observed with and without *meN-*, Object voice clauses never occur with *meN-*. The ostensible reason, according to this view, is that the subject position must be filled in Indonesian, so a DP object cannot be left low in the clause, which rules out *meN-*. If the object remains low,

⁹Object voice also places restrictions on the set of nominals that may occur as external argument; see Chung 1976; Cole et al. 2008b.

and no movement crosses the verb, the prediction is that *meN-* will be able to occur in Object voice clauses.

I test the prediction with verbs that embed CP complements. In Indonesian, leaving the grammatical subject position empty is possible with a CP complement, which may be left low. This is illustrated with the adjectival predicate *ternyata* ‘obvious’ and the passive verb *dipastikan* ‘confirmed’:

- (82) %Bahwa Ali menang sudah ternyata.
 C Ali win already obvious
 ‘That Ali won is obvious.’
- (83) Sudah ternyata bahwa Ali menang.
 already obvious C Ali win
 ‘It is obvious that Ali won.’
- (84) %Bahwa kopi ini adalah kopi asalan belum di-pasti-kan.
 C coffee this Cop coffee original not.yet PV-certain-Appl
 ‘That this coffee is original coffee is not yet confirmed.’
- (85) Belum di-pasti-kan bahwa kopi ini adalah kopi asalan.
 not.yet PV-certain-Appl C coffee this Cop coffee original
 ‘It is not yet confirmed that this coffee is original coffee.’

Not all Indonesian speakers accept sentential subjects as in (82) and (84).¹⁰ For all speakers however, a CP complement can remain in its base position, leaving the subject position empty, as in (83) and (85).

Now turning to Object voice clauses, the prediction is that when the CP complement remains in a low position, *meN-* will be possible on the verb since no movement, A or A-bar, has occurred. In the following Object clause sentences, the verb takes a CP as its object:

- (86) %Bahwa Ali menang akan saya umum-kan.
 C Ali win will 1s public-Appl
 ‘I will announce that Ali won.’

¹⁰Most of my consultants reject (82), except one speaker from Jakarta; however, examples of sentential subjects are frequent in the literature and in Indonesian grammars, which suggest regional variation and formality affect their acceptability.

- (87) Akan saya umum-kan bahwa Ali menang.
will 1s public-Appl C Ali win
'I will announce that Ali won.'
- (88) *Akan saya meng-umum-kan bahwa Ali menang.
will 1s AV-public-Appl C Ali win
- (89) %Bahwa kopi ini adalah kopi asli sudah ku-pasti-kan.
C coffee this Cop coffee original already 1s-certain-Appl
'I already made certain that this coffee is original coffee.'
- (90) Sudah ku-pasti-kan bahwa kopi ini adalah kopi asli.
already 1s-certain-Appl C coffee this Cop coffee original
'I already made certain that this coffee is original coffee.'
- (91) *Sudah saya mem-asti-kan bahwa kopi ini adalah kopi asli.
already 1s AV-certain-Appl C coffee this Cop coffee original

These are unambiguous Object voice clauses: the auxiliaries *akan* and *sudah* occur before the Agent, which is not possible in active clauses.¹¹ Furthermore, the clitic pronoun *ku* does not occur in subject position, which shows that the subject position must be empty in (90).

The prediction is not borne out: even when no movement crosses the verb, an Object voice verb cannot occur with *meN-*, as shown in (88) and (91). Therefore A-movement of the internal argument to subject position is not the source of the null voice morphology in Object voice. Rather, all Object voice verbs must have a null prefix, independent of movement. Since Object voice clauses constitute the primary piece of evidence in support of the view that A movement causes *meN-* deletion, there is little reason to adhere to this view. I conclude that *meN-* is only sensitive to A-bar movement in Indonesian, and that *meN-* cannot occur when a DP undergoes successive-cyclic movement through the edge of VoiceP. The obligatory null voice marking that replaces *meN-* in case of nominal A-bar movement retains the properties of wh-agreement. Furthermore, the morpheme that occurs on the verb in Object voice is a stored form that is always phonologically null.

¹¹Note that examples like (81) are ambiguous between object voice clauses and active clauses with a topicalized Theme and (optionally) unpronounced prefix *meN-*. To avoid this ambiguity, I include auxiliaries in these examples, which occur before the Agent in Object voice, in order to disambiguate these from active clauses (in which the Agent occurs before auxiliaries).

3.4.2. Other syntactic and semantic constraints on optionality

Returning to other deterministic contexts that require a certain pronunciation of *meN-/ber-*, the prefix is disallowed when the verb occurs as a modal or auxiliary. This applies to a small set of verbs, for example:

- (92) Aku { *coba* / *ber-coba* / *men-coba* } *men-cari informasi itu di google.*
1s try MV-try AV-try AV-find information that at google
'I tried to find that information on Google.'
- (93) *Informasi itu* { *coba* / **ber-coba* / **men-coba* } *aku cari di google.*
information that try MV-try AV-try 1s find at google
'I tried to find that information on Google.'
- (94) Rika { *suka* / *meny-uka-i* } *es krim.*
Rika like AV-like-Appl ice cream
'Rika likes ice cream.'
- (95) *Anak-anak* { *suka* / **meny-uka-i* } *men-angis.*
child-Redup like AV-like-Appl AV-cry
'Children often cry.'

The roots *coba* 'try' and *suka* 'like' are not typically identified as modals or aspectual auxiliaries, but the sentences above demonstrate that they can occur either as main verb or in positions reserved for modals/auxiliaries. In (93), the relative order of *coba* and the pronominal Agent *aku* indicate this is an Object voice clause; only modals and auxiliaries may occur in this position. In (95), *suka* has an aspectual meaning (frequency), which is distinct from its verbal meaning, 'like.' Where these forms as used as modal or auxiliary, they do not occur as category V, and do not have a Voice projection. Therefore *coba* and *suka* cannot take the verbal prefixes *meN-/ber-* when used as modal or auxiliary.¹²

Another deterministic condition on *meN-* is required by change of state verbs. For example, the following verbs (among others) must always be prefixed with *meN-*:

¹²See Jeoung 2018 for other verbs that may occur in the position of modal/auxiliary in Indonesian.

- (96) Aku {rasa / me-rasa} gatal dan muka-ku {*merah / me-merah}.
 1s feel AV-feel itchy and face-1s red AV-red
 ‘I felt itchy and my face reddened.’
- (97) Es di Kutub Utara sedang {*cair / men-cair}.
 ice at Pole North Prog liquid AV-liquid
 ‘Ice at the North Pole is melting.’

The un-prefixed forms are only possible with non-change of state meanings, i.e. ‘I felt itchy and my face *was red*’ (96); ‘Ice at the North Pole is *currently liquid*’ (97). Unlike other deterministic contexts that affect *meN-*, change of state verbs require an overt prefix rather than a bare verb.

The roots of these verbs are basic states (in the sense of Embick 2004; see also Koontz-Garboden and Beavers 2017), while the inflected form with *meN-* entails a change of state. Following the general approach in Hale and Keyser 2002, I assume that these intransitive change of state predicates are derived by combining the root with a structural head *v* that introduces semantics of change or becoming: [_v v-BECOME merah] (see also Embick 2004 in which the head bearing the feature [FIENT] derives the same result). When this structure combines with Voice, the prefix *meN-* is required; for verbs that do not denote a change of state, the root is combined with another type of *v* head. In the absence of *v-BECOME*, *meN-* remains optional.

Relatedly, Soh and Nomoto (2015) argue that for Malay, *meN-* is incompatible with verbs that express states, in contrast to a bare verb (without *meN-*), which may describe a state. (Diagnostics for this claim are presented in Soh and Nomoto 2009, 2015). *MeN-* prefixation requires that the verb express an event; this is because *meN-* occurs as *v*, which can only combine with a VP that describes eventualities with stages. Under this view, *meN-* is a light verb (*v*), rather than a voice morpheme.¹³

Soh and Nomoto (2015) are not alone in arguing that the category of *meN-* is *v*. Cole et al. (2008b) also suggest that *meN-* is hosted in *v* (although they claim that *meN-* marks active voice). Nuriah (2004) proposes that *meN-* is a categorial marker, marking non-verbal roots as category V.

¹³ Another claim made by Soh and Nomoto (2015) concerns degree achievement verbs in Malay that occur with *meN-*. The claim is that when *meN-* occurs on degree achievement verbs, they must receive atelic interpretations in Malay, in contrast to bare verbs, which can either be telic or atelic. See Soh and Nomoto 2015:150-152 for diagnostics that tease apart telic and atelic verbs. However, I find that the relevant contrasts do not exist in colloquial Indonesian; *meN-* verbs may occur in both atelic and telic frames. I therefore do not pursue this issue with degree achievement verbs.

Although Nuriah does not specifically address change of state verbs (the main argument is concerned with thematic roles and their mapping to *meN-* and *-kan*), the proposal is that *meN-* is a syntactic head that determines the category of the root, and that *meN-* must be overt when the root is a non-V category. This account assumes that the category of the root is transparent, based on intuitions about its most prominent meaning when it occurs in isolation (without affixation). However, this assumption is not well supported for many roots, i.e.:

- (98) *mem-banjir-i* ‘to flood, overflow’ (listed as verbal root)
banjir ‘(a) flood’; ‘to flood’
- (99) *mem-bisik-kan* ‘to whisper’ (listed as noun root)
bisik ‘(a) whisper’; ‘to whisper’
bisik-bisik ‘whispers’; ‘whispering’; ‘to whisper (to each other)’

The unprefix roots in these examples can often occur as either a verb or noun, which renders it difficult to distinguish which roots must be morphologically marked with *meN-*. Furthermore, the causative morpheme *per-* occurs between *meN-* and the root:

- (100) *Produk ini be-kerja mem-per-merah bibir dengan sempurna.*
 product this MV-work AV-Caus-red lip with perfect
 ‘This product works to redden lips perfectly.’

This is unexpected if the prefix *meN-* is a verbalizing head that determines the category of the root; under standard assumptions, no morpheme should intervene between the root and its category-defining head. This is also an unexplained problem for the view that *meN-* is a light verb, hosted in *v*.

I do not assume then, that *meN-* is a categorizing head, a light verb or *v*-BECOME. If *meN-* occurs as the head Voice, as I have suggested, it is reasonable to suggest that adjacency between VoiceP and *v*P makes the head Voice sensitive to the content of *v*. The presence of the head *v*-BECOME requires that the verb is prefixed with *meN-*.

Finally, all direct imperative verbs in Indonesian must occur without *meN-* or *ber-*:¹⁴

- (101) { Pakai / *me-makai } pensil yang bagus untuk ujian.
use AV-use pencil that good for test
'Use a good pencil for the test.'
- (102) { Cari-(lah) / *men-cari-(lah) } teman yang baik hati.
find-Imper AV-find-Imper friend C.Foc good heart
'Find a friend who is kind.'
- (103) { Bicara-(lah) / *ber-bicara-(lah) } kepada dokter anak.
speak-Imper MV-speak-Imper to doctor child
'Speak to the pediatrician.'

Like imperative clauses in many other languages, these clauses have a surface structure that is similar to declarative clauses. Following analyses of imperatives in other languages, I assume that imperative clauses have an imperative Operator located within CP, which derives an imperative interpretation (Han 2000). This imperative Operator is incompatible with inflection on the verb (in the form of *meN-/ber-*).

By dividing deterministic from non-deterministic conditions on the realization of *meN-/ber-*, I have considered only those factors that are deterministic, requiring or disallowing the prefix *meN-/ber-* in the specified context. The deterministic factors discussed are: A-bar movement, modal or auxiliary use, imperatives and change of state verbs. The first three of these disallow *meN-/ber-*, while change of state verbs require *meN-/ber-*. In these environments, syntactic and semantic constraints are determinants in the overall picture of variability, and any description of the distribution of *meN-/ber-* must be incomplete without reference to these. In the next section I consider non-deterministic properties related to *meN-* and *ber-*.

¹⁴Passive verbs can be used for polite or indirect commands in Indonesian, and these are reported to have the same imperative force as direct commands like (101). Interestingly, the passive prefix *di-* may occur in an imperative context:

- (1) Mohon di-bawa semua keperluan untuk ujian.
request PV-carry all supplies for test
'Please bring all supplies for the test.'

To my knowledge, different types of imperatives have not previously been studied in Indonesian or related languages. I leave this as an area that requires much further research.

3.5. Variable contexts and non-deterministic factors

3.5.1. Variable deletion rule and its effects

The prefixes *meN-* and *ber-* are optional so long as deterministic factors do not apply. Potential examples of verbs with optional *meN-/ber-* are numerous. The following sentences illustrate verbs from a range of semantic classes, and also show the optional morphology in both matrix and embedded clauses.

- (104) Saudara akan {alami / meng-alami} keberhasilan!
brother will experience AV-experience success
'You will experience success!'
- (105) Aku baru ingat Ibu {kirim / meng-irim} surat kepada pak Djoko.
1s new remember mother send AV-send letter to Mr Djoko
'I just remembered Mother sent a letter to pak Djoko.'
- (106) Apa-kah para murid {per-hati-kan / mem-per-hati-kan} guru dengan baik.
what-Q Pl.human student Caus-attention-Appl AV-Caus-attention-Appl teacher with good
'Students paid attention to the teacher well.'
- (107) Ayah sedang {per-baik-i / mem-per-baik-i} genteng.
Father Prog Caus-good-Appl AV-Caus-good-Appl roof.tiles
'Father is fixing the roof tiles.'
- (108) Alif {rahasia-kan / me-rahasia-kan} hadiah-nya untuk-ku.
Alif secret-Appl AV-secret-Appl gift-D for-1s
'Alif hid the gift for me.'
- (109) Dia {bahasa / ber-bahasa} Indonesia, aku {bahasa / ber-bahasa} Inggris.
3s language MV-language Indonesian 1s language MV-language English
'She spoke Indonesian, I spoke English.'
- (110) Mereka adalah orang-orang yang {pendidikan / ber-pendidikan} baik.
3p Cop person-Redup C.Foc education / MV-education good
'They are people that are highly educated.'
- (111) Dia {sedih / ber-sedih} karena harus ter-pisah.
3s sad MV-sad because must Invol-divide
'She is sad because they must be separated.'

- (112) Kami hanya { putar-putar / ber-putar-putar } naik motor, tanpa tujuan jelas.
 1p.Excl only turn-Redup MV-turn-Redup ride motorbike without goal clear
 ‘We only rode around and around on the motorbike aimlessly.’

The proposed source of variability in *meN-/ber-* is the variable rule stated in the Introduction, repeated here below:

- (113) Variable deletion

meN- → ∅

ber- → ∅

The rule applies in the morphophonology, i.e. a post-syntactic component of the grammar, after the syntactic structure has undergone spellout to PF. When the rule is applied, the prefix *meN-* or the prefix *ber-* is deleted, i.e. phonologically unrealized. I have stated one rule for *meN-* and another for *ber-*, so that each rule may be applied independently. Collapsing the two rules into a single abstract rule would result in a simpler generalization (e.g. *Voice* → ∅); however, it is not clear that *meN-* and *ber-* occur in the same syntactic position, to the exclusion of passive prefix *di-* (which is not variable). Furthermore, the variability of *meN-* is potentially affected by a different set of factors than *ber-*; stating the rule in this way also allows a different rate of application for each prefix.

Note that the rule is not stated over an abstract morpheme or syntactic node (e.g. *Voice*); instead, the rule targets the phonological content of each morpheme. This assumes that the rule applies after phonological strings have been inserted into nodes in the structure supplied by the syntax. Where *meN-* and *ber-* are not inserted, optionality does not exist, since variability arises from deletion of a phonological string. Recall that three of the deterministic conditions require a null prefix: A-bar movement, modal/auxiliary and imperatives. These three conditions are deterministic in the sense that the variable rule in (113) never applies. Insertion of a phonologically null morpheme, or (in the case of modal/auxiliary) lack of a structural position for *meN-/ber-*, effectively removes the possibility of optionality in *meN-/ber-*. These conditions bleed the application of (113), since no phonological string can be targeted for deletion.

One additional consequence of implementing a variable rule in the morphophonology, is that optionality is not captured by a system of deterministic rules for vocabulary insertion (within the framework of Distributed Morphology; cf. Halle and Marantz 1993, 1994). To illustrate this point, briefly consider the analysis for Standard Indonesian in Sato 2012. Sato discusses *meN-* and its obligatory absence in case of A-bar movement. Under Sato’s account, *meN-* is usually inserted as an AV morpheme. If a nominal undergoes successive-cyclic movement through the edge of the verbal domain (v^*P , following Chomsky 2000, 2001, 2008), it deletes an uninterpretable feature on the phase head, which Sato calls a D-feature. When this feature is deleted, the prefix *meN-* can no longer be inserted during vocabulary insertion. Instead, a null exponent is inserted as an elsewhere item. This is captured by the vocabulary items in (114):

- (114) $meN-$ \leftrightarrow [v -- [+D]]
 \emptyset_{meN-} \leftrightarrow [v -- [...]] (Sato 2012:41)

This captures the non-realization of *meN-* in Standard Indonesian, which is not otherwise variable (note that Sato states that *meN-* is obligatory when A-bar movement does not obtain) (2012:33). The vocabulary items in (114), then, are deterministically inserted after spell-out of syntactic structure. This approach, if applied to colloquial Indonesian, leaves the variability of *meN-* unaddressed. What I propose is that the variable rule in (113) may be applied after vocabulary insertion; the rule does not target the null form, but it does target the overt form, which can be deleted.

Since the rule in (113) is variably applied by individual speakers, it is subject to a variety of potential factors that determine its (probabilistic) application, including: differences between speakers, including regional and sociolinguistic differences (interspeaker variation); formality, register and stylistic variation (intraspeaker variation); phonological and prosodic factors such as speech rate; and quantitative distribution that speakers use to learn what is “conventional usage” in Indonesian. These factors are non-deterministic, so that deletion or non-deletion of *meN-/ber-* has no consequences for grammaticality.

I now return to some properties that have been attributed to *meN-* or *ber-* in the literature,

but which are not deterministic (as demonstrated by counter-examples already presented in the sections above). I argue that these non-deterministic factors do indeed affect *meN-* and/or *ber-*; but their effect has to do with probabilistic application of a variable rule. In other words, these factors determine whether *meN-/ber-* has a tendency to occur overtly or not.

Consider the observation that semantically transitive verbs occur with *meN-* if the object is not expressed (Fortin 2006). I have already suggested that this is a non-deterministic property, based on the grammaticality of (115) in colloquial Indonesian:

(115) Variable deletion acceptable in colloquial Indonesian

Ali sedang baca.
Ali Prog read
'Ali is reading.'

If the bare verb is taken as a probabilistic tendency, two potential factors are predicted to affect the rate of deletion of *meN-* in this example. The prefix is deleted more frequently when an object is present; the prefix is also deleted more frequently in informal contexts. In both cases, the correlation between transitivity and *meN-* is indirect, but transitivity is not a deterministic factor.

Aspectual semantics have also been previously discussed as a non-deterministic factor for *meN-*. Soh and Nomoto (2009, 2015) proposed that *meN-* is a progressive aspectual marker, given a contrast in interpretation between the two sentences below:

(116) Harga elektrik turun.
price electricity fall
'The price of electricity fell.'

(117) Harga elektrik men-urun.
price electricity meN-fall
'The price of electricity is falling.' (Soh and Nomoto 2009:151, ex. 7a-b)

I have suggested that the aspectual meaning is not obligatory, but a tendency. Variable deletion also accounts for this: *meN-* is deleted less often when the sentence has a progressive aspect; alternately, *meN-* is deleted more frequently with punctual verbs or non-progressive aspect.

Finally, I also noted that a number of authors claim that *meN*-prefixed verbs contrast with bare verbs in their discourse-pragmatic functions (Wallace 1979; Kana 1983; Kaswanti Purwo 1989; Voskuil 1996). Since neither *meN*- nor the bare form is obligatory in any of these contexts, these cannot be deterministic factors. Rather, information structure, the organization of discourse and other extra-syntactic factors affect the probabilistic rate of deletion. For example, following observations in Kaswanti Purwo 1989, *meN*- is less frequently deleted in the following contexts: 3 person narration, clauses that provide background information, in direct speech to an addressee, and in the presence of a modal.

3.5.2. Other consequences of variable rule application

Variable application of a deletion rule also provides a new perspective on some of the unpredictable aspects of *meN*-/*ber*- distribution. In this section I suggest that the variable deletion rule is applied at different rates for different roots, and that speakers learn overall rates of deletion for a particular root, as part of their language capacity.

Below, I present idiosyncratic verbs that require *meN*- or *ber*-, and idiosyncratic verbs that never take either prefix. The following lists are representative of the diversity in each set, but are not exhaustive.

(118) Verbs that require *meN*- or *ber*-

Intransitive

men-angis ‘cry’

me-rantau ‘migrate’

berada ‘exist’

ber-isteri ‘have a wife, be married’

ber-jenggot ‘grow a beard, have a beard’

ber-sahabat ‘be close friends’

ber-topi ‘wear a hat’

Transitive

meng-ajar 'teach'
mem-per-alat 'manipulate'
men-urut 'obey'
men-anya-i 'question'
meng-gila-i 'to be crazy (about)'
men-eliti 'research'

Importantly, these are not change of state verbs; they do not fall into one or more syntactic classes.

(119) Verbs that cannot take *meN-* or *ber-*

Intransitive

bangun 'get up'
duduk 'sit'
datang 'come'¹⁵
hujan 'rain'
kencing 'urinate'
kerimis 'drizzle'
pergi 'go'
pulang 'go home'
selesai 'finish'
tinggal 'stay, live (at a location)'

Transitive

cinta 'love'
ikut 'join'
lupa 'forget'
masuk 'enter'
naik 'go up, board (vehicle), travel (by vehicle)'

¹⁵In very formal Indonesian, *berdatang* is used as a polite quotative, as when a royal person or deity speaks.

punya ‘have, possess’

tahu ‘know’

tiba ‘arrive’

suka ‘like’

One interesting property that is shared among the verbs in (119) is that they are among the most frequently occurring in speech. These frequent verbs, then, implement the deletion rule at a maximum rate.

Both lists include intransitive and transitive verbs. These are idiosyncratic in the sense that they do not form any pattern, and must be learned. (Some authors, e.g. Wallace (1979) have attempted to categorize these roots according to semantic class, with the result that some categories have only 1-2 verbs.)

The morphology on these roots appears to be deterministic: either they must occur with *meN-*/*ber-*, or they must occur bare. However, I suggest that these idiosyncratic verbs are also a product of the variable rule. Consider the possibility that with some roots, the deletion rule is applied so infrequently that *meN-/ber-* almost always occur; conversely, with other roots, the deletion rule is applied at a near-ceiling rate, so that the verb almost always appears without *meN-/ber-*. If an Indonesian speaker never hears a certain verb without *meN-/ber-*, then the speaker learns that the deletion rule is rarely or never applied. Verbs that never occur with *meN-/ber-* are likewise learned as verbs that always apply the deletion rule. The suggestion that I make is that these are judged to require or disallow the prefix because rates of variable deletion are learned as part of conventional usage.

Excursus: Wh-agreement applies to all verbs

One puzzling property of the verbs in (118) is that although they cannot occur with *meN-* alone or *ber-* alone,¹⁶ most can occur with both *meN-* and an applicative suffix (*-kan* or *-i*):

¹⁶When *tinggal* ‘stay, live’ is affixed with *meN-* its meaning is changed, ‘die.’ When the root *cinta* ‘love’ occurs with *ber-* its meaning is also changed, ‘have sex.’

- (120) Bare verbs that can be affixed with both *meN-* and *-kan/-i*
- mem-bangun-kan ‘rouse (someone)’
 meny-elesai-kan ‘finish (something)’
 men-inggal-kan ‘leave behind’
 men-duduk-i ‘occupy’
 menge-tahu-i ‘discover, ascertain’
 me-masuk-i ‘enter into a space’
 men-cinta-i ‘love’
 meny-uka-i ‘like’

The affixation does not have a predictable effect; in some cases, it is simply a change in transitivity (*membangunkan, menyelesaikan, meninggalkan*). For other roots, a related meaning results from affixation (*menduduki, mengetahui, memasuki*). The verbs *cinta* and *suka* are exceptions in that their affixed forms are not distinct from their bare forms, either syntactically or semantically.

Since some verbs never occur with the prefix *meN-*, and many more only optionally occur with *meN-*, the question arises whether *wh*-agreement still applies if the form of the verb after movement has not changed (in other words, the verb appears bare both before and after A-bar movement). I argue that *wh*-movement applies in all cases of A-bar movement, based on the behavior of verbs that require a prefix.

We have already seen that the prefix is optional with many roots. With the roots below, *meN-* is required in clauses without extraction:

- (121) Saiful { *jinjing / men-jinjing } tas-nya.
 Saiful hold AV-hold bag-D
 ‘Saiful held (up) her bag.’
- (122) Lani { suka / *suka-i / meny-uka-i } bunga mawar.
 Lani like like-Appl AV-like-Appl flower rose
 ‘Lani likes roses.’
- (123) Dia { *rangka / *rangka-i / me-rangka-i } bunga.
 3s arrange arrange-Appl AV-arrange-Appl flower

'He arranged the flowers.'

- (124) Tono { *gila / *gila-i / meng-gila-i } penyanyi itu.
Tono crazy crazy-Appl AV-crazy-Appl singer that
'Tono is crazy about that singer.'
- (125) Penjahat itu { *per-alat / mem-per-alat } se-orang gadis kecil.
criminal that Caus-tool AV-Caus-tool one-person girl little
'The criminal manipulated the little girl.'

While these roots usually require *meN-* to be well formed, they also require a null prefix when wh-agreement applies. Extraction of the object is only possible without *meN-*:

- (126) Apa yang Saiful { jinjing / *men-jinjing } ___ ?
what C.Foc Saiful hold AV-hold
'What did Saiful hold (up)?'
- (127) Bunga apa yang Lani { suka / suka-i / *meny-uka-i } ___?
flower what C.Foc Lani like like-Appl AV-like-Appl
'What flower does Lani like?'
- (128) Apa yang dia { *rangka / rangka-i / *me-rangka-i } ___?
what C.Foc 3s arrange arrange-Appl AV-arrange-Appl
'What did he arrange?'
- (129) Siapa yang Tono { *gila / gila-i / *meng-gila-i } ___?
who C.Foc Tono crazy crazy-Appl AV-crazy-Appl
'Who is Tono is crazy about?'¹⁷
- (130) gadis kecil yang penjahat itu { per-alat / *mem-per-alat } ____.
girl little C.Foc criminal that Caus-tool AV-Caus-tool
'the little girl that the criminal manipulated'

Many of the verbs that show this pattern (i.e. do not allow optional deletion of *meN-* in basic clauses) are those that take the suffix *-i* (*menyukai*, *merangkai*, *menggilai*). The function of this suffix is debated in the literature. It has been called a locative applicative, or a semantically empty morpheme that is required for well-formedness. For some authors, the fact that *suka-i*, *rangka-i* and *gila-i* are not possible in (127-129) has been taken to mean that the voice morpheme is the circumfix *meN-i*. Wh-agreement patterns suggest that this view is not correct, or at least that *meN-* and *-i* are

¹⁷The bare roots *gila* and *rangka* can occur as adjective 'crazy' and noun 'order, frame' respectively.

separate morphemes: these roots must occur with both *meN-* and *-i* when there is no extraction, yet in case of object extraction, *meN-* does not occur.

The examples above show that morphological wh-agreement does not depend on whether the surface form of the verb may occur with *meN-*; these are independent factors in Indonesian and the one does not depend on the other. Wh-agreement applies to all verbs, including those that optionally occur with *meN-/ber-* and those that never occur with *meN-/ber-*.

Returning to idiosyncratic roots, I have suggested that individual roots, which require *meN-* or *ber-* to be well-formed, do not belong to the category of deterministic conditions. Rather, rates of variable deletion are learned by speakers, so that with certain roots the prefix is always deleted or never deleted. One prediction of this view is that different speakers may have learned different rates of deletion for certain roots; for a particular root, one speaker may have a high rate of *meN-/ber-* deletion, while another speaker never deletes *meN-/ber-*.

This prediction is confirmed by conflicting judgments from consultants about the optionality of *meN-* and *ber-* with the roots listed below. The *meN-* form is always accepted, but not all consultants accept the bare form; this is evidence that the optionality of these prefixes arises from variable deletion. The forms that were only accepted by some consultants are marked with %.

(131) Verbs for which consultants disagreed

Transitive

% sapa / meny-apa ‘greet’

% jinjin / men-jinjin ‘carry (suspended by the hand)’

% rangkai / me-rangkai ‘arrange’

% per-satu-kan / mem-per-satu-kan ‘unify, bring together’

% per-masalah-kan / mem-per-masalah-kan ‘problematize’

% per-indah / mem-per-indah ‘beautify, adorn’

% percaya-kan / mem-percaya-kan ‘entrust’

% ber-daya-kan / mem-ber-daya-kan ‘empower’

% dalam-i / men-dalam-i 'fathom, understand deeply'

% curiga-i / men-curiga-i 'suspect'

Furthermore, for some of the roots in (131), consultants indicated that they were not certain whether the bare form was possible or not. This, too, is consistent with a variable deletion rule that is applied at a near-maximum for selected roots.

Since I have already shown that A-bar movement requires a bare verb, the question arises what effect A-bar movement will have, when crossing over verbs that always appear with *meN-/ber-* (118) and verbs for which consultants disagreed or were uncertain about (131). First, the examples below show A-movement with verbs that must occur with *meN-*:¹⁸

(132) Ibu Tina { meng-ajar / *ajar } matematika.
Mrs Tina AV-teach teach mathematics
'Mrs Tina teaches mathematics.'

(133) mata pelajaran yang Ibu Tina { *meng-ajar / ajar } —
Subject C.Foc Mrs Tina AV-teach teach
'the subject that Mrs Tina teaches'

(134) Penjahat itu { mem-per-alat / *per-alat } se-orang gadis kecil.
criminal that AV-Caus-instrument Caus-instrument one-person girl small
'The criminal used/manipulated a little girl.'

(135) Siapa yang penjahat itu { *mem-per-alat / per-alat } —?
who C.Foc criminal that AV-Caus-instrument Caus-instrument
'Who did the criminal use/manipulate?'

(136) Andi sangat { meng-gila-i / *gila-i } Arema.
Andi very AV-crazy-Appl crazy-Appl Arema
'Andi is very crazy about Arema (football club).'

(137) tim futbol yang Andi { *meng-gila-i / gila-i } —
team football C.Foc Andi AV-crazy-Appl crazy-Appl
'the team that Andi is crazy about'

(138) Didik { men-eliti / *teliti } monyet selama 4 tahun.
Didik AV-thorough thorough monkey for 4 year
'Didik researched monkeys for 4 years.'

¹⁸No verbs that require *ber-* are transitive, so A-movement does not occur. Change of state verbs, which require *meN-*, are also intransitive.

- (139) Apa yang Didik { *men-eliti / teliti } ____ selama 4 tahun?
 what C.Foc Didik AV-thorough thorough for 4 year
 ‘What did Didik research for 4 years?’

These sentences show that verbs which usually require *meN-/ber-* still occur bare in case of A-bar movement. This provides a contrast between a deterministic condition for *meN-/ber-* realization (morphological wh-agreement) and a non-deterministic operation (variable deletion). Even for verbs that are judged ill-formed without *meN-/ber-*, these must occur as bare verbs in case of A-bar movement. Unsurprisingly, for verbs that speakers disagree about, these verbs must also occur bare in case of A-bar movement:

- (140) Ibu-ku { %rangkai / me-rangkai } bunga tadi pagi.
 mother-1s arrange AV-arrange flower before morning
 ‘My mother arranged flowers this morning.’
- (141) Apa yang ibu-ku { rangkai / *me-rangkai } ____ tadi pagi?
 what C.Foc mother-1s arrange AV-arrange flower before morning
 ‘What did my mother arranged this morning?’
- (142) Kalau sudah me-nikah, tidak boleh { %per-masalah-kan / mem-per-masalah-kan }
 if already AV-marry Neg may Caus-problem-Appl AV-Caus-problem-Appl
 masa lalu.
 era past
 ‘If (one is) already married, (one) should not problematize the past.’
- (143) hal-hal yang dia { %per-masalah-kan / *mem-per-masalah-kan } ____
 things C.Foc 3s Caus-problem-Appl AV-Caus-problem-Appl
 ‘things that he problematized’
- (144) Andi telah { %percaya-kan / mem-percaya-kan } buku harian-nya kepada-ku.
 Andi Perf trust-Appl AV-trust-Appl book daily-D to-1s
 ‘Andi has entrusted her diary to me.’
- (145) buku yang Andi telah { percaya-kan / *mem-percaya-kan } ____ kepada-ku.
 book C.Foc Andi Perf trust-Appl AV-trust-Appl to-1s
 ‘the book that Andi entrusted to me’

I conclude the reason that idiosyncratic verbs always occur with *meN-/ber-* or without *meN-/ber-* is that the variable deletion rule is applied at a maximum rate for some roots; the rule is never

applied for other roots. This predicts that speakers learn overall rates of deletion, as well as root-specific rates of deletion. This is consistent with the fact that the roots which always occur bare (119) are among the most common in the language; frequent input helps speakers learn that the deletion rule is always applied and the bare form occurs.

Extra-syntactic factors (e.g. verb frequency and rates of deletion) are responsible for the contrast between verbs that allow optional *meN-* and *ber-*, and those that do not. In addition to the intraspeaker variability shown by these prefixes, there is some degree of interspeaker variability with regard to which roots may occur without *meN-*, and some roots for which speakers themselves have uncertain judgments. This too falls out from different rates of deletion for individual verbs.

3.6. Chapter summary

The overall thesis that I have advanced in this chapter is that the distribution of these functional morphemes in Indonesian can only be accounted for with a combination of 1) semantico-syntactic factors that affect the realization of these morphemes in a deterministic way, as well as 2) extra-syntactic factors that influence a post-syntactic variable deletion rule, such that application of the rule is probabilistic rather than deterministic. I demonstrated that the distinction between deterministic and non-deterministic factors is an important one, not only to capture the realization of *meN-* and *ber-*, but also for analyses of their respective functions. I have attempted to show that this approach makes sense of a number of previously observed properties of *meN-* and *ber-*.

Additionally, this chapter extends wh-agreement to the verbal domain. The well-known correlation between nominal extraction and null verbal morphology is re-framed here as morphological wh-agreement resulting from phase-based successive cyclic movement. I also argue against the assumption that A movement has the same effect on verbal morphology: by demonstrating that Object voice clauses independently require null verbal morphology, I showed that A-movement of the Theme to subject position does not cause *meN-* to occur as a null morpheme.

CHAPTER 4 Possessor Extraction

4.1. Introduction

4.1.1. Background

In this chapter I discuss possessor sub-extraction in Indonesian, with additional data showing possessor movement in Javanese and Madurese.¹ I show that a nominal possessor can be extracted from a possessive DP via A-bar movement, and surface as a clefted element at the left periphery of the clause, as in (2).

- (1) Rumah-(nya) Adi di-rata-kan kemarin.
house-D Adi PV-flat-Appl yesterday
'Adi's house was destroyed yesterday.'
- (2) orang yang rumah-nya ___ di-rata-kan kemarin
person C.Foc house-D PV-flat-Appl yesterday
'the person whose house was destroyed yesterday'

The suffix *-nya* is optional in possessive DPs (1), but A-bar movement requires *-nya* to occur on the possessum that remains in situ (2). I argue that pronunciation of *-nya* in these cases is a type of wh-agreement, which is a morphological reflex of A-bar movement. This chapter, then, extends the wh-agreement analysis that is argued for throughout this thesis: wh-agreement on complementizers is discussed in Chapter 2, and wh-agreement on verbs is discussed in Chapter 3. In this chapter I argue that wh-agreement also occurs in the nominal domain, when possessors are extracted from the DP.

Morphology marking wh-agreement within the DP is cross-linguistically unusual. In Chapters 2 and 3 I have discussed wh-agreement in relation to complementizers and verbal prefixes, respectively, which is consistent with the cross-linguistic generalization that morphological wh-agreement occurs in these two domains. Zaenen (1983) makes the claim that other domains do not

¹The variety of Indonesian discussed in this chapter is that of educated speakers living in urban areas East Java; some of these speakers were raised in other areas but attended university in East Java and use Indonesian on a daily basis. One of my consultants speaks only Indonesian as a first language; all other consultants also speak Javanese or Madurese. My Madurese consultants are from Bangkalan and Jember; however, all Madurese data cited in this paper are from the Bangkalan (western) variety. My Javanese consultants are from also from East Java.

show the effects of A-bar or wh-movement:

(3) Zaenen's generalization:

Only complementizers and verbal morphology are affected by wh-movement.

(Zaenen 1983, formulated in Watanabe 1996:177)

According to Zaenen, wh-agreement falls under a wider set of phenomena that she calls syntactic binding. Wh-agreement is marked on complementizers and verbs because these fall within the "binding domain," which is defined as projections that dominate a bindee but not the binder. I argue in this chapter that possessor extraction in Indonesian instantiates wh-agreement within the DP domain: when a possessor undergoes successive cyclic A-bar movement, it must escape its DP by movement through the specifier of DP. The result of this movement is obligatory pronunciation of D, which is realized as a suffix on the possessum. The Indonesian data, then, provides a pattern of wh-agreement that suggests Zaenen's generalization is incomplete, and that DPs can also show morphological reflexes of A-bar (wh-) movement.

Broadly speaking, possessor sub-extraction provides support for A-bar movement through the edge of DP, with movement proceeding cyclically via syntactic phases. This suggests that D is a phase head in Indonesian, in addition to phase heads Voice and C. In addition, long-distance extraction in Indonesian shows another cross-linguistically unusual pattern: all complementizers crossed by movement must show wh-agreement, but the highest C is marked differently than intermediate C (see Chapter 2). Marking A-bar movement through DP with the suffix *-nya* also means that movement through DP is marked differently than movement through clauses.

4.1.2. *Internal and external possession*

In this chapter I present data showing both internal and external possession. When an overt possessor occurs within a possessive DP as in (4), I refer to this as internal possession. When the possessor occurs outside of the possessive DP containing the possessum with which it has a possession relation, as in (5), this is a type of external possession:

- (4) Adik mem-baca [buku Siti].
 younger.sibling AV-read book Siti
 ‘Little brother read Siti’s book.’
- (5) Siapa yang adik baca [buku-nya ___]?
 who C.Foc younger.sibling read book-D
 ‘Who is it that little brother is reading (her) book?’

In other languages, many cases of external possession have been analyzed as possessor raising or possessor ascension (see Keenan 1972; Perlmutter and Postal 1983; Baker 1999; Landau 1999; Payne and Barshi 1999; Deal 2013a,b for various approaches.) In many of these cases, the external possessor appears in an A position (and is marked for Case in this position). In contrast, the type of external possession that I discuss in this chapter is distinct from possessor raising. Sentences such as (5) are derived by A-bar movement of the possessor, and the possessor always occurs in an A-bar position rather than an A position. I argue that possessor extraction in Indonesian is subject to the same principles and constraints that apply to A-bar movement of other nominals in this language. This type of possessor movement is also discussed in Ross 1986; Szabolcsi 1992; Coon 2009; Gavruseva 2000 for various languages. To my knowledge however, possessor extraction has not been analyzed in Indonesian and its closely related languages.²

Although I describe the DP-internal arguments as “possessum” and “possessor,” a broad range of semantic relations are possible in this genitive construction, including part-whole, entity-origin and container-content. The arguments may also be alienable or inalienable, concrete or abstract:

- (6) tangan-ku
 hand-1s
 ‘my hand’
- (7) pinggir jalan
 edge road
 ‘the side of the road’
- (8) Gubernur Jawa Timur
 governor Java east

²Much of this chapter appears in an earlier form in Jeoung (to appear).

‘Governor of East Java’

- (9) kemenangan tim putri Indonesia
victory team daughter Indonesia
‘the victory of the Indonesian women’s team’
- (10) inti ajaran agama Hindu
core teaching religion Hindu
‘the core of Hindu religious teachings’

In the following discussion, I present an analysis in which syntactic and morphological operations apply across this type of construction without making a distinction among these relations. The “possessor” in possessor extraction, then, is defined structurally, as a DP argument generated in the specifier of an NP (see Section 4.4.1).

4.1.3. Addressing variability in *-nya*

A key component of my analysis is to re-analyze the suffix *-nya* in light of its variable pronunciation in internal possession, and to extend this understanding to the obligatory nature of *-nya* in cases of possessor extraction. The analysis relies on data from the variety of colloquial Indonesian used in everyday situations, which often deviates from formal or Standard Indonesian. By formal Indonesian I refer to language that is spoken in official and professional settings, while Standard Indonesian is the prescriptive variety taught in schools and also used in writing. Colloquial Indonesian deviates from formal and Standard Indonesian in various ways, although speakers often switch between these depending on pragmatic context.

In Standard Indonesian, possessive DPs with an overt internal possessor do not occur with the suffix *-nya* on the possessum:

- (11) Rumah Adi di-rata-kan kemarin.
house Adi PV-flat-Appl yesterday
‘Adi’s house was destroyed yesterday.’

Rumah Adi is the possessive DP form cited in formal grammars (e.g. Dardjowidjojo 1978; Wolff et al. 1992; Sneddon et al. 2012). In colloquial Indonesian, however, the morpheme *-nya* can op-

tionally occur as a suffix on the possessum. This is illustrated below:

- (12) Rumah(-nya) Adi di-rata-kan kemarin.
house-D Adi PV-flat-Appl yesterday
'Adi's house was destroyed yesterday.'
- (13) Buku(-nya) dia biru, kalau buku(-nya) Desy kuning.
book-D 3s blue as.for book-D Desy yellow
'His book is blue, but Desy's book is yellow.'

The nature and extent of this variability in the possessive suffix *-nya* has not yet been studied and is not well defined. However, for my consultants, the sentences above are possible both with and without *-nya*, and I find no semantic difference resulting from pronunciation of *-nya* in these examples (see further discussion in Section 4.3.3). Anecdotally, my consultants report that some speakers frequently use the suffix *-nya* with internal possession, while other speakers rarely do (i.e., except for contexts in which it is required). It is possible that idiosyncratic usage or sociolinguistic factors (such as age, geography or influence from other languages) determine the variable expression of *-nya*. Crucially however, all consultants agreed that the grammaticality of (11-13) was not affected by (non-)pronunciation of *-nya*.

I proceed with the assumption that *-nya* is an optional morpheme that occurs in internal possession. In anticipation of the analysis below, *-nya* in possessive DPs is glossed as the functional head D.

4.1.4. *Outline of chapter*

In Section 4.2 I present possessor extraction patterns in Indonesian, focusing on the possibility of sub-extraction from subject position, object position and other positions within the clause. These patterns support an analysis in which the possessor undergoes successive-cyclic A-bar movement before landing in its surface position at the edge of the clause. I also include possessor extraction data from two related languages of Indonesia, Javanese and Madurese, which support the analysis proposed for Indonesian. In Section 4.3 I discuss the status of the suffix *-nya* in possessive DPs. I argue that *-nya* is not a 3 pronoun in possessives, but rather the pronunciation of the head D, which

is a type of agreement with a possessor. The interaction of definiteness and pronunciation of *-nya* is also discussed. Section 4.4 presents a derivational analysis for possessor extraction, including movement through the edge of DP and VoiceP, driven by the edge feature [EPP-D]. The implications of possessor extraction are discussed in Section 4.5. Section 4.6 concludes.

4.2. Patterns of possessor extraction in Indonesian

Argument extraction in Indonesian is a well-studied topic, and it is well known that subjects in Indonesian may be freely extracted, while extraction of arguments that begin lower than Voice/*v* require a null prefix instead of the active prefix *meN-* (Saddy 1991; Voskuil 1996, 2000; Cole and Hermon 2005; Aldridge 2008; Arka and Manning 2008; Cole et al. 2008b; Fortin 2009; Davies 2010; Yanti 2010; Sato 2012). The goal of the discussion in this section is to show that this generalization profitably extends to a previously overlooked set of nominals: possessors. When a possessive DP is in subject position, the possessor may be sub-extracted from its DP and appear in an A-bar position at the left edge of the clause. This movement is without consequences for verbal morphology. However, sub-extraction of a possessor originating in a lower argument requires a null prefix on the verb instead of the active prefix *meN-*. I argue that this pattern shows that possessors undergo movement in the same way that verbal arguments do, via successive-cyclic A-bar movement.³ I argue that base generation of the possessor in a high position does not account for the extraction patterns observed in Indonesian, as well as similar patterns in related languages such as Javanese and Madurese.

4.2.1. Sub-extraction of subject possessors

I begin with possessor sub-extraction from the position of grammatical subjects. The fact that possessors can escape from a subject in Indonesian has previously been noted by Musgrave (2001) and Sneddon et al. (2012). In (14), the possessive DP occurs in the canonical preverbal position of grammatical subjects. This possessive DP may undergo A-bar movement to the left edge of the clause, followed by *yang* (15). The possessor may also be extracted alone as in (16-17), in which case the possessum remains in situ, with the obligatory suffix *-nya*.

³Sub-extraction from ditransitive arguments appears to be an exception; see Section 4.2.4.

- (14) Rumah Adi di-rata-kan kemarin.
house Adi PV-flat-Appl yesterday
'Adi's house was destroyed yesterday.'
- (15) Rumah Adi yang ___ di-rata-kan kemarin.
house Adi C.Foc PV-flat-Appl yesterday
'It was Adi's house that was destroyed yesterday.'
- (16) orang yang rumah-nya ___ di-rata-kan kemarin
person C.Foc house-D PV-flat-Appl yesterday
'the person whose house was destroyed yesterday'
- (17) Siapa yang rumah-nya ___ di-rata-kan kemarin?
who C.Foc house-D PV-flat-Appl yesterday
'Who is it that their house was destroyed yesterday?'

Note that the extracted possessor can be a lexical possessor such as *orang* 'person' (16) or the wh possessor *siapa* 'who' (17). Extracted wh possessors are usually human. In other words, *siapa* 'who' may be sub-extracted, but *apa* 'what' resists sub-extraction as illustrated in (19). However, pied piping of the wh phrase makes the question acceptable (20). Lexical possessors, on the other hand, are not sensitive to a human/non-human distinction, so the head of a relative clause can be a non-human possessor, as in (21).

- (18) Atap Balai Kota di-bongkar kemarin.
roof hall city PV-tear.down yesterday
'The roof of City Hall was torn down yesterday.'
- (19) ?Apa yang atap-nya ___ di-bongkar kemarin?
what C.Foc roof-D PV-tear.down yesterday
('What is it that its roof was destroyed yesterday?')
- (20) Atap apa yang ___ di-bongkar kemarin?
roof what C.Foc PV-tear.down yesterday
'The roof of what was destroyed yesterday?'
- (21) Kami me-lihat gedung yang atap-nya ___ di-bongkar.
1s.Excl AV-see building C.Foc roof-D PV-tear.down
'We saw a building whose roof was torn down.'

As previously mentioned, possessor extraction is possible for DPs with a range of semantic relations, i.e. inalienability and affectedness are not required. Neither is possessor extraction dependent

on thematic role. (14-21) are passive clauses in which the possessor extracts from a Theme in subject position, but possessors may also extract from Agents and Experiencers in active clauses, as illustrated in examples below.

Grammatical subjects in Indonesian are marked as definite or specific, otherwise they are interpreted as generic (Sneddon et al. 2012). This rules out single question words and indefinites in subject position:⁴

- (22) Sepatu ini kena air.
shoe this get water
'These shoes got wet.'
- (23) *Sepatu kena air.
shoe get water
('A shoe/some shoes got wet.')
- (24) Rumah Adi di-rata-kan kemarin.
house Adi PV-flat-Appl yesterday
'Adi's house was destroyed yesterday.'
- (25) *Apa di-rata-kan kemarin?
what PV-flat-Appl yesterday
('What was destroyed yesterday?')

Interestingly, if the subject is a possessive DP, it is the definiteness of the possessor, rather than the definiteness of the possessum, that fulfills the subject requirement.

- (26) Sepatu Melly kena air.
shoe Melly get water
'Melly's shoes got wet.'
- (27) Sepatu orang itu kena air.
shoe Melly get water
'Melly's shoes got wet.'
- (28) *Sepatu siapa kena air?
shoe who get water
('Whose shoes got wet?')

⁴Cole et al. (2005) suggest that the inability of wh phrases to occur in subject position arises from the interaction of syntax and information structure; they also report variation in this requirement in Colloquial Jakartan Indonesian.

In contrast to the definite possessors *Melly* and *orang itu*, the generic noun *orang* ‘person’ and the wh phrase *siapa* ‘who’ cannot occur as possessors in subject position. The definiteness requirement means that a subject with a wh-possessor cannot be left in-situ; a wh possessor forces either subject extraction (29) or possessor sub-extraction (30).

- (29) Sepatu siapa yang ___ kena basah?
 shoe who C.Foc get wet
 ‘Whose shoes got wet?’
- (30) Siapa yang sepatu-nya ___ kena basah?
 who C.Foc shoe-D get wet
 ‘Who is it that (her) shoes got wet?’

4.2.2. Sub-extraction of object possessors

Next I turn to possessive DPs in object position. I show that A-bar movement of the possessor over an active verb is only possible with special morphology on the verb; this morphological requirement is the same for arguments extracted from object position. For a possessive DP that is the internal argument in a transitive clause, it may be extracted as in (32). Its possessor may also be sub-extracted (33).

- (31) Adik mem-baca buku gadis itu.
 younger.sibling AV-read book girl that
 ‘Little brother is reading that girls book.’
- (32) Buku gadis itu yang adik { baca / *mem-baca } ____.
 book girl that C.Foc younger.sibling read AV-read
 ‘It is that girls book that little brother is reading.’
- (33) Siapa yang adik { baca / *mem-baca } buku-nya ____?
 who C.Foc younger.sibling read AV-read book-D
 ‘Who is it that little brother is reading (her) book?’

In both cases of extraction, the active voice prefix *meN-* cannot occur on the verb. Instead, the verb *baca* must occur in its “bare form,” with no active voice morphology. Note that this requirement applies only to A-bar movement: in passive clauses, A-movement of the internal argument over the

verb does not require null voice morphology. Rather, the verb occurs with the passive prefix *di-* (see examples 14-17).

This morphological requirement for object extraction contrasts with subject extraction from the same clause, which is well-formed with the prefix *meN-*.⁵

- (34) Siapa yang ___ mem-baca buku gadis itu?
who C.Foc AV-read book girl that
'Who is it that is reading the girls book?'

This is consistent with the general pattern of argument extraction in Indonesian: the active prefix *meN-* is disallowed when extraction crosses the verb. This means that in situ object questions are compatible with active voice morphology on the verb, including an in situ wh possessor:

- (35) Adik mem-baca buku siapa?
younger.sibling AV-read book who
'Whose book is little brother reading?'

- (36) Desy me-lihat foto siapa?
Desy AV-see photo who
'Whose photo did Desy see?'

(35-36) are interpreted as matrix questions, but since overt movement across the verb has not applied, no special requirement is placed on verbal morphology. Note that unlike subject arguments, non-subjects are not required to be definite or specific in Indonesian. Thus wh-possessors and generic possessors are licit in object position.

4.2.3. Long distance sub-extraction

As discussed in Chapter 2, DPs may undergo long distance extraction in Indonesian. This predicts that A-bar movement will be possible for possessors as well.

However, when A-bar movement crosses more than one clause boundary, the sentence is

⁵In this chapter I do not detail the optional realization of *meN-* in clauses without A-bar movement, a topic that is discussed in Chapter 3. Verbs that are shown with the active prefix *meN-* may sometimes occur without it; for example, 34 is grammatical with the bare verb *baca*, but the availability of the prefix contrasts with cases of A-bar movement over the verb, in which *meN-* is always disallowed, as in (32-33).

sometimes judged unnatural or degraded; this is not particular to possessors, but true of argument extraction in general. Acceptability improves for many speakers if the matrix subject is a 1 or 2 pronoun. The following illustrates long distance extraction of a possessor, with a 1 and 2 pronouns as matrix subject:

- (37) Saya { kira / meng-ira } rumah Adi di-rata-kan kemarin.
 1s think AV-think house Adi PV-flat-Appl yesterday
 ‘I think Adi’s house was destroyed yesterday.’
- (38) Siapa yang kamu { kira / *meng-ira } rumah-nya ___ di-rata-kan kemarin?
 who C.Foc 2s think AV-think house-D PV-flat-Appl yesterday
 ‘Who is it that you think (his) house was destroyed yesterday?’
- (39) Kamu { kira / meng-ira } siapa yang rumah-nya ___ di-rata-kan kemarin?
 2s think AV-think who C.Foc house-D PV-flat-Appl yesterday
 ‘Who is it that you think (his) house was destroyed yesterday?’

The matrix verb in (38) must occur in its bare form *kira*, since the possessor begins as the object of the embedded clause and crosses this verb. In (39) however, the *wh* possessor moves to an intermediate position at the left edge of the embedded clause; it has not crossed the matrix verb, which may occur with the active prefix *meN-*. This pattern demonstrates that it is overt movement over the verb that requires null morphology on the active verb, and that this requirement applies iteratively in long distance movement.⁶

To summarize the discussion thus far, previous literature has already established the correlation between obligatory null voice morphology in active clauses and extraction from object position: in active transitive clauses in Indonesian, a bare verb is obligatory with DP extraction from object position, but not subject position. The data presented here show that the pattern also applies to the sub-extraction of possessors. A bare verb is obligatory when the external possessor is associated with the possessum in object position, but not when associated with a subject possessum. This supports a movement analysis for possessors that occur on the left periphery of the sentence. I assume that the same mechanism for A-bar movement of nominal arguments is also the strategy used for

⁶In order to simplify the discussion, I have not included the optional complementizer *bahwa* in (37) and the obligatory null C in (38). See Chapter 2 for discussion.

possessor movement.

4.2.4. Sub-extraction from other positions

In this section I examine possessor extraction from other argument positions in the clause, namely adjuncts and ditransitives.

Adjuncts are islands for A-bar movement in Indonesian. For instance, although PP adjuncts can occur in various positions in the clause, they cannot be clefted with the morpheme *yang*, which is a property of A-bar movement:

- (40) Aku mau rapat di sekolah-nya Dodi.
1s want meet at school-D Dodi
'I am going to a meeting at Dodi's school.'
- (41) Di sekolah-nya Dodi (*yang) aku mau rapat.
at school-D Dodi C.Foc 1s want meet
'At Dodi's school I want to meet.'

Sub-extraction from an adjunct clause is also impossible:

- (42) Ibu sedih karena kesehatan nenek mulai me-nurun.
mother sad because health grandmother begin AV-descend
'Mother was sad because Grandmother's health began to worsen.'
- (43) *Apa yang Ibu sedih karena ___ mulai me-nurun?
what C.Foc mother sad because begin AV-descend
('What was Mother was sad because began to worsen?')

These facts predict that possessor sub-extraction from an adjunct will be impossible. This is the case for both a PP adjunct (44) and an adjunct clause (45):

- (44) *Siapa yang aku mau rapat di sekolah-nya ___?
who C.Foc 1s want meet at school-D
('Who is it that I am going to a meeting at (his) school?')
- (45) *Siapa yang Ibu sedih karena kesehatan-nya ___ mulai me-nurun?
who C.Foc mother sad because health-D begin AV-descend
('Who is it that Mother is sad because (her) health began to worsen?')

Next, I turn to possessor extraction from ditransitive clauses. Indonesian allows extraction of the higher of the two ditransitive arguments, e.g. a Beneficiary, Goal or Recipient:

- (46) Kamu mem-beli-kan ibu-mu bunga.
2s AV-buy-Appl mother-2s flower
'You bought your mother flowers.'
- (47) Siapa yang kamu (*mem-)beli-kan ____ bunga?
who C.Foc 2s AV-buy-Appl flower
'For whom did you buy flowers?'

On the other hand, extraction of the lower argument (*bunga* 'flower') is not possible.⁷ Since the ability of an argument to undergo extraction is correlated with the possibility of possessor sub-extraction from the same position, this predicts that possessor will be able to extract from a ditransitive object.

Contrary to expectation however, a possessor cannot be sub-extracted from a Beneficiary, Goal or Recipient:

- (48) *orang yang kamu beli-kan istri-nya ____ bunga
person C.Foc 2s buy-Appl wife-D flower
('the person that you bought (his) wife flowers')
- (49) *Siapa yang kamu beli-kan istri-nya ____ bunga?
who C.Foc 2s buy-Appl wife-D flower
('Who is it that you bought (his) wife flowers?')

Extraction is not possible even when the verb does not bear the active voice prefix. It appears therefore that ditransitive objects are an exception to the previously noted correlation, between the extractability of a subject or monotransitive object and the extractability of its possessor. However, the ditransitive case does not challenge the previous generalization that when A-bar movement occurs over an active verb, verbal morphology must mark the movement.

It is not clear at this point why the possessor cannot extract from this position. One possible

⁷This contrasts with the data reported in Sato 2012 for formal/Standard Indonesian and Kendal Javanese. See his examples (18b, 19b) in which the Theme can be extracted over the Goal/ Recipient; this is not grammatical for the Indonesian and Javanese speakers that I consulted. The pattern is the same with give-type ditransitives (cf. Kaswanti Purwo (1995)).

explanation is that thematic specifiers in these languages may resist sub-extraction. Several properties make the applicative object position (which I assume to be the specifier of ApplP) similar to the base position of external arguments (SpecVoiceP) (cf. Pylkkanen 2002; Harley 2013; Legate 2014). First, SpecApplP and SpecVoiceP are both thematic positions, in contrast to the derived subject position, SpecIP (or SpecTP). A DP argument may move out of its thematic position: the external argument moves to SpecIP in active clauses, and the applicative object raises to SpecIP in a passive ditransitive clause (see Kaswanti Purwo 1995; Son and Cole 2008 for discussion of ditransitive constructions in Indonesian). The DP argument may also remain in situ, in its thematic position: in Object voice the external argument remains in SpecVoiceP and a Theme becomes grammatical subject (Guilfoyle et al. 1992; Cole et al. 2008a; Legate 2014). Crucially, when the external argument remains in SpecVoiceP in Object voice, it does not allow movement (sub-extraction) from within the DP: Musgrave (2001) shows that this argument cannot launch floating quantifiers in Indonesian. (See also Legate 2014 for the impossibility of floating quantifiers from an Agent in Acehnese Object voice; and Arka 2003 for similar observations in Balinese). An applicative object that remains in situ likewise disallows movement of its possessor. If this explanation is on the right track, possessor sub-extraction from ditransitives is ruled out for an independent reason, that thematic specifiers disallow sub-extraction.

4.2.5. *Interim summary*

To summarize the discussion until this point, I have shown that the properties of A-bar extraction in Indonesian also apply to possessor extraction. Null verbal morphology is required for object extraction over an active verb, and also for possessor sub-extraction from objects. No such requirement is placed on subject extraction or sub-extraction. I take this as evidence that possessors at the left periphery of the clause are not base generated in their surface positions, but undergo successive-cyclic A-bar movement, like verbal arguments.

Let us briefly consider an alternate analysis in which the possessor is base-generated in its surface position at the left periphery. This approach faces several challenges. First, it must stipulate that there is an asymmetry between subjects and non-subjects. This approach fails to explain

why a high base-generated possessor requires a null active voice prefix when the possessum is the object, but not when the possessum is the grammatical subject. By contrast, the pattern is explained under the present analysis: A-bar movement of a possessor is subject to the same constraints as A-bar movement of an argument. Given that subjects may be freely extracted in Indonesian, but objects may be extracted only with null active voice morphology, it follows that the possessors of these arguments show the same pattern. This is consistent with much previous work in Indonesian languages that observes obligatory null verbal morphology as indicative of A-bar movement of the internal object (Voskuil 2000; Cole and Hermon 2005; Aldridge 2008; Cole et al. 2008b; Fortin 2009; Davies 2010; Yanti 2010; Sato 2012; Jeoung 2017, among others). The possessor also obligatorily occurs with the morpheme *yang*, which I have argued elsewhere is required in all cases of A-bar extraction of nominals. Finally, a base-generated possessor should not be sensitive to islands, since no syntactic movement occurs. There is no obvious reason why the possessum could not occur within syntactic islands such as adjuncts, or as a ditransitive argument.

4.2.6. *Possessor extraction in related languages*

Possessor sub-extraction is also possible in two languages closely related to Indonesian, Madurese and Javanese. These languages have DP-internal structure parallel to Indonesian, and a voice system similar to that of Indonesian (for discussion of Indonesian-type languages, particularly in contrast to Philippine-type languages, see Arka 2002; Himmelmann 2002; Cole et al. 2008a; Blust 2013). In this section I show that possessors in Javanese and Madurese extract via A-bar movement. Similar to Indonesian, these languages show an asymmetry between subjects and non-subjects with regard to the verbal morphology required for extraction. Additionally, a split between extraction patterns in two registers in Madurese further supports the present movement analysis for possessors.

Like Indonesian, Javanese and Madurese are SVO languages. Nominals that undergo A-bar movement appear at the left edge of the clause:

(50) Javanese

- a. Lina senengi kue.
Lina like cake

'Lina likes cake.'

- b. kue sing Lina senengi ____
cake C.Foc Lina like
'the cake that Lina likes'

(51) Madurese (familiar)

- a. Adi ng-akan mie.
Adi AV-eat noodles
'Adi is eating noodles.'
- b. Sapah se ____ ng-akan mie?
who C.Foc AV-eat noodles
'Who is eating noodles?'

A-bar movement is the strategy used for moved-wh questions as well as relatives. The extracted nominal is the head of a relative or cleft, and the nominal must be followed by the morpheme *sing* (Javanese) or *se* (Madurese), which separates it from the rest of the clause.⁸

Subject extraction is shown below for each language, followed by possessor sub-extraction from subject position, with the possessum remaining in situ in subject position. Like Indonesian, the extracted possessor may either be a wh possessor or a lexical possessor.

(52) Javanese

- a. Buku-ne Rini di-woco adik.
book-D Rini PV-read younger.sibling
'Rini's book was read by little brother.'
- b. Buku-ne Rini sing ____ di-woco adik.
book-D Rini C.Foc PV-read younger.sibling
'It was Rini's book that was read by little brother.'
- c. Sopo sing buku-ne ____ di-woco adik?
who C.Foc book-D PV-read younger.sibling
'Who is it that (her) book was read by little brother?'

(53) Madurese (familiar)

- a. Kalambhi-nah Joko e-sasa.
clothing-D Joko PV-wash

⁸I assume that the pseudo-cleft structure that I have proposed for Indonesian also applies to Javanese and Madurese.

'Joko's clothing was washed.'

- b. Kalambhi-nah Joko se ____ e-sasa.
clothing-D Joko C.Foc PV-wash
'It was Joko's clothing that was washed.'
- c. oreng se kalambhi-nah ____ e-sasa
person C.Foc clothing-D PV-wash
'the person whose clothing was washed'

For extraction of an object from an active transitive clause, the pattern in Javanese is similar to that of Indonesian. Object extraction requires a bare verb, without the active voice prefix *N-*.⁹ Either the possessive DP or the possessor may be extracted:

(54) Javanese

- a. Aku kate m-oco buku-ne penulis iku.
1s will AV-read buku-D writer that
'I will read that writer's book.'
- b. Buku-ne penulis iku sing aku kate { woco / *m-oco } ____.
book-D writer that C.Foc 1s will read AV-read
'It is that writer's book that I will read.'
- c. penulis sing aku kate { woco / *m-oco } buku-ne ____
writer C.Foc 1s will read AV-read book-D
'the writer that I will read (her) book'

In the familiar register of Madurese however, objects cannot be extracted, regardless of verbal morphology. Both bare verbs and verbs with the active prefix *N-* disallow extraction.

(55) Madurese (familiar)

- a. Ale' m-acah buku-nah Tono.
younger.sibling AV-read book-D Tono
'Little brother read Tono's book.'
- b. *Buku-nah Tono se ale { m-acah / bacah } ____.
book-D Tono C.Foc younger.sibling AV-read read
('It was Tono's book that little brother was read.')

⁹The prefix is a homorganic nasal segment whose phonological realization is determined by the first segment of the stem that it combines with.

- c. *Apa se ale { m-acah / bacah } ___?
 what C.Foc younger.sibling AV-read read
 ('What did little brother read?')

If external possessors undergo sub-extraction via A-bar movement, this predicts that the possessor will not be able to extract from object position, and this is the case:

(56) Madurese (familiar)

- a. *Siapa se ale { m-acah / bacah } buku-nah ___?
 who C.Foc younger.sibling AV-read read book-D
 ('Who is it that little brother is reading (his) book?')
- b. *orang se ale { m-acah / bacah } buku-nah ___
 person C.Foc younger.sibling AV-read read book-D
 ('the person that little brother is reading (his) book')

An unusual fact about Madurese is that A-bar movement in the familiar register is more restricted than in the polite register (Jeoung 2017). Only subjects can undergo A-bar movement in the familiar register, but in the polite register, both subject (57b) and object (57d) extraction are possible:

(57) Madurese (polite)

- a. Buku-epon pak ustadz e-maos sareng rajih-epon.
 book-D Mr teacher PV-read by wife-D
 'Teacher's book was read by his wife.'
- b. Buku-epon pak ustadz se ___ e-maos sareng rajih-epon.
 book-D Mr teacher C.Foc PV-read by wife-D
 'It was Teacher's book that was read by his wife.'
- c. Kaulah lastareh { m-acah / *bacah } buku-epon imam ka'dissah.
 1s Perf AV-read read book-D imam that
 'I already read that imam's book.'
- d. Buku-epon imam ka'dissah se kaulah lastareh { *m-acah / bacah } ____.
 book-D imam that C.Foc 1s already AV-read read
 'It was that imam's book that I already read.'

Consistent with Indonesian and Javanese, extraction from object position requires the bare form of the verb in (57d). The availability of both subject and object extraction in the polite register also

predicts that possessors will be able to sub-extract from both positions. This prediction is also borne out:

(58) Madurese (polite)

- a. Paserah se buku-epon ___ e-maos sareng rajih-epon?
who C.Foc book-D PV-read by wife-D
'Who was it that (his) book was read by his wife?'
- b. Imam ka'dissah se kaulah lastareh { *m-acah / bacah } buku-epon ____.
imam that C.Foc 1s already AV-read read book-D
'It was that imam that I already read (his) book.'

Also predicted is the fact that the verbal morphology is required to be bare when the external possessor is associated with the object in (58b).

Given that subjects may be freely extracted in Indonesian, Javanese and polite Madurese, but objects may be extracted only with null active voice morphology, it follows that the possessors of these arguments show the same pattern. Similarly, subjects and their possessors may be extracted in familiar Madurese, but objects and their possessors may not. Like other Indonesian languages, obligatory null verbal morphology is required when the internal object undergoes A-bar movement (Saddy 1991; Voskuil 1996, 2000; Cole and Hermon 2005; Aldridge 2008; Arka and Manning 2008; Cole et al. 2008a; Sato 2008a; Davies 2010; Yanti 2010; Legate 2012, 2014; Jeoung 2017; among others.)

DP-internal morphology in Javanese and Madurese also parallels that in Indonesian. The possessum is suffixed with *-ne* in Javanese, *-Nah* in familiar Madurese and *-epon* in polite Madurese, as illustrated below. This suffix is obligatory in all possessive DPs, including internal possession (a-examples below), with non-overt possessors (b-examples) and with possessor extraction as discussed previously (c-examples).

(59) Javanese

- a. Buku-ne Rini
book-D Rini
'Rini's book'

- b. Buku-ne
book-D
'his/her book' or 'the book'
- c. Sopo sing buku-ne ___ di-woco adik?
who C.Foc book-D PV-read younger.sibling
'Who is it that (her) book was read by little brother?'

(60) Madurese (familiar)

- a. Kalambhi-nah Joko
clothing-D Joko
'Joko's clothing'
- b. Kalambhi-nah
clothing-D
'his/her clothing' or 'the book'
- c. oreng se kalambhi-nah ___ e-sasa
person C.Foc clothing-D PV-wash
'the person whose clothing was washed'

(61) Madurese (polite)

- a. Buku-epon pak ustadz
book-D Mr teacher
'teacher's book'
- b. Buku-epon
book-D
'his/her book' or 'the book'
- c. Paserah se buku-epon ___ e-maos sareng rajih-epon?
who C.Foc book-D PV-read by wife-D
'Who was it that (his) book was read by his wife?'

Also parallel to Indonesian *-nya*, Javanese *-ne* and Madurese *-Nahl-epon* also occur as the definite morpheme suffixed to nouns, as indicated by the glosses in (59b, 60b, 61b). The morphology of Indonesian possessive DPs differs only in that the (*-nya*) is optional in internal possession; otherwise, *-nya* is required. In the following section, I propose to treat the possessive suffix *-nya* on par with Javanese *-ne* and Madurese *-Nahl-epon*, which are not pronominal but rather a type of possessor agreement.

4.3. The suffix *-nya*

4.3.1. *-Nya* as possessor agreement

I now turn to the status of the Indonesian suffix *-nya*, which occurs in both possessive contexts and non-possessive contexts. It is clear that *-nya* serves several different functions in non-possessive contexts. I briefly describe some of these below, but I do not attempt to provide an analysis of all functions of *-nya*. Here I focus on the status of *-nya* only in possessive DPs. *-Nya* occurs in three types of possessive DPs :

(62) Internal possession (optional *-nya*)

Buku(-nya) Siti
book-D Siti
'Siti's book'

(63) *pro* possessor (obligatory *-nya*)

buku*(-nya) *pro*
book-D *pro*
'his/her book'

(64) Possessor extraction (obligatory *-nya*)

Siapa yang buku*(-nya) ___ di-baca oleh adik?
who C.Foc book-D PV-read by younger.sibling
'Who is it that (her) book was read by little brother?'

As previously mentioned, possessor extraction has not been studied in depth for Indonesian or closely related languages, but possessive DPs are mentioned in previous work on Indonesian. In internal possession (62), with an overt possessor that has not been extracted, previous authors have treated optional *-nya* as a linker between possessum and possessor, although no analysis has been proposed for the structural position of this linker (e.g. see Sneddon et al. 2012). When the possessor is not overt as in (63), *-nya* is most commonly treated as a 3 singular possessive pronoun; this leads to the analysis of *-nya* as a resumptive pronoun in cases of extraction such as (64) (Voskuil 2000; Musgrave 2001; Chung 2008).

I depart from the view that *-nya* is a 3 singular possessive pronoun in (63), and by extension, a resumptive pronoun in (64). One difficulty of the pronominal analysis is that it requires two distinct *-nya* morphemes, even though they occur in the same position in (62-64). *-Nya* is assumed to be a pronoun in cases without an overt possessor (63-64), but this assumption cannot be extended to cases in which an overt possessor is present (62). In the latter case, *-nya* is simply said to be a different morpheme, i.e. a possessive linker. However, identity in form and distribution suggest a uniform treatment for *-nya* in possessive DPs.

In the analysis that follows, I propose that in possessive contexts, the head D may be pronounced as *-nya*, which is realized as a suffix on the head noun (possessum). The pronunciation of D is a type of agreement with the possessor: when an overt possessor occurs (internal possession), possessor agreement is optional, as seen in (62). When the possessor is *pro*, possessor agreement is obligatory (63). When the possessor undergoes A-bar movement (64), agreement in the form of *-nya* is also required.

Several pieces of evidence support the analysis of *-nya* in possessive DPs as pronunciation of D rather than a pronoun: (i) The definite marker *-nya* and possessive *-nya* occur in the same position, on the head noun. (ii) *-Nya* can co-occur with lexical and pronominal possessors within possessive DPs. This is unexplained if *-nya* is a pronoun. (iii) *-Nya* is obligatory in possessor extraction, not only with 3 person arguments, but also with 1 or 2 person arguments. (iv) *-Nya* cannot occur as a resumptive element in general (non-possessive) argument extraction. I discuss each of these in turn.

First, the linear position of possessive *-nya* is the same as that of the definite morpheme *-nya*. (65) is a non-possessive context in which *-nya* marks only definiteness on the subject:

- (65) { Tempat-nya / Tempat itu } bagus buat foto-foto.
 place-D place that good make photo-Red
 ‘The place was good for photos.’

Definiteness may be marked on the noun with *-nya*, or with the demonstrative *itu*.¹⁰ Recall that

¹⁰*Itu* has come to function as a definite morpheme in Indonesian, and can be bleached of deictic or demonstrative meaning...

subjects in Indonesian must be definite or specific, so the subject in (65) is ill-formed without *-nya* or *itu*. Turning to possessive DPs, the possessum suffixed with *-nya* is identical to definite DPs. Compare (65) with the possessive contexts below:

- (66) Tempat(-nya) Tono bagus buat foto-foto.
 place-D Tono good make photo-Red
 ‘Tono’s place is good for photos.’
- (67) Tono pikir tempat-nya bagus buat foto-foto.
 Tono think place-D good make photo-Red
 ‘Tono thinks his place is good for photos.’

The identity of form and position between the definite suffix and the possessive suffix suggests that these are both pronunciations of the same functional element, which I take to be the head D. Furthermore, a limited set of modifiers can occur between N and *-nya*; this set remains constant whether the context is definiteness or possession. Single modifiers, including those denoting size, quality, material or origin, may occur between N and *-nya*, interpreted either as a definite or a possessive.

- (68) a. rumah besar-nya
 house big-D
 ‘his/her big house’ or ‘the big house’
- b. *rumah besar dan mewah -nya
 house big and fancy D
 (‘his/her big and fancy house’ or ‘the big and fancy house’)
- c. rumah besar dan mewah
 house big and fancy
 ‘a big and fancy house’
- d. rumah besar dan mewah itu
 house big and fancy that
 ‘the big and fancy house’
- (69) a. kain sutera-nya
 cloth silk-D
 ‘his/her silk cloth’ or ‘the silk cloth’
- b. *kain sutera panjang -nya
 cloth silk long D

(‘his/her long silk cloth’ or ‘the long silk cloth’)

- c. kain sutera panjang
cloth silk long
‘(some) long silk cloth’
- d. kain sutera panjang itu
cloth silk long that
‘the long silk cloth’

Complex modifiers such as *besar dan mewah* ‘big and fancy’ and multiple adjectives such as *sutera panjang* ‘long silk’ cannot occur inside *-nya*. In definite contexts, these can occur with the demonstrative *itu*. This pattern is striking because the distribution of definite *-nya* differs from that of demonstrative *itu* (which always occurs in final linear position), yet definite *-nya* and possessive *-nya* share the same distribution.

A second piece of evidence that argues against *-nya* as a pronoun is that it can co-occur with pronominal possessors. Consider the range of possessive DPs shown with various possessors in Table 4.^{11,12}

| | free possessor | clitic possessor | <i>-nya</i> + possessor (optional <i>-nya</i>) | possessor extraction (obligatory <i>-nya</i>) |
|------------|----------------|------------------|---|--|
| 1s | rumah aku | rumah-ku | rumah-nya aku | aku yang ... rumah-nya |
| 2s | rumah kamu | rumah-mu | rumah-nya kamu | kamu ... rumah-nya |
| 3s | rumah dia | - | rumah-nya dia | dia ... rumah-nya |
| 1p Incl | rumah kita | - | rumah-nya kita | kita ... rumah-nya |
| 1p Excl | rumah kami | - | rumah-nya kami | kami ... rumah-nya |
| 2p | rumah kalian | - | rumah-nya kalian | kalian ... rumah-nya |
| 3p | rumah mereka | - | rumah-nya mereka | mereka ... rumah-nya |
| ‘whose’ | rumah siapa | - | rumah-nya siapa | siapa ... rumah-nya |
| ‘Rika’s’ | rumah Rika | - | rumah-nya Rika | Rika ... rumah-nya |
| <i>pro</i> | - | - | rumah-nya <i>pro</i> | - |

Table 4: Possessive DPs with possessum *rumah* ‘house’

The fourth column illustrates the optional use of *-nya* between possessum and possessor in colloquial Indonesian. In standard (prescriptive) Indonesian, possessive DPs with overt possessors do not occur with *-nya* (cf. Sneddon et al. 2012). However, for my educated Indonesian consultants from

¹¹The 1 singular clitic *ku* can occur as enclitic or proclitic.

¹²There is variation in whether *rumah-nya* can be used for 3 plural

the island of Java, *-nya* may optionally occur with an overt internal possessor, and this is reported to be common in speech. The co-occurrence of *-nya* with an overt possessor is reported to be more frequent in emphatic contexts or affective discourse styles, for example, when the speaker is angry. However, this is not a strict requirement, as illustrated in the examples below, where *-nya* co-occurs with an overt possessor in a non-emphatic context:

- (70) Buku(-nya) dia biru, kalau buku(-nya) Desy kuning.
 book-D 3s blue as.for book-D Desy yellow
 ‘His book is blue, but Desy’s book is yellow.’
- (71) Uang(-nya) orang kaya cepat di-keluar-kan.
 money-D person rich quick PV-exit-Appl
 ‘Rich people’s money is quickly spent.’

I conclude then, that *-nya* may optionally occur with an overt internal possessor, and that this variability depends on stylistic or idiosyncratic usage. (I return shortly to the question of whether *-nya* always encodes definiteness.)

Table 4 departs from previous descriptions of *-nya* in another significant way. Previous authors consider *rumah-nya* to be the cliticized variant of *rumah dia* ‘his/her house’ or *rumah mereka* ‘their house.’ Under the present analysis however, *rumah-nya* contains no overt possessor, but rather a *pro* possessor (see the last row of Table 4). Since Indonesian is a *pro*-drop language that allows null arguments when they are understood within discourse context, it is reasonable to assume a *pro* possessor, which occurs when the possessor is already understood or has been previously mentioned. *Pro* possessors require agreement on the possessum in the form of *-nya*. Accordingly, 3 person pronouns *dia* and *mereka* do not have a corresponding possessive clitic form at all. Rather, the *pro* possessor is interpreted as 3 person by default, rendering ‘his/her/their house’ as possible readings for *rumah-nya*.

In short, *-nya* may occur with all possessors across the forms in Table 4, which is explained if there is a single *-nya* morpheme that is optional in possessive contexts. This analysis accounts for the facts in a straightforward way, rather than positing two *-nya* morphemes that occur in the same position in possessive DPs.

A third piece of evidence that possessive *-nya* is not pronominal comes from possessor extraction, which results in obligatory *-nya* (see last column of Table 4). Any possessor that is extracted requires *-nya* on the possessum, regardless of the type of DP (lexical possessor, wh phrase) and person/number features:

- (72) { Aku/ Kamu/ Dia/ Kita } yang rumah-nya ___ di-rata-kan kemarin.
 1s 2s 3s 1p.Incl C.Foc house-D PV-flat-Appl yesterday
 ‘It is I/you/he/we whose house was destroyed yesterday.’
- (73) { Rumah aku/ Rumah kamu/ Rumah dia/ Rumah kita } yang ___ di-rata-kan
 house 1s house 2s house 3s house 1p.Incl C.Foc PV-flat-Appl
 kemarin.
 yesterday
 ‘It is my house/your house/his house/our house that was destroyed yesterday.’

I should note that extraction of a pronominal possessor (as in 72) is judged to be awkward by some consultants, although no more so for a 1 or 2 possessor than for a 3 possessor. All consultants preferred to extract the possessive DP as in (73), with intonational stress on the pronoun; this is reported to be more common than pronoun extraction (72). However, when a context is presented that forces a pronominal possessor to be focused and extracted as in (72), then the only acceptable suffix on the possessum is *-nya*; this judgment was shared by all consultants. The traditional view, which treats *-nya* as a 3 person resumptive pronoun, does not explain the fact that *-nya* is compatible with the extraction of 1 and 2 person arguments in (72). Furthermore, if *-nya* is a resumptive 3 pronoun, then this predicts that the extraction of 1 or 2 person arguments will occur with a resumptive pronoun as well. However, 1 or 2 resumptive pronouns are ungrammatical with possessor extraction:

- (74) *Aku yang rumah-ku di-rata-kan kemarin.
 1s C.Foc house-1s PV-flat-Appl yesterday
 (‘It is I whose house was destroyed yesterday.’)
- (75) *Kamu yang rumah-mu di-rata-kan kemarin.
 2s C.Foc house-2s PV-flat-Appl yesterday
 (‘It is you whose house was destroyed yesterday.’)

This demonstrates that in possessor extraction, the obligatory suffix *-nya* does not carry phi-features,

but rather marks the extraction of any possessor out of the DP. This pattern also supports an analysis of *-nya* as an instance of wh-agreement. One of the properties of wh-agreement is that it does not register DP-internal features of the moved argument, i.e. phi-features (Reintges et al. 2006).

Finally, if the suffix *-nya* in possessor extraction is a 3 person pronoun that occurs resumptively, this predicts that *-nya* will occur resumptively in general (non-possessive) argument extraction. However, this is impossible:

- (76) Apa yang adik baca (*nya)?
 what C.Foc younger.sibling read NYA
 ('What is little brother reading?')
- (77) Siapa yang Desy lihat (*nya)?
 who C.Foc Desy see NYA
 ('Who did Desy see?')

(76-77) show that *-nya* only occurs with possessor extraction, not general argument extraction. It is worth noting here that this fact does not fall out from a morphological requirement that *-nya* attach to nominal arguments only. *-Nya* can cliticize to other categories in Indonesian, such as transitive verbs (Musgrave 2001; Sneddon et al. 2012; also see Kroeger 2014 on Malay):

- (78) Ibu mem-per-indah baju-ku dengan men-jahit (-nya).
 Mother AV-Caus-beautiful shirt-1s with AV-sew NYA
 'Mother adorned my shirt by sewing (it).'
- (79) Saya sudah mem-baca-nya.
 1s already AV-read-NYA
 'I've read it/them.' (modified from Musgrave 2001:91)

Since *-nya* can cliticize to verbs, that is not the source of the ungrammaticality in (76-77). The fact that *-nya* does not occur in the place of the moved argument in (76-77) is unexplained if it is a resumptive pronoun, but it is expected under the current analysis, because *-nya* is required only when an argument is sub-extracted from DP. Since there is no sub-extraction from DP in (76-77), wh-agreement in the form of *-nya* is not required; instead, extraction of a DP from object position as in (76-77) instead requires wh-agreement on the verb (i.e. a bare verb form).

4.3.2. *-Nya* in non-possessive contexts

Having argued that *-nya* is not a pronoun in possessive DPs, let me briefly note some other functions of the form *-nya* in non-possessive contexts. In the examples below, I gloss *-nya* simply as NYA. We have already seen in (78-79) that *-nya* appears to occur in place of an object argument, and is optional embedded in the adjunct PP in (78). It is also possible for *-nya* to occur as a prepositional object as in (80):

- (80) Saya di-jemput (oleh dia / oleh-nya).
1s PV-meet by 3s by-NYA
'I was met by him.' (modified from Musgrave 2001:140)
- (81) Saya di-jemput (dia).
1s PV-meet 3s
'I was met by him.'
- (82) Saya di-jemput (-nya).
1s PV-meet NYA
'I was met by him.'

The preposition *oleh* may be null in passive clauses, as shown in (81-82) (cf. Kroeger 2014; Jeoung and Biggs 2017). With a null P, *-nya* can cliticize directly to the verb in (82). It is impossible, however, for *-nya* to occur as a subject, even in embedded clauses introduced by an overt complementizer that might host a clitic:

- (83) *Tono bilang bahwa -nya suka kue coklat.
Tono say C -NYA like cake chocolate
(‘Tono said that he likes chocolate cake.’)

-Nya can also occur in object position with a left-dislocated topic:

- (84) Surat ini, saya yang me-nulis { -nya / itu }.
letter this 1s C.Foc AV-write NYA that
'As for this letter, it is I who wrote it.' (modified from Voskuil 2000:207)

It is clear that this is not a case of A-bar extraction: the topic *surat ini* is not followed by the

morpheme *yang*, and verb is not required to occur in bare form, both of which are required with A-bar extraction. I have also shown that *-nya* cannot occur in the position of the gap or trace when the object is extracted (see 76-77). In (84) then, I assume the topic is base-generated in its surface position; here *-nya* appears to be the associate of the topic.

Besides cases in which *-nya* appears to function as a nominal argument, *-nya* is also used in the formation of some adverbial expressions:

- (85) Ke-lihat-an-nya Djoko sakit demam.
 Nmlz-see-Nmlz-NYA sick fever
 ‘Apparently, Djoko was sick with fever.’
- (86) Pokok-nya kita saling me-maham-i.
 primary-NYA 1p.Incl each.other AV-understand-Apppl
 ‘Most importantly, we understand each other.’

Quantified expressions with *ke-* + numeral also occur with *-nya*:

- (87) ke-dua orang itu
 KE-two person that
 ‘both of those people’
- (88) ke-dua-nya
 KE-two-NYA
 ‘both of them’

I have proposed that in possessive DPs, *-nya* is a type of agreement with a possessor DP, without assuming that this analysis applies to other uses of *-nya*. However, it seems possible that where *-nya* appears to replace a nominal argument, the present analysis might be extended to these cases if *-nya* agrees with a *pro* argument, just as with a *pro* possessor in (63). (80 and 82) are repeated below with *pro* arguments:

- (89) Saya di-jemput (oleh-nya *pro*).
 1s PV-meet by-NYA PRO
 ‘I was met by him.’
- (90) Saya di-jemput oleh (-nya) *pro*.
 1s PV-meet by NYA PRO

'I was met by him.'

Similarly, a silent *pro* argument in (84) and (88) could be compatible with *-nya* as well. However, I leave the status of *-nya* in non-possessive contexts for further research.

4.3.3. *Definiteness effects and -nya*

I have proposed that *-nya* is a type of agreement in possessive DPs, and that pronunciation of the head D in this particular construction does not have definite semantics. This predicts that possessive *-nya* will not affect the definiteness of possessive DPs in constructions that require either definiteness or indefiniteness. To test this, I examine the behavior of possessive DPs with respect to two definiteness effects previously observed in Indonesian.

The first definiteness effect is that grammatical subjects in Indonesian must be definite or specific, or else interpreted as generic (Sneddon et al. 2012). As previously discussed in Section 4.2.1, the definiteness of the possessor determines whether the DP is definite; in subject position, possessive DPs must have definite possessors. This rules out (92) and (93):

- (91) Sepatu Melly kena basah.
shoe Melly get wet
'Melly's shoes got wet.'
- (92) *Sepatu orang kena basah.
shoe person get wet
(‘The shoes of someone got wet.’)
- (93) *Sepatu siapa kena basah?
shoe who get wet
(‘The shoes of who got wet?’)

If *-nya* adds definite semantics, it might be expected to ameliorate the ungrammaticality of these indefinite subjects. However, (94) and (95) show that adding *-nya* does not count for subject definiteness:

- (94) *Sepatu-nya orang kena basah.
shoe-D person get wet

(‘The shoes of someone got wet.’)

- (95) *Sepatu-nya siapa kena basah?
shoe-D who get wet
(‘The shoes of who got wet?’)
- (96) Dia ambil sepatu-(nya) orang!
3s take shoe-D person
‘He’s taking someone’s shoes!’
- (97) Sepatu-(nya) siapa yang kena basah?
shoe-D who C.Foc get wet
‘Whose shoes was it that got wet?’

The possessive DPs are licit in object position and as head of a cleft (96-97), showing that it is only the definiteness requirement for subjects that rules out (94-95).

Existential clauses formed with *ada* provide a second diagnostic for definiteness. An argument that follows *ada* must be indefinite:¹³

- (98) Ada tiga orang dengan nama Wahyu di keluarga kami.
Exist three person with name Wahyu at family 1p.Excl
‘There are three people with the name Wahyu in our family.’
- (99) *Ada orang itu dengan nama Wahyu.
Exist person that with name Wahyu
(‘There is the person with the name Wahyu.’)
- (100) *Ada bola itu di atas atap.
Exist ball that at top roof
(‘There is the ball on the roof.’)

As expected, a possessive DP with a definite possessor cannot follow *ada*, as in (101). However, some possessive DPs with a definite possessor can have an indefinite interpretation if they are not unique, as in (102-103).

- (101) *Ada walikota Surabaya di stadion kemarin.
Exist mayor Surabaya at stadium yesterday
(‘There was the mayor of Surabaya at the stadium yesterday.’)

¹³Nominal arguments that precede existential *ada* may be definite.

- (102) Ada presiden Indonesia yang perempuan.
Exist president Indonesian C.Foc female
'There was a president of Indonesia that was female.'
- (103) Ada bola Tono di atas atap.
Exist ball Tono at top roof
'There is a ball of Tono's on the roof.'

The question is whether *-nya* adds definiteness to these possessive DPs. If the possessive suffix *-nya* encodes definiteness in possessive contexts, then we expect it will be incompatible with the existential *ada* construction. This is shown in (104):

- (104) Ada bola-nya Tono di atas atap.
Exist ball-D Tono at top roof
'There is a ball of Tono's on the roof.'

This shows that the indefinite interpretation is still possible with the possessive suffix *-nya*, which does not cause the DP to pattern as a definite.

To summarize this section, I have laid out evidence to support two complementary claims, that *-nya* is not a resumptive pronoun in possessor extraction; and that *-nya* is a type of agreement. I have proposed that *-nya* in possessive DPs as agreement with a possessor, whether the possessor is in situ, *pro* or A-bar moved. The distribution and form of this morpheme suggests that it is the pronunciation of the functional head D, which encodes definiteness elsewhere in the language.

Extending this analysis to related languages, we have previously seen that in Javanese and Madurese, the suffix on the possessum is always obligatory, even with internal possessors (refer to 59a-60c). Javanese and Madurese, then, have obligatory possessor agreement on the noun (possessum), whether the possessor is overt, *pro* or extracted from the DP. Indonesian differs from this pattern only in that possessor agreement is optional with an internal possessor.

4.4. A derivational analysis for possessor extraction

Having argued for a movement analysis for clefted possessors at the left periphery, in this section I propose that a possessor undergoes A-bar movement through the edge of DP, then the edge of VoiceP, before landing in its surface position in the CP domain. This movement is driven by edge features on D and Voice, which are phase heads.

Other accounts of possessor extraction also propose that the DP specifier acts as an “escape hatch” for further movement. Szabolcsi (1992) observes that in Hungarian possessor extraction, possessors cannot move directly from their base-generated positions, but must pass through SpecDP. Szabolcsi notes that the specifier of DP is an operator position, and takes this to support the claim that DP and CP have parallel structure and operations, since possessor movement to SpecDP parallels the movement of Hungarian subjects to clause-initial position. This approach is also taken up in Gavruseva’s (2000) analysis of possessor extraction in Hungarian, Tzotzil and Chamorro. Gavruseva proposes that possessors extract via the specifier of DP, driven by a strong uninterpretable Q feature on matrix D which attracts the possessor. I follow both Gavruseva and Szabolcsi in assuming that the possessor extracts through the specifier of DP; see also discussion in Boeckx 2003. Cole and Hermon (1998) reach a similar conclusion for the movement of a null operator in complementizer-gap relativization in Malay.

This approach has two implications. For Indonesian, I show that possessor extraction data support movement through the edge of DP, indicating that D is a phase head that drives successive-cyclic movement of nominals. Second, I draw a parallel between the obligatory morphology on the possessum (*-nya*) and the morphological effects of nominal movement through VoiceP and CP in this language, which I have argued are a morphological wh-agreement. In Indonesian, I attempt to unify the morphological effects of A-bar movement through the edge of DP, VoiceP and CP as wh-agreement.

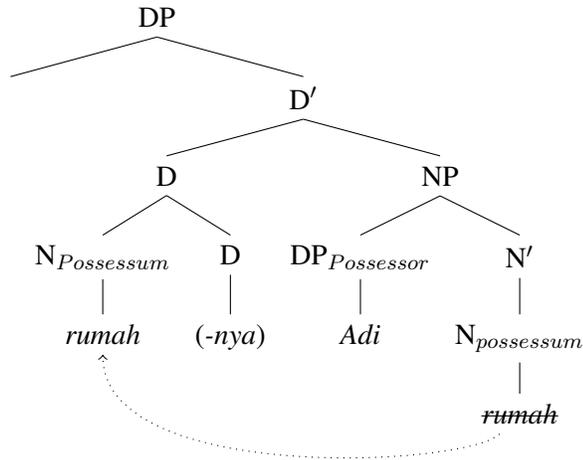
4.4.1. Movement through SpecDP

I propose a derivation for possessor extraction in two steps: A-bar movement through the edge of DP, followed by A-bar movement through the edge of VoiceP. I assume that all cases of A-bar extraction result in a pseudo-cleft, and that it is a null Operator that moves phase-cyclically through the clause. Arguments for null Operator movement are given in Chapter 2, and apply to possessor movement as well. In the discussion below, I refer generally to “possessor movement” or “possessor extraction” for ease of exposition, and represent movement with a possessor, as the distinction does not significantly affect the discussion. For the pseudo-cleft structure with null Operator movement, see the tree structure in (123).

I assume that the possessor is generated in the specifier of NP, as an external argument of N; recall that the possessor can have a number of different semantic relations with the possessum in these languages, and is not limited to relations that are inalienable, part-whole etc. The possessum is the head N, while the suffix *-nya* is the head D. The word order of possessive DPs is derived by head movement of N to D, resulting in *-nya* optionally suffixed to the possessum. This structure is shown in (106):

(105) rumah(-nya) Adi
house-D Adi
 ‘Adi’s house’

(106) Possessive DP structure



The following examples provide evidence for the structure in (106). First consider possessive DPs with postnominal complements:

- (107) buku [cerita anak]
 book story child
 'book of children's stories'
- (108) *buku [cerita anak] Siti
 book story child Siti
- (109) *buku [cerita anak] -nya Siti
 book story child -D Siti
- (110) *buku-nya [cerita anak] Siti
 book-D story child Siti
- (111) buku Siti [tentang cerita anak]
 book Siti about story child
 'Siti's book about children's stories'

When a possessor is present, the nominal complements cannot occur between the possessum and possessor (108-110). Instead, a periphrastic construction must be used, with the possessor embedded in a PP (111). A similar pattern holds for PP adjuncts that modify the head noun. The PP is usually postnominal, but cannot occur between possessum and possessor:

- (112) murid [dari Jakarta]
student from Jakarta
'student from Jakarta'
- (113) murid-(nya) Siti [dari Jakarta]
student-D Siti from Jakarta
'Siti's student from Jakarta'
- (114) *murid-(nya) [dari Jakarta] Siti
student-D from Jakarta Siti
- (115) *murid [dari Jakarta] -nya Siti
student from Jakarta D Siti

An adjective modifying the possessum also cannot occur in its usual postnominal position. Instead a relative strategy is used (117):

- (116) murid pintar
student smart
'smart student'
- (117) murid-(nya) Siti yang pintar
student-D Siti C.Foc smart
'Siti's student that is smart'
- (118) *murid pintar-nya Siti
student smart-D Siti
- (119) *murid-nya pintar Siti
student-D smart Siti

These data support a head movement analysis for a simple possessum (N), which is always raised to D in possessive contexts. When the possessum is complex, nominal complements and modifiers cannot occur between the possessum and possessor, but rather require periphrastic constructions.¹⁴

¹⁴A few adjectives can occur between the possessum and possessor, and may occur inside the suffix *-nya*. For example, adjectives denoting size, material or origin:

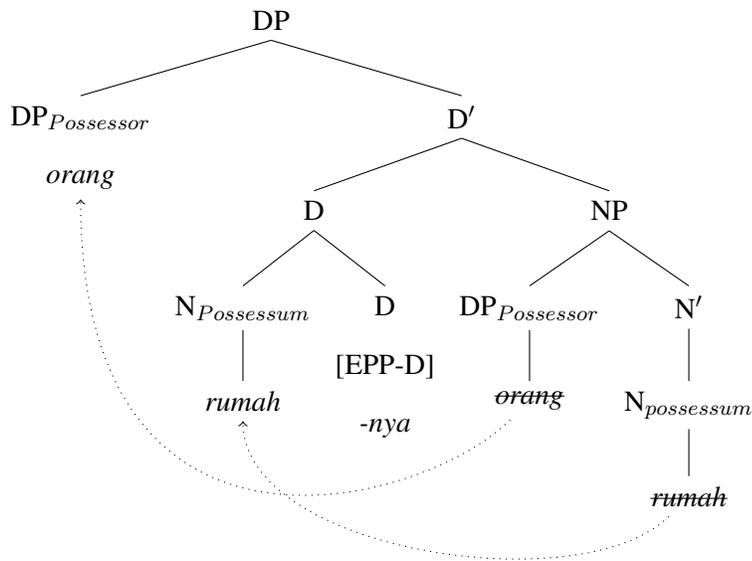
- (1) rumah besar-(nya) Siti
house large-D Siti
'Siti's big house'
- (2) kain sutera-(nya) bu Henny
cloth silk-D Mrs Henny
'Mrs Henny's silk cloth'

Given the limited number of these adjectives, and the high frequency with which they are used with the head N, I treat these as compounds that may also undergo head-raising to D.

In the first step of possessor extraction, the possessor undergoes movement driven by an edge feature on the functional head D. I call this edge feature [EPP-D], because it does not check an argument in situ, but must be satisfied by attracting a DP to its specifier; and the feature targets DPs only. A DP constituent may be clefted or relativized, whereas PP, AP and other adjuncts cannot.

(120) orang yang rumah-nya ___ di-rata-kan kemarin
 person C.Foc house-D PV-flat-Appl yesterday
 ‘the person whose house was destroyed yesterday’

(121) Possessor sub-extraction



[EPP-D] on the head D probes downward for the closest DP in its c-command domain and finds the possessor *orang*. This DP is attracted to the specifier of DP. Szabolcsi (1992), writing about possessor extraction in Hungarian, noted the similarity between C and D in terms of movement possibilities: “both are functional categories whose SPEC is a designated landing site for operators and serves as an escape hatch for movement” (1992:43).

A possessor that has first moved to the specifier of a possessive DP is available for further movement. As previously discussed, either the maximal DP projection (the possessive DP) or its specifier (the possessor) can be extracted. The difference between these has to do with focus: all

extracted nominals are focused since they result in a cleft construction. If the possessor DP is focused, it can be extracted; otherwise, the matrix DP is focused and the entire possessive DP is extracted. Therefore I assume that D bears both [EPP-D] and [Focus] features, and that these features probe together. [EPP-D] targets only DPs, and [Focus] targets the DP that is marked for pragmatic prominence. In the derivation discussed here, I assume that the possessor DP is marked for focus, but the matrix DP is not.

From SpecDP, the possessor next undergoes successive-cyclic movement through the edge of VoiceP, before landing in its surface position in the CP domain. I assume that the extended verbal structure of the clause includes both vP and VoiceP, which I take to be the highest verbal projection and the domain relevant for successive-cyclic movement, or a phase.¹⁵ The voice prefixes on the verb are the spellout of the functional head Voice. In basic active clauses, this head is realized as *meN-*. It is this head that is also phonologically null when there is extraction from object position. External arguments are generated in the specifier of VoiceP (Pylkkanen 2002; Harley 2013; Legate 2014) and raised to the position of grammatical subjects, SpecIP, to satisfy [EPP] on I.^{16,17}

In an active clause, the derivation of possessor extraction from object position begins with an active Voice head that bears the feature [EPP-D]. The [EPP-D] feature was previously discussed for DP-internal movement: it attracts the closest DP to its specifier, but does not target other categories such as PP or AP. [EPP-D] on Voice triggers movement through the edge of the verbal domain. This type of movement is proposed in Chomsky 1986, 2000 (through the edge of vP) and also proposed for a number of analyses in Indonesian and Austronesian languages (Rackowski and Richards 2005; Aldridge 2008, 2017; Cole et al. 2008b; Sato 2008b, 2012; Legate 2014; van Urk and Richards 2015). These authors agree that in these languages, one of the functions of voice morphology is to mark nominal movement through the edge of the phase.

For nominals in object position, the feature [EPP-D] on Voice attracts the closest DP that

¹⁵Causative prefixes in these languages are hosted in vP (cf. Legate 2014). These prefixes do not interact with possessor extraction.

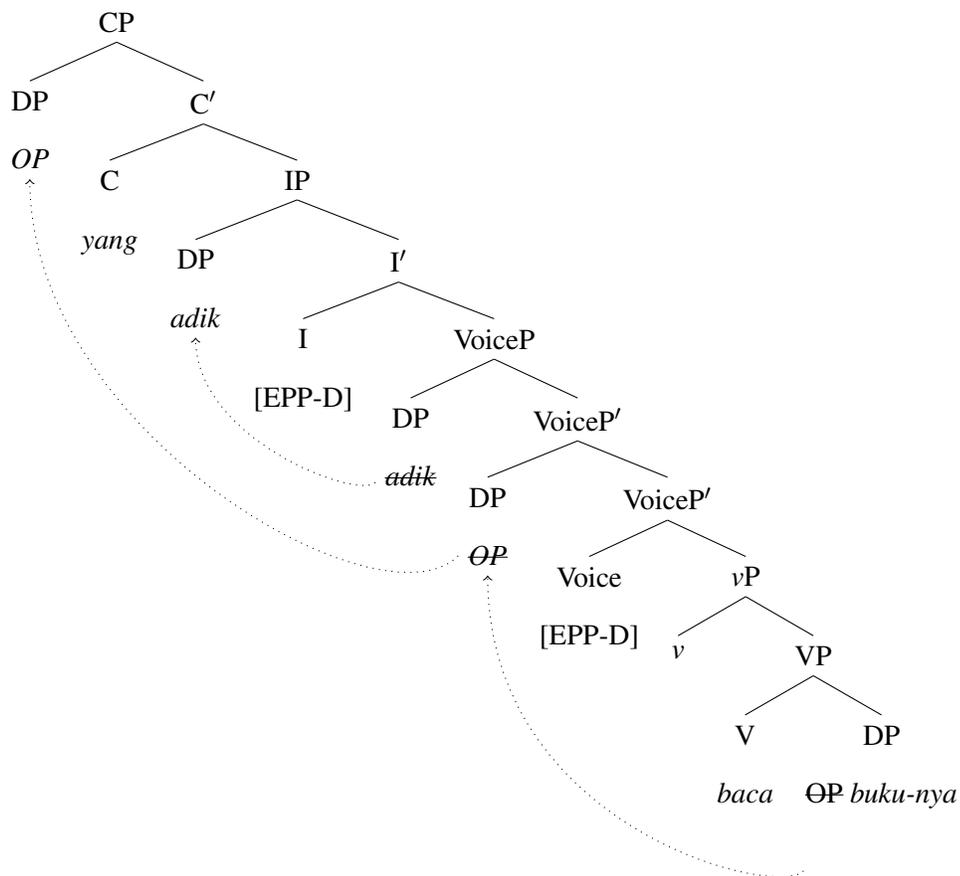
¹⁶In active voice, the derived subject position generally must be filled in these languages. VP fronting is also possible in active voice, resulting in variations in word order, but does not occur with the clefts and relatives that are discussed here.

¹⁷I use IP instead of TP because tense is not overtly marked in Indonesian clauses.

it c-commands, and raises it to SpecVoiceP where [EPP-D] is checked. This is illustrated below. Note that the tree in (123) shows the possessor as a null Operator, which undergoes phase-cyclic movement within the clause. This CP is embedded in a copular sentence in (124), with the lexical possessor *orang* generated as predicate (see Chapter 2 for discussion of the cleft structure).

(122) orang yang adik baca buku-nya —
 person C.Foc younger.sibling read book-D
 'the person that little brother read (his) book'

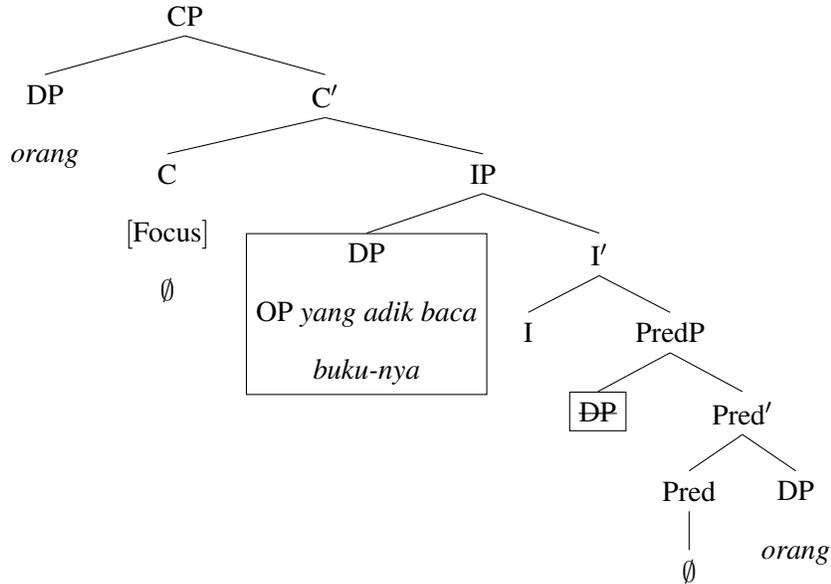
(123) Possessor sub-extraction from object position^{18,19}



(124) Matrix copular clause of pseudo-cleft

¹⁸In (123) the possessor is tucked in below the specifier hosting the external argument (cf. Richards 1999). Assuming [EPP-D] targets the closest DP eligible for movement, the external argument rather than the possessor is raised to SpecIP.

¹⁹(123) shows the root *baca* read in the node V. The formation of the phonological word that includes the heads Voice-v-V is not shown here.



The moved nominal may be either the object, or the possessor of the object, depending on which DP is focused. (DP-internal movement is not shown in 123.) If the possessor is focused, it moves to the edge of the DP; from SpecVoiceP, the possessor is available for further movement. If the possessive DP is focused, no DP-internal movement is required. A [Focus] feature on C probes for a DP with [Focus] and finds the possessor in SpecVoiceP. Although Voice in active clauses is usually pronounced as the prefix *meN-*, when Voice bears the feature [EPP-D], triggering obligatory movement, it is realized as a null prefix, \emptyset - (for related discussion of the null prefix in cases of extraction, see Cole and Hermon 2005; Cole et al. 2008b; Sato 2008b, 2012).

This analysis holds for Javanese and polite Madurese as well, as these languages allow object extraction and object possessor extraction. In familiar Madurese however, since objects and their possessors cannot extract, the active Voice head cannot bear the [EPP-D] feature. Nominals cannot raise to the edge of VoiceP from object position; consequently, [Focus] on C does not find an object bearing a [Focus] feature because it cannot probe into a lower phase. As a result, objects and their possessors cannot be extracted in familiar Madurese; and since Voice never bears the [EPP-D] feature, active verbs always bear a voice prefix. For the external argument (and its possessor) in an active clause, movement to a higher position does not require [EPP-D] on the Voice head, because the external argument is generated in the specifier of Voice and therefore is already on the edge

of the phase; it is visible for further movement. The [EPP-D] feature on I first raises the external argument to the grammatical subject position in SpecIP, where it may then be found by a probe on C. Since the external argument in an active clause does not interact with [EPP-D] on the Voice head, extraction of this nominal never has consequences for voice morphology; a bare verb is not required for subject extraction. Similarly, the internal argument (and its possessor) in a passive clause are first raised to grammatical subject, and can be extracted from SpecIP. The passive voice morphology is required when the internal argument occurs as subject; however, further extraction from subject position is not reflected by a change of verbal morphology.

4.5. Implications of possessor extraction in Indonesian

4.5.1. Phase head D and wh-agreement

Successive-cyclic movement through vP and CP was proposed in Chomsky 2000 and Chomsky 2001, and implemented in analyses of many languages. For discussion in Indonesian-type languages, see Aldridge 2008; Cole et al. 2008; Sato 2008, 2010; Legate 2014. In this thesis, I have agreed with previous work in Indonesian that supports successive-cyclic movement through the edge of the verbal domain, and through the edge of CP. This chapter argues, based on possessor extraction, that phase-cyclic movement can proceed through the edge of DP as well. I have argued that in Indonesian, possessors are DP-internal arguments that undergo A-bar movement through the edge of DP before continuing on to SpecVoiceP and SpecCP.

The main evidence for successive-cyclic movement in Indonesian is the required morphology on the phase head, D. The obligatory suffix *-nya* is a type of wh-agreement: it marks DP-internal A-bar movement to a specifier position. I have proposed that D can bear an [EPP-D] feature, which targets DPs only, and can only be checked by overt movement. The other phase heads, Voice and C, can also bear [EPP-D], which drives A-bar movement in local steps through each phase. The presence of [EPP-D] requires wh-agreement on each phase head.

The phasehood of DP in Indonesian provides a new pattern of wh-agreement. Recall that

Zaenen's Generalization did not include the nominal domain:

(125) Zaenen's generalization:

Only complementizers and verbal morphology are affected by wh-movement.

(Zaenen 1983, formulated in Watanabe 1996:177)

Possessor extraction extends this generalization beyond complementizers and verbs, since wh-agreement is also marked within the DP. The generalization is at least incomplete, and should be revised to include wh-agreement observed in the nominal domain. But I also suggest that Zaenen's generalization is better reframed with reference to syntactic phases instead of syntactic binding domains: Indonesian shows wh-agreement in the nominal, verbal and complementizer domains because DP, VoiceP and CP are syntactic phases in this language. This expands the range of possible wh-agreement patterns to any syntactic phase that allows A-bar movement through its edge. As noted by Zaenen (1983) and Reintges et al. (2006), not all languages have morphological wh-agreement when A-bar (wh-) movement obtains. Therefore wh-agreement is expected to be possible, but not required, whenever phase-based movement applies through a phase edge.

4.5.2. *EPP and subject extraction*

Possessor extraction also shows that in Indonesian, the position of grammatical subject generally must be filled, and that this requirement is satisfied before A-bar extraction from the clause.

As previously discussed, subject extraction is possible in Indonesian (as well as Javanese and Madurese), and does not impose special requirements on verbal morphology like object extraction; see examples throughout this chapter. I have assumed that grammatical subjects occupy a structural position, the specifier of IP, and that all subjects are derived by A-movement to this position. Since Indonesian generally requires a grammatical subject (whether overt or *pro*), I take this to mean that an [EPP] feature on I is filled by A movement.

In the context of areal languages, Indonesian-type languages and Philippine-type languages are historically related. The syntax of Indonesian-type languages deviates from Philippine-type

syntax to varying degrees, especially with regard to the voice system and argument extraction. (For discussion of Indonesian-type syntax and Philippine-type syntax, see Wouk and Ross 2002; Himmelmann 2005; Cole et al. 2008b). In Philippine-type languages such as Tagalog, it is possible to extract various thematic arguments from within the clause. This extraction does not require overt movement to a “subject” position; rather, verbal morphology marks which argument has been extracted. (For various analyses of Tagalog syntax, see Kroeger 1993; Aldridge 2004; Rackowski and Richards 2005; Aldridge 2017, among others.)

Indonesian deviates from Philippine languages in filling the structural subject position, independent of extraction. However, this is usually not directly observed. Consider the following sentences in active, object and passive voices. In the (a)-examples, the grammatical subject is shown in its surface position, and the position from which it has undergone movement is shown as a gap. The external argument in (126) is generated in SpecVoiceP; the internal arguments in (127) and (128) are generated in post-verbal position as the complement of V. In the (b)-examples, the clefted (extracted) argument has the same thematic role as the grammatical subject in (a)-examples. Note however, that it is not obvious if the argument has been extracted from derived subject position, or from its base position.

(126) Active voice

- a. **Teman-nya adik** akan ___ mem-baca buku itu.
friend-D younger.sib will AV-read book that
'Little brother's friend will read the book.'
- b. **Siapa** yang ___ akan ___ mem-baca buku itu?
who C.Foc will AV-read book that
'Who will read the book?'

(127) Object voice

- a. **Buku Siti** akan adik baca ____.
book Siti will younger.sib read
'Little brother will read Siti's book.'
- b. **Apa** yang ___ akan adik baca ____?
what C.Foc will younger.sib read
'What will little brother read?'

(128) Passive voice

a. **Buku Siti** di-baca ____.
book Siti PV-read
'Siti's book was read.'

b. **Apa** yang ____ di-baca ____?
what C.Foc PV-read
'What was read?'

The question is whether the clefted DP in the b-examples can be extracted directly from its base (thematic) position, without movement to SpecIP. In other words, it is not clear whether the EPP feature on I must first be satisfied in cases of A-bar extraction. If EPP does not need to be satisfied when extraction occurs, this would mean that extraction in Indonesian preserves a Philippine-type flavor.

Possessor extraction provides a way to test this type of movement. I have previously noted that DP extraction from a particular position is correlated with possessor sub-extraction from the same argument. If a possessor is able to extract directly from thematic position, then the possessum should be able to occur in situ, with no grammatical subject. The following examples test this possibility:

(129) Active voice

***Siapa** yang akan teman-nya ____ mem-baca buku itu?
who C.Foc will friend-D AV-read book that
'Who is it that (his) friend will read the book?'

(130) Object voice

***Siapa** yang akan adik ____ baca buku-nya ____?
who C.Foc will younger.sib read book-D
'Whos it that little brother will read (his) book?'

(131) Passive voice

***Siapa** yang di-baca buku-nya ____?
who C.Foc PV-read book-D
'Who was it that (his) book was read?'

This data shows that when the grammatical subject position has not been filled, DP extraction directly from thematic base position is not possible. The EPP feature on I, then, must be satisfied in order for the sentence to be well-formed; after the DP raises to the specifier of IP to check EPP, it is available for further A-bar movement. (See also Jeoung 2017:28 for a similar argument in Madurese, which also requires the grammatical subject position to be filled.)

It is worth noting that although the subject position generally must be filled in Indonesian, there are exceptions. For example, if a passive clause or adjectival predicate embeds a CP, then the subject position remains empty:

- (132) Sudah dike-tahu-i bahwa ada upaya kudeta.
 already PV-know-Appl C exist effort coup.d’etat
 ‘It is already known that there was an attempted coup d’etat.’ (From Arrahmah News, July 16, 2016, “Erdogan Minta Rakyat Turki Lawan Kudeta Militer.”)
- (133) Sangat jelas bahwa tujuan mereka adalah mem-buat Lee Hsien Loong turun dari
 very clear C goal 3p Cop AV-make Lee Hsien Loong descend from
 jabatan-nya sebagai PM...
 position-D as PM
 ‘It is clear that their goal is to make Lee Hsien Loong step down as PM...’ (Jawa Pos Online, July 4, 2017, “Aroma Politis di Konflik Keluarga PM.”)

The English expletive *it* is used to translate these sentences; Indonesian does not have a corresponding expletive element that occurs in subject position. In (132-133), satisfaction of EPP conflicts with another language-internal constraint: sentential subjects are not possible for my consultants, nor can CPs be extracted with the cleft strategy (see Chapter 2):

- (134) a. Sudah di-umum-kan { bahwa / \emptyset } Ali menang.
 already PV-public-Appl C Ali win
 ‘It was already announced that Ali won.’
- b. *{ Bahwa / \emptyset } Ali menang sudah di-umum-kan.
 C Ali win already PV-public-Appl
 (‘That Ali won was already announced.’)
- c. *{ Bahwa / \emptyset } Ali menang yang sudah di-umum-kan.
 C Ali win C.Foc already PV-public-Appl

(‘That Ali won was already announced.’)

Since CPs are disallowed as subjects in SpecIP, and no other local DP is available to check EPP, the subject position remains unfilled. Furthermore, sub-extraction from the embedded CP is not possible for sentences like (134a), in which no argument has been raised to subject position in the matrix clause:

- (135) *Siapa yang sudah di-umum-kan ___ menang?
who C.Foc already PV-public-Appl win
(‘Who is it that it was already announced (he) won?’)

4.5.3. *Revisiting A-bar extraction in Indonesian-type languages*

Possessor extraction also brings new insight to analyses of Indonesian-type voice systems (like that of Indonesian, Javanese and Madurese). I discuss two issues: the role of Voice in regulating DP movement through the clause, and Cole, Hermon and Yanti’s Voice Agreement Hypothesis for Indonesian.

First, A-bar movement of possessors in Indonesian shows that Voice is concerned with the movement of all DPs, not only verbal arguments, because possessors are arguments of N rather than arguments of V. Whereas morphological voice marking typically indicates the position of verbal arguments (for example, the Theme in an active clause remains low, while the Theme in a passive clause occurs as subject), when nominal extraction has occurred, null voice morphology also serves as a reflex of nominal movement through VoiceP. When sub-extraction of the possessor from object position occurs in Indonesian, Javanese and polite Madurese, the null voice prefix is required (just as when the full object DP is extracted). The object of the verb has not shifted, since the head N, the possessum, remains in its merged position. Yet extraction of a non-argument, the possessor, requires that the voice morphology reflect that a DP has moved through the edge of VoiceP.

It is interesting that movement of the possessor is constrained by Voice in the same way that verbal arguments are; this shows that the voice system regulates the movement of any DP that shifts out of VoiceP. This is not only true of possessors, but also arguments extracted from an embedded

clause (see Chapter 3). This means that the head Voice not only determines argument structure of the clause (i.e. transitivity; or whether an external argument is hosted in its specifier), but also A-bar movement of all VoiceP-internal arguments.

Possessor extraction In Indonesian, Javanese and Madurese also calls for a re-examination of analyses of DP extraction in Indonesian-type languages. Recall that in these languages, null voice morphology marks object extraction. Cole, Hermon and Yanti (2008) propose the Voice Agreement Hypothesis for Indonesian, a language-specific morphological rule that requires agreement between the shifted nominal and the morphological voice marker on the verb. Either the case or the thematic role of the DP must be marked on the verb; a morphological filter prevents conflicting features on the verb and the extracted DP. Possessors present a challenge to this type of analysis because both the object (matrix DP) and its possessor can be extracted with a null voice marker. I assume that abstract genitive or possessive case is assigned to the possessor, while its matrix DP bears abstract accusative case. Under the Voice Agreement Hypothesis, both types of nominals could not extract with the same null voice prefix. A similar argument applies to a mismatch in theta roles between an argument and its possessor.²⁰

Despite the difficulties that possessor extraction poses to this analysis, the Voice Agreement Hypothesis essentially proposes that object extraction requires morphological agreement. In this thesis I have also adopted this general approach, that the null voice morphology required in cases of object extraction is wh-agreement. This type of wh-agreement does not register any particular features of the A-bar moved nominal (at least in Indonesian); neither Case features nor thematic role is morphologically marked in wh-agreement. Like other wh-agreement in other languages, DP-internal features are also not marked. Rather, it is only movement through the edge of VoiceP that is marked by the null prefix.

²⁰This analysis faces further challenges when a subject undergoes long distance extraction from an embedded clause (see Saddy 1991 for extended discussion). Long-distance subject movement also requires a bare verb, collapsing the case/thematic distinction between subjects and objects.

4.5.4. *Implications for the left periphery in Indonesian*

Possessor extraction also provides new observations about the organization of the left periphery of the clause. I have argued that the availability of possessor extraction from object position, and the extraction of objects in general, is regulated low in the clause, by Voice. If a nominal can escape VoiceP (or is already on the edge of VoiceP), then it can undergo A-bar extraction. Languages like familiar Madurese, in which subjects can be extracted but objects cannot, do not allow Voice to bear an edge feature [EPP-D] that raises nominals to SpecVoiceP.

Other theories offer an alternative explanation for the impossibility of object extraction in a language such as familiar Madurese. Some authors have accounted for “subject-only” extraction in some Austronesian languages (like familiar Madurese) by invoking a variation of Feature Inheritance (Chomsky 2008 and subsequent). Under C-T Feature Inheritance, uninterpretable features on T originate on C but are passed down (inherited). Among the relevant set of features involved, A-bar features remain on C while A features are passed down to T, thus deriving a distinction between A and A-bar positions and movement.

Inheritance is motivated by the close relationship between C and T cross-linguistically. In Austronesian languages, Inheritance is further motivated by the properties of the highest or leftmost argument, which appears to have both A and A-bar properties. This follows if C and T are not distinct in these languages, but rather form a combined projection that hosts a single subject/topic/pivot argument. This suggests that both A and A-bar features remain on a single head, and that Inheritance does not apply. (See Richards 2001; Pearson 2005; Rackowski and Richards 2005; Fortuny 2008; Legate 2011, 2014 for discussion in various Austronesian languages; also see Aldridge (2017), Aldridge, to appear for an alternative view on the extraction restriction.)

The theory of Feature Inheritance can be applied to familiar Madurese to account for the impossibility of object extraction in the following way. Legate (2011, 2014) proposes Underinheritance, in which the formal features on C fail to be inherited by T. The result is that CP and TP are not projected separately, but rather form a single combined projection: let us call this CTP. The

single specifier of CTP allows for only one argument to occupy this high position (a subject DP or topic DP), leaving no structural position for another DP. Therefore, object extraction is impossible in active clauses, because the external argument already occupies the single position available at the left periphery; this derives the subject-only restriction. The effect of Under-inheritance is that only one nominal can occur in a high position in the clause, which is empirically attested in a number of Austronesian languages including familiar Madurese.

Alternatively, recent theories of head-splitting (Martinović 2015; see also Erlewine 2017) posit that some of the features on a functional head can split off and re-project a new head. Head-splitting of CTP can also derive the traditional division between TP and CP: T is merged first, but some of its features split into a separate head, C. If this head does not split however, the single head (CT) could host only one argument in its specifier; just as with under-inheritance, a preverbal subject would prevent another argument from moving to a high position.

According to Under-inheritance and failure of head-splitting, a structural explanation derives the subject-only restriction on extraction. If this type of structural account is correct for a language that does not allow any object extraction, then we expect that familiar Madurese could not host two arguments at the left periphery of the clause, because a single CTP projection hosts only one argument. Recall that objects cannot extract in familiar Madurese (55c is repeated here as (136):

(136) Familiar Madurese

*Apa se ale { m-acah / bacah } ____?
 what C.Foc younger.sibling AV-read read
 ('What did little brother read?')

By examining possessor sub-extraction from subject position however, we see that a structural account cannot be correct:

(137) oreng se kalambhi-nah ____ e-sasa
 person C.Foc clothing-D PV-wash
 'the person whose clothing was washed'

The possessor and the possessum simultaneously occur in separate A and A-bar positions, the latter separated from the rest of the clause by the relative morpheme *se*. Possessor extraction, therefore, shows that under-inheritance and head splitting do not account for the impossibility of object extraction, at least in familiar Madurese. In familiar Madurese, a traditional split between C and T(I) remains, with (at least) two positions for nominals at the left periphery of the clause (SpecCP and SpecIP). For other languages of the area that do not allow object extraction, possessor extraction provides a useful diagnostic to test whether the left periphery of the clause allows two nominals.

4.6. Chapter summary

I have shown that syntactic movement, rather than base generation, derives the surface position of external possessors in these languages. Possessors first escape their possessive DPs, triggering *wh*-agreement on the possessum; the suffix is not a resumptive pronoun as previously assumed, but pronunciation of the head D. The possessor is then extracted through the edge of VoiceP: a null voice prefix on the verb not only marks the extraction of an object, but must also mark the extraction of a possessor from object position. This is evidence that the functional head Voice regulates A-bar extraction of all nominals passing through its specifier.

Possessor extraction provides support for D, Voice and C as phase heads in Indonesian (and related languages that have Indonesian-type syntax such as Javanese and Madurese). I have argued that an extracted possessor undergoes successive-cyclic movement through the edge of DP, then the edge of VoiceP, before landing in its surface position in CP. This syntactic movement is driven by the feature [EPP-D] on each phase head. Another implication of possessor extraction data is that the organization of the left periphery in these languages shows a structural distinction between CP and TP(IP), with a traditional division of features associated with C and T; theories of under-inheritance of features or head-splitting do not adequately account for subject-only extraction. In sum, this paper has attempted to bring new data involving possessors to the rich discussion on voice and nominal extraction in Indonesian-type languages.

Finally, this chapter has taken seriously the optional realization of *-nya* in Indonesian posses-

sive DPs. By treating the suffix as an optional morpheme in internal possession, I re-analyze *-nya* as an agreement morpheme, rather than a pronoun. This approach also allows comparison between Indonesian and related languages such as Javanese and Madurese, which have analogous suffixes in possessive DPs.

CHAPTER 5 Conclusion

This dissertation has investigated several morphemes that display surface optionality in Indonesian, with a focus on syntactic environments that affect their variable realization. The investigation contributes novel data in Indonesian (as well as Javanese and Madurese, in Chapter 4), to describe syntactic operations and contexts that affect optional realization. Several of the patterns presented, concerning the form and distribution of variable morphemes, were previously unreported in the literature; I highlight these here.

For the system of declarative complementizers (Chapter 2), I have demonstrated, based on patterns in both local extraction and long-distance extraction, that *yang* is a form of C that occurs when a nominal undergoes A-bar movement and lands in its specifier. I also observed that *yang* and null C display characteristic properties of morphological wh-agreement, and that the pattern of wh-agreement in Indonesian has not been previously attested: the highest C crossed by movement is marked differently from intermediate C. Another novel observation is that *bahwa* is disallowed in wh-in situ questions, while *kalau* can occur in some cases, a fact that is not accounted for by A-bar movement. Based on these patterns, I developed an analysis of wh-phrases and the structure of wh questions in Indonesian, incorporating insights from pseudo-cleft structure.

For the verbal prefixes *meN-* and *ber-* (Chapter 3), I observed that the properties that have been reported for these morphemes fall into two types, deterministic factors and non-deterministic factors. This distinction revealed that several syntactic and semantic conditions disallow optional realization, requiring either *meN-/ber-* or a bare verb. In contrast, non-deterministic factors only affect probabilistic realization; this is implemented via variable rule. I also argued that the loss of optionality in case of A-bar movement is a type of wh-agreement on the verb. I strengthened the case for wh-agreement by showing that A movement does not cause *meN-* to delete, contrary to prior assumptions.

In Chapter 4, I presented novel patterns of possessor extraction in Indonesian, Javanese and Madurese. I demonstrated that possessors can escape their DPs via successive-cyclic A-bar move-

ment, and showed that this results in a cross-linguistically unusual case of wh-agreement. I developed a novel analysis for the suffix *-nya* that is required in possessor extraction. *-Nya* was previously assumed to be a possessive pronoun; under my analysis, *-nya* is a pronunciation of the head D, and instantiates wh-agreement in the nominal domain. I also argued that wh-agreement provides evidence that DP is a phase in this language.

Each case study stands independently as an investigation of a particular aspect of optional morphology in this language, and the findings presented in each case further our understanding of the syntax and morphology of Indonesian complementizers, verbal prefixes and possessive DPs. The individual case studies are also unified under an analysis of morphological wh-agreement. I have claimed that wh-agreement occurs across three domains in Indonesian: complementizers, verbs and nominals. This analysis depends on acknowledging optional morphemes within a syntactic analysis. In the case of complementizers, the loss of optionality is indicative of wh-agreement. In the verbal domain, wh-agreement on Voice is marked by the loss of optionality in *meN-/ber-*. Likewise, the analysis of wh-agreement in the nominal domain depends on the observation that the suffix *-nya* is an optional form in possessive DPs in colloquial Indonesian. By considering optional realization as a property that merits attention, this dissertation presents an approach to analyzing optional morphemes which integrates syntactic and extra-syntactic factors that affect surface variability.

APPENDIX: List of Abbreviations

| | |
|--------|-------------------|
| Abil | abilitative |
| Advers | adversative |
| Appl | applicative |
| AV | active voice |
| Caus | causative |
| Class | classifier |
| Cop | copula |
| Def | definite |
| Excl | exclusive |
| Exist | existential |
| Foc | focus |
| Imper | imperative |
| Incl | inclusive |
| Invol | involtive |
| Loc | locative |
| MV | middle voice |
| Neg | negation |
| Nmlz | nominalizer |
| OP | operator |
| OV | object voice |
| Perf | perfective |
| Pl | plural marker |
| Prog | progressive |
| PV | passive voice |
| Q | question particle |
| Redup | reduplicant |
| Stat | stative |
| VAR | variable |

Bibliography

- Aldridge, E. (2004). *Ergativity and Word Order in Austronesian Languages*. PhD thesis, Cornell University.
- Aldridge, E. (2008). Phase-based account of extraction in Indonesian. *Lingua*, 118(10):1440–1469.
- Aldridge, E. (2017). Phi-feature competition: A unified approach to the Austronesian extraction restriction. *Proceedings of the 52nd meeting of the Chicago Linguistics Society (CLS52)*, 52.
- Amaechi, M. and Georgi, D. (2017). On subject/non-subject extraction asymmetries in igbo. In *Handout, Quirks of subject extraction workshop, National University of Singapore*.
- Anderson, E. A. (1983). The meaning of variation in Indonesian. *Linguistic Studies of Indonesian and Other Languages in Indonesia (NUSA 15)*, pages 1–26.
- Aoun, J. and Li, Y. (1993). Wh-elements in situ: Syntax or LF? *Linguistic Inquiry*, 24(2):199–238.
- Arka, I. (2000). Control and argument structure: Explaining control into subject in Indonesian. Paper presented at the 4th International Malay and Indonesian Symposium.
- Arka, I. (2002). Voice systems in the Austronesian languages of Nusantara: Typology, symmetry and undergoer orientation. URL <http://hdl.handle.net/1885/41058>.
- Arka, I. (2003). *Balinese Morphosyntax: A Lexical-functional Approach*. The Australian National University.
- Arka, I. (2008). Voice and being core: Evidence from (eastern) Indonesian languages. In Austin, P. K. and Musgrave, S., editors, *Voice and Grammatical Relations in Austronesian Languages*. CSLI Publications.
- Arka, I. (2009). On the zero (voice) prefix and bare verbs in Austronesian languages of Nusa Tenggara, Indonesia. In Evans, B., editor, *Discovering history through language: papers in honour of Malcolm Ross*, pages 247–270. Pacific Linguistics, Canberra.

- Arka, I. and Manning, C. (1998). On the three subjects in Indonesian: Evidence from binding. In *On-line Proceedings of the LFG98 Conference*, <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG3-1998/lfg98arkamanning.pdf>. CSLI Publications.
- Arka, I. and Manning, C. (2008). Voice and grammatical relations in Indonesian: A new perspective. In Austin, P. K. and Musgrave, S., editors, *Voice and Grammatical Relations in Austronesian Languages*, pages 45–69. CSLI Publications.
- Arka, I. and Wouk, F. (2014). Voice-related constructions in the Austronesian languages of Flores. In *Argument Realisations and Related Constructions in Austronesian Languages: Papers from 12-ICAL*, pages 313–333. Asia-Pacific Linguistics.
- Baker, M. (1999). External possession in mohawk: Body parts, incorporation and argument structure. In Payne, D. and Barshi, I., editors, *External possession*, pages 293–323. John Benjamins, Philadelphia.
- Bayer, J. (2006). Wh-in-situ. In Everaert, M. and van Riemsdijk, H., editors, *The Blackwell Companion to Syntax*, volume I, pages 376–. Blackwell, Malden, MA.
- Beck, S. (2006). Intervention effects follow from focus interpretation. *Natural Language Semantics*, 14(1):1–56.
- Beck, S. and Kim, S.-S. (2006). Intervention effects in alternative questions. *The Journal of Comparative Germanic Linguistics*, 9(3):165–208.
- Blust, R. (2013). *The Austronesian Languages*. Asia-Pacific Linguistics.
- Bobaljik, J. (2002). A-chains at the PF-interface: Copies and ‘covert’ movement. *Natural Language and Linguistic Theory*, 20:197–267.
- Boeckx, C. (2003). *Islands and Chains: Resumption as Stranding*, volume 63. John Benjamins.
- Bošković, Željko. (2014). Now I’m a phase, now I’m not a phase: On the variability of phases with

- extraction and ellipsis. *Linguistic Inquiry*, 45(1):27–89.
- Bresnan, J. (1977). Variables in the theory of transformations. In Peter W. Culicover, Thomas Wasow, A. A., editor, *Formal Syntax*, pages 157–196. Academic Press, New York.
- Cheng, L. (2003). Wh-in-situ. *Glott International*, 7(4):103–109.
- Cheng, L. (2009). Wh-in-situ, from the 1980s to now. *Language and Linguistics Compass*, 3:767–791.
- Cheng, L. and Rooryck, J. (2000). Licensing wh-in-situ. *Syntax*, 3(1).
- Choe, J.-W. (1984). LF WH-movement: A case of pied-piping? Ms., University of Massachusetts Amherst.
- Chomsky, N. (1977). On WH-movement. In Peter W. Culicover, T. W. and Akmajian, A., editors, *Formal Syntax*, pages 71–133. Academic Press.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press.
- Chomsky, N. (2000). Minimalist inquiries. The framework. In Martin, R., Michaels, D., and Uriagereka, J., editors, *Step by Step*. Essays on Minimalist Syntax in Honor of Howard Lasnik, pages 89–155. MIT Press, Cambridge, MA.
- Chomsky, N. (2001). Derivation by phase. In Kenstowicz, M., editor, *Ken Hale*. A life in language, pages 1–52. MIT Press, Cambridge, MA.
- Chomsky, N. (2008). On phases. In Freidin, R., Otero, C. P., and Zubizarreta, M. L., editors, *Foundational Issues in Linguistic Theory*. Essays in Honor of Jean-Roger Vergnaud, pages 133–166. MIT Press, Cambridge, MA.
- Chomsky, N. and Lasnik, H. (1977). Filters and control. *Linguistic Inquiry*, 8:425–504.
- Chung, S. (1976). On the subject of two passives in Indonesian. In Li, C. N., editor, *Subject and Topic*. Academic Press.

- Chung, S. (1982). Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry*, 13:39–77.
- Chung, S. (1998). *The Design of Agreement*. The University of Chicago Press.
- Chung, S. (2008). Possessors and definiteness effects in two Austronesian languages. In Matthewson, L., editor, *Quantification: A Cross-linguistic Perspective*, number 64 in North Holland Linguistic Series: Linguistic Variations, pages 179–224. Emerald, Chicago.
- Chung, S. and Georgopoulos, C. (1988). Agreement with gaps in Chamorro and Palauan. In Barlow, M. and Ferguson, C. A., editors, *Agreement in Natural Language*. Center for the Study of Language and Information, Stanford, CA.
- Cole, P. and Hermon, G. (1998). The typology of WH-movement: WH-questions in Malay. *Syntax*, 1:221–258.
- Cole, P. and Hermon, G. (2005). Subject and non-subject relativization in Indonesian. *Journal of East Asian Linguistics*, 14:59–88.
- Cole, P., Hermon, G., and Aman, N. (1999). Clefted questions in Malay. Ms.
- Cole, P., Hermon, G., and Tjung, Y. (2005). How irregular is WH in situ in Indonesian? In *Studies in Language*, volume 29, pages 553–581. John Benjamins.
- Cole, P., Hermon, G., Tjung, Y., Sim, C.-Y., and Kim, C. (2008a). A binding theory exempt anaphor. *Trends in Linguistics*, pages 577–590.
- Cole, P., Hermon, G., and Yanti (2008b). Voice in Malay/Indonesian. *Lingua*, 118:1500–1553.
- Coon, J. (2009). Interrogative possessors and the problem with pied-piping in Chol. *Linguistic Inquiry*, 40(1):165–175.
- Culicover, P. W. (1993). Evidence against ECP accounts of the that-t effect. *Linguistic Inquiry*, 24:557–561.
- Dardjowidjojo, S. (1978). *Sentence Patterns of Indonesian*. University of Hawaii Press.

- Davies, W. (2005). The richness of Madurese voice. In Arka, I. and Ross, M., editors, *The Many Faces of Austronesian Voice Systems: Some New Empirical Studies*. The Australian National University, Canberra.
- Davies, W. D. (2010). *A Grammar of Madurese*. DeGruyter Mouton.
- Deal, A. R. (2013a). External possession and possessor raising. In Everaert, M. and Riemsdijk, H. V., editors, *The Companion to Syntax*. Wiley-Blackwell, 2nd edition.
- Deal, A. R. (2013b). Possessor raising. *Linguistic Inquiry*, 44(3):391–432.
- Donohue, M. (2007). Variation in voice in Indonesian/Malay: Historical and synchronic perspectives. In Matsumoto, Y., Oshima, D., Robinson, O., and Sells, P., editors, *Diversity in Language: Perspectives and Implications*, pages 71–129. CSLI Publications.
- Donohue, M. (2008). Malay as a mirror of Austronesian: Voice development and voice variation. *Lingua*, 118:1470–1499.
- Embick, D. (2004). On the structure of resultative participles in English. *Linguistic Inquiry*, 35(3):355–392.
- Erlewine, M. (2017). Why the null complementizer is special in complementizer-trace effects. In Claire Halpert, H. K. and van Urk, C., editors, *A Pesky Set: Papers for David Pesetsky*, volume 80, pages 371–380. MIT Working Papers in Linguistics, Cambridge, MA.
- Fortin, C. (2006). Reconciling meng- and NP movement in Indonesian. In *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistic Society*, volume 32, pages 47–58.
- Fortin, C. (2007). *Indonesian Sluicing and Verb Phrase Ellipsis: Description and Explanation in a Minimalist Framework*. PhD thesis, University of Michigan.
- Fortin, C. (2009). On the left periphery in Indonesian. In Chung, S., Finer, D., Paul, I., and Potsdam, E., editors, *Proceedings of the 16th Meeting of the Austronesian Formal Linguistics Association*, volume 3, pages 29–43.

- Fortuny, J. (2008). *The Emergence of Order in Syntax*, volume 119. John Benjamins.
- Gavruseva, E. (2000). On the syntax of possessor extraction. *Lingua*, 110(10):743–772.
- Georgopoulos, C. (1985). Variables in Palauan syntax. *Natural Language and Linguistic Theory*, 3(1):59–94.
- Gil, D. (2002). The prefixes di- and N- in Malay/Indonesian dialects. In Wouk, F. and Ross, M., editors, *The History and Typology of Western Austronesian Voice Systems*, pages 241–283. Pacific Linguistics.
- Groat, E. and O’Neil, J. (1996). Spell-out at the LF interface: Achieving a unified syntactic and computational system in the Minimalist framework. In Abraham, W., Epstein, S., Thrainsson, H., and Zwart, C., editors, *Minimal Ideas: Syntactic Studies in the Minimalist Framework*, pages 113–139. John Benjamins.
- Guilfoyle, E., Hung, H., and Travis, L. (1992). Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory*, pages 375–414.
- Haik, I. (1990). Anaphoric, pronominal and referential infl. *Natural Language and Linguistic Theory*, 8(3):347–374.
- Hale, K. and Keyser, S. (2002). *Prolegomenon to a Theory of Argument Structure*, volume 39. MIT Press.
- Halle, M. and Marantz, A. (1993). Distributed Morphology and the pieces of inflection. In Hale, K. L. and Keyser, S. J., editors, *The View From Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, number 111-176. MIT Press.
- Halle, M. and Marantz, A. (1994). Some key features of Distributed Morphology. In *MIT Working Papers in Linguistics*, volume 21, pages 275–288. MIT Press.
- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10:41–53.

- Han, C. (2000). *The Structure and Interpretation of Imperatives: Mood and Force in Universal Grammar*. Psychology Press.
- Han, C. and Romero, M. (2004). Disjunction, focus and scope. *Linguistic Inquiry*, 35(2):179–217.
- Harley, H. (2013). External arguments and the Mirror Principle: On the distinctness of Voice and v. *Lingua*, 125:34–57.
- Hidajat, L. (2013). On the distribution of the nasal prefix N- in basilectal Jakarta Indonesian. *NUSA*, 55.
- Himmelman, N. P. (2002). Voice in western Austronesian: An update. In Wouk, F. and Ross, M., editors, *The History and Typology of Western Austronesian Voice Systems*. Pacific Linguistics, Canberra.
- Himmelman, N. P. (2005). The Austronesian languages of Asia and Madagascar: Typological characteristics. In Adelaar, K. A. and Himmelmann, N. P., editors, *The Austronesian languages of Asia and Madagascar*, pages 110–181. Routledge.
- Hopper, P. J. and Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, 56(2):251–299.
- Huang, C. (1982). *Logical Relations in Chinese and the Theory of Grammar*. PhD thesis, MIT, Cambridge, MA.
- Ishii, T. (2004). The phase impenetrability condition, the vacuous movement hypothesis and that-t effects. *Lingua*, 114(2):183–215.
- Jeoung, H. (2017). On the number of voices in Madurese. *Journal of the Southeast Asian Linguistic Society*.
- Jeoung, H. (2018). Revisiting crossed control in Indonesian. Presentation given at the International Symposium on Malay/Indonesian Linguistics, May 11-12, 2018, UCLA.

- Jeoung, H. and Biggs, A. (2017). Variants of Indonesian prepositions as intra-speaker variability at PF. In Djaerv, K. and Goodwin-Davies, A., editors, *University of Pennsylvania Working Papers in Linguistics*, volume 22.2.
- Kader, M. (1976). *The Syntax of Malay Interrogatives*. PhD thesis, Simon Fraser University.
- Kana, M. (1983). The syntax and semantics of verbal prefixes in Bahasa Indonesia. In *Papers from the Third International Conference on Austronesian Linguistics*, volume 4, pages 21–33. Dept. of Linguistics, Australian National University.
- Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, 1(1):3–44.
- Kaswanti Purwo, B. (1989). Diatesis di dalam Bahasa Indonesia: Telaah wacana [Voice in Indonesian: A study in discourse]. In *Serpih-serpih Telaah Pasif Bahasa Indonesia*, pages 345–441. Kanisius, Yogyakarta.
- Kaswanti Purwo, B. (1995). The two proto-types of ditransitive verbs: The Indonesian evidence. In W. Abraham, T. Givón, S. T., editor, *Discourse Grammar and Typology: Papers in Honor of John W.M. Verhaar*, pages 77–99. John Benjamins, Philadelphia.
- Keenan, E. L. (1972). Relative clause formation in malagasy. In et al., P. M. P., editor, *The Chicago Witch Hunt: Papers from the Relative Clause Festival*, pages 169–189. Chicago Linguistics Society.
- Koontz-Garboden, A. and Beavers, J. (2017). Change of state verbs and the semantics of roots. In *Proceedings of the 34th West Coast Conference on Formal Linguistics*, Summerville. Cascadilla Proceedings Project.
- Kridalaksana, H. (1975). Beberapa ciri Bahasa Indonesia Standar. *Majalah Pengajaran Bahasa dan Sastra*, 1(1).
- Kroeger, P. (1993). *Phrase Structure and Grammatical Relations in Tagalog*. CSLI Publications.
- Kroeger, P. (2009). Malagasy clefts from a Western Malayo-Polynesian perspective: Comments on

- the paper by Hans-Martin Gärtner. *Natural Language and Linguistic Theory*, 27:817–838.
- Kroeger, P. (2014). External negation in Malay/Indonesian. *Language*, 90(1).
- Kurniawan, F. (2016). Nasal assimilation and substitution in Standard Indonesian: Evidence from loanwords production task. *Linguistik Indonesia*, 34(1):31–55.
- Kurniawan, F. O. (2015). Nasal assimilation in Jakarta Indonesian. In Camp, A., Otsuka, Y., Stabile, C., and Tanaka, N., editors, *Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association*, pages 149–166.
- Kushartanti, B. (2014). *The Acquisition of Stylistic Variation by Jakarta Indonesian Children*. PhD thesis, Utrecht University.
- Labov, W. (1969). Contraction, deletion and inherent variability of the English copula. *Language*, 45(4):715–762.
- Landau, I. (1999). Possessor raising and the structure of vp. *Lingua*, 107:1–37.
- Larson, R. (1985). On the syntax of disjunction scope. *Natural Language and Linguistic Theory*, 3:217–264.
- Lasnik, H. and Saito, M. (1984). On the nature of proper government. *Linguistic Inquiry*, 15:235–289.
- Legate, J. A. (2003). Some interface properties of the phase. *Linguistic Inquiry*, 34(3):506–515.
- Legate, J. A. (2011). Under-inheritance. NELS handout.
- Legate, J. A. (2012). Subjects in Acehnese and the nature of the passive. *Language*, 88(3):495–525.
- Legate, J. A. (2014). *Voice and v: Lessons from Acehnese*. MIT Press.
- Macdonald, R. and Dardjowidjojo, S. (1967). *Indonesian Reference Grammar*. Georgetown University Press.

- Martinović, M. (2015). *Feature Geometry and head-splitting: Evidence from the morphosyntax of the Wolof clausal periphery*. PhD thesis, University of Chicago.
- Massam, D. (2003). Questions and the left periphery in Niuean. *Cornell Working Papers in Linguistics*, 19:94–106.
- McCloskey, J. (2001). The morphosyntax of WH-extraction in Irish. *Journal of Linguistics*, 37(1):67–100.
- McCloskey, J. (2002). Resumption, successive cyclicity and the locality of operations. In Epstein, S. and Seely, T., editors, *Derivation and Explanation in the Minimalist Program*, volume 4, chapter 8, pages 184–226. Blackwell, Oxford.
- McKinnon, T., Cole, P., and Hermon, G. (2011). Object agreement and 'pro-drop' in Kerinci Malay. *Language*, 87(4):715–750.
- Musgrave, S. (2001). *Non-subject arguments in Indonesian*. PhD thesis, University of Melbourne.
- Nishigauchi, T. (1984). Japanese LF: subjacency vs. ECP. In *Seoul Papers On Formal Grammar*, pages 71–105. Linguistic Society of Korea.
- Nomoto, H. (2006). Voice in colloquial Malay relatives. In Tsuruga, Y., Miyake, N., and Kawaguchi, Y., editors, *Gengojouhougaku Kenkyuuhoukoku*, volume 12, pages 97–116, Tokyo. Tokyo University of Foreign Studies.
- Nomoto, H. (2013). On the optionality of grammatical markers: A case study of voice marking in Malay/Indonesian. In Adelaar, A., editor, *Voice Variation in Austronesian Languages of Indonesia*, volume 54 of *NUSA*, pages 121–143.
- Nomoto, H. and Shoho, I. (2007). Voice in relative clauses in Malay: A comparison of written and spoken language. In Yuji Kawaguchi, Toshihiro Takagaki, N. T. and Tsuruga, Y., editors, *Corpus-based Perspectives in Linguistics*, pages 353–370. John Benjamins, Amsterdam.
- Nomoto, H. and Wahab, K. A. (2014). Person restriction on passive agents in Malay: Information

- and syntax. *NUSA: Linguistic Studies of Languages In and Around Indonesia*, 57:31–50.
- Nuriah, Z. (2004). The relation of verbal affixes meN- and -kan with argument structure. MA Thesis, Utrecht University.
- Paul, I. (2001). Concealed pseudo-clefts. *Lingua*, 111:707–727.
- Payne, D. and Barshi, I., editors (1999). *External Possession*, volume 39. John Benjamins.
- Pearson, M. (2001). *The Clause Structure of Malagasy: A Minimalist Approach*. PhD thesis, UCLA.
- Pearson, M. (2005). The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory*, 23(2):381–457.
- Perlmutter, D. (1968). *Deep and Surface Structure Constraints in Syntax*. PhD thesis, MIT.
- Perlmutter, D. (1971). *Deep and Surface Constraints in Generative Grammar*. Holt, Rinehart and Winston, New York.
- Perlmutter, D. (1978). Impersonal passives and the unaccusative hypothesis. In *Annual meeting of the Berkeley Linguistics Society*, volume 4, pages 157–190.
- Perlmutter, D. and Postal, P. (1983). The relational succession law. In Perlmutter, D., editor, *Studies in Relational Grammar 1*, pages 30–80. University of Chicago Press.
- Pesetsky, D. (1982). Complementizer-trace phenomena and the nominative islands condition. *The Linguistic Review*, 1(3):297–344.
- Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. In Reuland, E. and ter Meulen, A., editors, *The Representation of (In)definiteness*, pages 98–129. MIT Press, Cambridge, MA.
- Pesetsky, D. (2017). Complementizer-trace effects. In *The Wiley Blackwell Companion to Syntax, Second Edition*, pages 1–34. Wiley-Blackwell.
- Pesetsky, D. and Torrego, E. (2000). T-to-C Movement: Causes and consequences. In Kenstowicz,

- M., editor, *Ken Hale: A Life in Language*. MIT Press.
- Pylkkanen, L. (2002). *Introducing Arguments*. PhD thesis, MIT.
- Pylkkanen, L. (2008). *Introducing Arguments*. MIT Press.
- Rackowski, A. (2002). *The Structure of Tagalog: Specificity, Voice and the Distribution of Arguments*. PhD thesis, MIT.
- Rackowski, A. and Richards, N. (2005). Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry*, 36(4):565–599.
- Reintges, C. H., LeSourd, P., and Chung, S. (2006). Movement, wh-agreement and apparent wh-in-situ. In Cheng, L. and Corver, N., editors, *WH-Movement: Moving On*, volume 42, pages 165–194. MIT Press.
- Richards, N. (1999). Featural cyclicity and the ordering of multiple specifiers. *Current Studies in Linguistics*, 32:127–158.
- Richards, N. (2001). *Movement in Language: Interactions and Architectures*. Oxford University Press.
- Rizzi, L. (1982). Negation, wh-movement and the null subject parameter. In *Issues in Italian Syntax*, pages 117–185. Foris, Dordrecht Holland.
- Rizzi, L. (1997). *The fine structure of the left periphery*, pages 281–337. Springer Netherlands.
- Rooth, M. and Partee, B. (1982). Conjunction, type ambiguity and wide scope or. In *Proceedings of the first west coast conference on formal linguistics*, volume 1. Linguistics Dept., Stanford University.
- Ross, J. (1986). *Infinite Syntax*. Ablex Publishing Corporation.
- Saad, N., Jalaluddin, N., and Ahmad, Z. (2015). Penelitian semula makna imbuhan meN-: Kajian eksplisit vs implisit. In *Seminar Linguistik Kebangsaan (SLiK2015)*, Kuala Lumpur. Dewan

Bahasa dan Pustaka.

- Saddy, D. (1991). WH scope mechanisms in Bahasa Indonesia. In Cheng, L. L. S. and Demirdash, H., editors, *MIT Working Papers in Linguistics: More Papers on WH-Movement*, number 15, pages 183–218. MIT Press, Cambridge, MA.
- Sato, Y. (2008a). The distribution of active voice morphology in Javanese and vP phases. In *Proceedings of WECOL 2007*, pages 235–246.
- Sato, Y. (2008b). *Minimalist Interfaces: Selected Issues in Indonesian and Javanese*. PhD thesis, University of Arizona.
- Sato, Y. (2012). Successive cyclicity at the syntax-morphology interface: Evidence from Standard Indonesian and Kendal Javanese. *Studia Linguistica*, 66:32–57.
- Sneddon, J. (2003). *The Indonesian Language: Its History and Role in Modern Society*. UNSW Press, Sydney.
- Sneddon, J. N. (1996). *Indonesian: A Comprehensive Grammar*. Routledge.
- Sneddon, J. N. (2006). *Colloquial Jakartan Indonesian*. Pacific Linguistics, The Australian National University.
- Sneddon, J. N., Adelaar, A., Djenar, D. N., and Ewing, M. (2012). *Indonesian: A Comprehensive Grammar*. Routledge.
- Soh, H. (2010). Voice and aspect: Some notes from Malay. In *Proceedings of the workshop on Indonesian-type voice system*, pages 25–35. Tokyo University of Foreign Studies.
- Soh, H. and Nomoto, H. (2011). Malay verbal prefix meN- and the unergative/unaccusative distinction. *Journal of East Asian Linguistics*, 20(1):77–106.
- Soh, H. L. (1998). Certain restrictions on A-bar movement in Malay. In Pearson, M., editor, *Proceedings of the Third and Fourth Meetings of the Austronesian Formal Linguistics Association*

- 1996-1999, pages 295–308. Dept of Linguistics, UCLA.
- Soh, H. L. (2013). Voice and aspect: Some notes from Malaysian Malay. *NUSA: More on Voice in Languages of Indonesia*, 54:159–173.
- Soh, H. L. and Nomoto, H. (2009). Progressive aspect, the verbal prefix meN- and stative sentences in Malay. *Oceanic Linguistics*, 48:148–171.
- Soh, H. L. and Nomoto, H. (2015). Degree achievements, telicity and the verbal prefix meN- in Malay. *Journal of Linguistics*, 51:147–183.
- Son, M.-J. and Cole, P. (2008). An event-based account of -kan constructions in Standard Indonesian. *Language*, 84(1):120–160.
- Szabolcsi, A. (1992). *The Noun Phrase*. PhD thesis, UCLA.
- Tuller, L. (1986). *Bijective Relations in Universal Grammar*. PhD thesis, UCLA.
- van Urk, C. and Richards, N. (2015). Two components of long-distance extraction: Successive-cyclicity in Dinka. *Linguistic Inquiry*, 46(1):113–155.
- Voskuil, J. E. (1996). *Comparative Morphology: Verb Taxonomy in Indonesian, Tagalog and Dutch*. Holland Academic Graphics.
- Voskuil, J. E. (2000). Indonesian voice and A-bar movement. In Ileana Paul, V. P. and Travis, L., editors, *Formal Issues in Austronesian Linguistics*, pages 195–213. Kluwer Academic Publishers.
- Wallace, S. (1979). Voice, mode or aspect? the semantics of verbal prefixes in Jakarta Malay. In van Coetsem, F. and Waugh, L. R., editors, *Contributions to Grammatical Studies, Semantics and Syntax (Cornell Linguistic Contributions)*, volume 2, pages 151–178. Brill, Leiden.
- Watanabe, A. (1992). Subjacency and S-structure movement of wh-in-situ. *Journal of East Asian Linguistics*, 1:255–291.

- Watanabe, A. (1996). *Case Absorption and WH-Agreement*, volume 37 of *Studies in Natural Language and Linguistic Theory*. Kluwer Academic Publishers, Dordrecht.
- Watanabe, A. (2001). Wh-in-situ languages. In Baltin, M. and Collins, C., editors, *The Handbook of Contemporary Syntactic Theory*, pages 203–225. Blackwell, Oxford.
- Wolff, J. (1990). *Formal Indonesian*. Cornell University, Southeast Asia Program, Ithaca, NY.
- Wolff, J. U., Oetomo, D., and Fietkiewicz, D. (1992). *Beginning Indonesian Through Self-Instruction*, volume 3. Cornell University, Southeast Asia Program.
- Wouk, F. (2004). Transitive bare verb stems in spoken Jakarta Indonesian. *NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia*, 53:29–46.
- Wouk, F. and Ross, M. (2002). *The History and Typology of Western Austronesian Voice Systems*. Pacific Linguistics, The Australian National University.
- Yanti (2010). *A Reference Grammar of Jambi Malay*. PhD thesis, University of Delaware.
- Zaenen, A. (1983). On syntactic binding. *Linguistic Inquiry*, 14:469–504.