

MINIMALIST INTERFACES: SELECTED ISSUES IN INDONESIAN AND JAVANESE

by

Yosuke Sato

A Dissertation to the Faculty of the
DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2008

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Yosuke Sato entitled Minimalist Interfaces: Selected Issues in Indonesian and Javanese and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Heidi Harley Date: April 21st, 2008

Andrew Barss Date: April 21st, 2008

Andrew Carnie Date: April 21st, 2008

Simin Karimi Date: April 21st, 2008

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director: Heidi Harley Date: April 21st, 2008

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Yosuke Sato

ACKNOWLEDGEMENTS

Having been a student for about 20 years, time has finally come for me to put an end to this very long journey and thank countless many people who have had influences on my life before I forget who they are. The real reason I am the way I am is because I have been blessed with great teachers, scholars, friends, and families that have provided me with encouragement, trust, care, and love.

My first thanks go to my dissertation committee: Heidi Harley, Andrew Carnie, Andy Barss, and Simin Karimi. Getting Heidi's praise on my work and progress has been the major drive for my life in Tucson. Since I met her back in September 2, 2004, Heidi always has served as my role model; brilliant, kind, friendly, sensitive, highly knowledgeable; has interests in whatever topics I have worked on during my graduate life and always has something very important to say about my work. Among her many good qualities, Heidi is the most humble person I have ever met. She has taken great pains not only as my academic supervisor but also as one human being in walking with me through various stages of my progress, academic or life, sometimes dark and sometimes challenging, so that I am doing okay. I don't really remember how many times I have cried in front of her due to many obstacles that have overwhelmed me, just because she is so caring and kind, but I do remember that she always has listened to me very patiently and gave appropriate advice so I could overcome those obstacles and become a grown-up individual. Her extensive comments, suggestions, criticisms, and questions on all my work have been unparalleled. They often exceed the amount of texts I wrote myself! Nobody should be able to expect such high-quality feedback from anybody in this busy world. In short, she has made me want to be a better scholar. Andrew has also been a very important person for me. He has provided new directions to pursue in my work, suggested a list of many papers that independently came up with the ideas I thought were mine, taught me the "real" state of arts on the journal submission, paper presentation, job applications and interviews, conference organization, and, most importantly, how to manage my life as a scholar without getting too stressed out to produce good work: "Reward yourself; you are worth it!", that is what Andrew said in October 2007. It really worked, Andrew! I will keep this sentence as my mantra whenever I feel I am being overwhelmed with too many stuff to do in the rest of my academic career. Andy has always given me warm encouragement and provided extensive comments on my work since I took his seminar on ellipsis and sluicing in Fall 2006, however busy he was with his little Xander. The material in chapter 3 developed from our discussion on Indonesian sluicing. I will also never forget his wonderful,

flowing lectures on formal semantics and syntax (especially, binding theory). Finally, Simin has just been great in every respect: as my professor, as my colleague, as my teaching supervisor, and as a human being. She read all of some 40 papers I wrote during 2004-2007 and left extensive feedback on each of them. Noticing as early as Fall 2006 that I would be burned out if I had continued my work habit at that time, she immediately told me to stop working and relax, drive, watch, enjoy nature in the desert, and so on. This warning also worked for the better. It is also one of my privileges to have been able to work with Simin on the paper on modality and as her TA for LING 210. I will never forget her warm, motherly smile every time we met!

I would like to thank many other professors at the Department of Linguistics for giving me the best education possible: Amy Fountain, Mike Hammond, Cecile McKee, Terry Langendoen, Janet Nicol, Massimo Piattelli-Palmarini, Adam Ussishkin, Natasha Warner, Andy Wedel, and Mary Willie. Special thanks to Amy for being such a great professor whose teaching philosophy influenced me most, to Andy and Adam for getting rid of my irrational fear of phonology/phonetics and letting me enjoy it, and to Massimo for introducing me to many linguistic giants from all over the world so that I can get the best human networking possible. I also thank the following staff at our department for making my life much easier: Jennifer Columbus, Marian Wiseley, Jennie Bradley, Kimberly Young, and Emery Rosemary.

My thanks, of course, also go to my fellow students at the University of Arizona for sharing joy, happiness, sadness, and many other experiences with me and/or being willing to give me a hand when I need one badly. I thank the following colleagues: Michael Anderson, Jeff Berry, Shannon Bischoff, Jordan Brewer, Lindsay Butler, Jerid Francom, Jason Ginsburg, Erin Good, Jason Haugen, Mosa Hulden, Scott Jackson, Sunjing Ji, Greg Key, Amy LaCross, Charles Lin, David Medeiros, Kumiko Nakamura, Peter Norquest, Naomi Ogasawara, Jaime Parchment, Jeff Punske, Sumayya Racy, Peter Richtsmeier, Dan Siddiqi, Tatyana Slobodchikoff, Azita Taleghani, Alex Trueman, Mercedes Tubino-Blanco, Stacey Oberly, Sylvia Reed and Dainon Woudstra. These people have all marked big impacts on my life but among these great colleagues, I would like to give special thanks to: Mosa for always encouraging me through his peculiar Finnish-style ironical critique of my work, taking me to Frisbees and drink, driving me back and forth between the Tucson Int'l airport and my apartment, and, most importantly, for just being such a brilliant soul mate; to Shannon for helping me overcome many issues, academic or otherwise, through his diverse experiences in life and linguistics; to Jaime and Jeff for giving me shoulders to cry on and general help in

overcoming many obstacles that got in my way; to Kumiko for deep sympathy and kindness as another civilian from Niigata Prefecture in Japan.

I have been benefited from many scholars on many occasions whose comments/criticisms/suggestions/ encouragements via personal discussion, e-mail correspondences, meetings, lectures and many other forms deeply influenced my work and academic progress. Here is some loose list of people I would like to thank: David Adger, Seiki Ayano, Mark Baker, Bob Berwick, Cedric Boeckx, Noam Chomsky, Sandy Chung, Barbara Citko, Peter Culicover, William Davies, Henry Davis, Marcel den Dikken, Anna-Maria Di Sciullo, Malcolm Elliott, Catherine Fortin, Shin Fukuda, Naoki Fukui, Elly van Gleeneden, Donna Gerts, Annastasia Ginnakidou, Kleanthes Grohmann, Youssef Haddad, Nobuko Hasegawa, Hui-Yu Huang, Jiro Inaba, Yuriko Hatori, Kyle Johnson, Taro Kageyama, Ayasha Kidwai, Masaaki Kamiya, Richard Kayne, Maki Kishida, Hideki Kishimoto, Paul Kroeger, Richard Larson, Howard Lasnik, Roger Martin, Bradley McDonnell, Jason Merchant, Shigeru Miyagawa, Tomohiro Miyake, Takashi Munakata, Pam Munro, Masanori Nakamura, Kunio Nishiyama, Masayuki Oishi, Masha Polinsky, Jerrold Saddock, Mamoru Saito, Ivan Ortega-Santos, Lisa Selkirk, Peter Sells, Kayono Shiobara, Tom Stroik, Hisao Tokizaki, Takashi Toyoshima, Lisa Travis, Juan Uriagereka, James Yoon, Akira Watanabe, and Paul Willis. I also wish to take this opportunity to express my deep gratitude to the anonymous reviewers for providing many detailed comments to improve the quality of my articles: *Gengo Kenkyu*, *English Linguistics*, *Journal of East Asian Linguistics*, *Journal of Linguistics*, *Journal of South East Asian Linguistic Society*, *Language*, *Lingua*, *Linguistic Inquiry*, *Natural Language & Linguistic Theory*, *Natural Language Semantics*, *Snippets*, *Syntax*, *Studia Linguistica*, and *The Linguistic Review*.

I have been extremely fortunate to meet many scholars, teachers and colleagues in Sendai and Niigata whose encouragements and support made it possible for me to continue my linguistic research. I thank Masaru Nakamura, Yoshiaki Kaneko, Yoshiki Ogawa, Nobuhiro Miyoshi, Yoshihito Dobashi, Dana Lupsa, Mika Takahashi, Tomohiro Oshima, Takahiro Tozawa, Kazuyuki Nakano, Shinichiro Hatade, Akito Kurogi, and Yukie Hara. I don't know how to thank *Nakamura-sensei* enough for providing me with support and encouragement to make sure that I am doing all right in Tucson since I left Sendai in Fall 2004 and for giving me a fundamental sense of direction in linguistics. I regret that I could not be in Sendai to celebrate his retirement from Tohoku University in March, 2008. I just hope that this dissertation would serve as a small token of my appreciation for his constant guidance and support and as a sign that I accomplished what I promised him to do in America. Special thanks also go to *Kaneko-sensei*, *Ogawa-sensei*, *Miyoshi-san*, and

Dobashi-san for providing many critical comments and suggestions on my half-baked ideas and/or supporting my decision to study abroad. My thanks also extend to Takamichi Aki and Tsuyoshi Oishi for first introducing me to the exciting field of linguistics called *generative grammar* and sending me go to Sendai to study more about this discipline. Thanks also go to Yuso Yoshi for many conversations regarding English learning/teaching and for telling me that English is worth living for, way back when I was 18. Your words will always be in my heart.

My life could not come this far without bread and butter. My first two years of graduate life have been supported by the Fulbright Fellowship provided by the International Institute of Education/IIIE and the Japan-United States Education Commission/JUSEC; the other two years have been supported by Teaching Assistantships from the Department of Linguistics. I am very grateful to these organizations for supporting my life here in Tucson. Special thanks to Mizuho Iwata and Miyuki Ito (from JUSEC), Emily Bosio Gutridge (from Denver IIE), as well as Shadi Moqbel and Selim Ben-Said (my Fulbright fellows), for always making sure that I will complete my program as promised. I also thank the Graduate & Professional Student Council of the University of Arizona for the Professional Opportunities Grant to make my planned conference end with a great success and thereby to start a new tradition that I want to persist after I leave Tucson.

Finally, I want to give deepest thanks to my family back in Niigata, Japan and Kendal, Central Java, Indonesia. I am enormously grateful to my parents, Toyoki Sato and Yuko Sato, my grandparents, Sakae Sato and Fusa Sato, my two sisters Asami Sato and Naomi Sato, and my late great grandmother Torano Sato for their love, trust, and unshakable faith in me. Whether they know what I have done in Tucson, I can at least tell that I had the greatest time here thanks to your support. I favorably remember that they have been always proud to say that I have been doing something *worthwhile* in Tucson. None of my family members in Niigata would have ever imagined that the poor kid who got F's on all subjects in elementary school and whose only good quality was to play *Kendo* well would come this far to finish the PhD program in linguistics. My warm thanks also go to my family in Indonesia who welcomed me into their house with very big smiles at first sight and taught me by example that life contains many lessons that are equally or more important than linguistics. Thank you very much to you all for your care, love, and support.

My life so far has been very dramatic. In retrospect, my stay at SUNY Buffalo, NY during July 14-August 4, 2008 as the pre-academic orientation for graduate students was THE turning point in my life, when I finally met the person I wanted to grow old with and become happy with. (Thanks again to the Fulbright, for that matter, for sending me to Buffalo!) She was to be my ambassador of happiness from Indonesia. She came from

Bowling Green, OH all the way into the southwest desert in Tucson in May 2006 for the sole purpose of living with me and supporting me in every way imaginable until this very moment. She has always been with me in times of sadness and happiness in too many forms to list: as my sister, as my consultant, as my mother, as my wife, as my best cook, as my best teacher, as my best soulmate, as my best colleague, and as my everything good. She understands everything I would say before I actually do; she understands every emotion I have; her love and care for me is unparalleled. She has many superb qualities but her inner beauty is what I respect the most. I feel in tremendous awe and at a loss for words when I recall that such a great human being like her chose to complete me. Thinking this way makes me try to even stand her frequent *kentuts*, far more potent than my socks... I give my warmest thanks to Dwi Hesti Yuliani with love and respect. *Teri makasi, sayangku tercinta. Aku cinta kamu.*

DEDICATION

To my families back in Niigata, Japan and Kendal, Indonesia,
with deepest appreciation, respect, and love

TABLE OF CONTENTS

List of Figures.....	14
List of Tables.....	15
List of Abbreviations.....	16
Abstract.....	17
 CHAPTER 1: INTRODUCTION.....	 19
Section 1: Minimalist Interfaces.....	19
Section 2: Overview of the Dissertation.....	27
 CHAPTER 2: SUCCESSIVE CYCLICITY AND PHASE THEORY AT THE SYNTAX- PHONOLOGY INTERFACE: “VOICES” FROM INDONESIAN AND JAVANESE.....	 38
Section 1: Introduction.....	38
Section 2: The Distribution of the Active Voice Morphology in Indonesian and Javanese.....	42
Section 2.1: Cole and Hermon’s (1998) Generalization in Malay/Indonesian.....	42
Section 2.2: The Distribution of the Nasal Prefix <i>ng-</i> in Javanese.....	53
Section 2.3: A Brief Excursus on Madurese.....	56
Section 3: Successive Cyclicity and the Role of <i>vP</i> Phases at the Syntax-Phonology Interface.....	60
Section 3.1: Kayne’s (1989) Analysis of Participle Agreement in French.....	61
Section 3.2: Phase Theory.....	66
Section 3.3: A Phase-Theoretic Analysis of the Active Voice Deletion.....	73
Section 3.4: <i>meN-/ng-</i> Deletion as Failure of Vocabulary Insertion.....	86
Section 3.5: Section Summary.....	88
Section 4: Other Alternative Accounts.....	89
Section 4.1: NP Accessibility and the Subject-Only Restriction.....	89
Section 4.1.1: Indonesian.....	92
Section 4.1.2: Javanese.....	98
Section 4.2: Soh’s (1998) Relativized Minimality Account.....	100
Section 4.3: Voskuil’s (2000) <i>Pro</i> -Based Account.....	104
Section 4.4: Fortin’s (In press) Anti-Passive Account.....	109
Section 4.5: Anti-Agreement-Based Accounts.....	117
Section 4.6: Case-Based Accounts.....	119
Section 5: Conclusions.....	123
Section 5.1: The Derivational Nature of the Syntax-Phonology Interface.....	124
Section 5.2: Minimalist Interfaces.....	125

TABLE OF CONTENTS — *Continued*

CHAPTER 3: PREPOSITION STRANDING UNDER SLUICING AND REPAIR	
BY DELETION: WHY INDONESIAN IS (NOT) SO SPECIAL?.....	129
Section 1: Introduction.....	129
Section 2: Sluicing Constructions.....	136
Section 2.1: Merchant's (2001) Analysis of Sluicing and the P-Stranding Generalization..	137
Section 2.2: P-Stranding under Sluicing in languages without P-Stranding.....	141
Section 2.2.1: Brazilian Portuguese.....	142
Section 2.2.2: Polish.....	145
Section 2.2.3: Mandarin Chinese.....	150
Section 2.2.4: Malagasy.....	153
Section 2.2.5: Serbo-Croatian.....	156
Section 2.2.6: "Noises" in Italian, Greek, and Polish.....	163
Section 2.3: Section Summary.....	163
Section 3: The Syntax of Sluicing in Indonesian.....	164
Section 3.1: Indonesian Sluicing = Pseudosluicing?.....	165
Section 3.2: Cheng (1991): A Reduced Cleft Analysis.....	180
Section 3.3: Cole et al. (To appear): A Headless Relative Clause Analysis.....	183
Section 3.4: Sluicing in Indonesian ≠ Pseudosluicing: Why Indonesian is So Special.....	189
Section 4: Repair by Deletion at the Syntax-Phonology Interface.....	190
Section 4.1: Feature Percolation, Minimality, D-to-P Incorporation, and Interface Repair.....	191
Section 4.1.1: Whether the [+wh] Feature Percolates or Not.....	192
Section 4.1.2: D-P Coalescence as D-to-P Incorporation.....	196
Section 4.1.3: Repair of "Imperfections" at the Syntax-Phonology Interface.....	202
Section 4.2: Towards an Etiology of P-Stranding Violations across Languages.....	207
Section 4.2.1: English.....	209
Section 4.2.2: Indonesian.....	210
Section 4.2.3: French.....	216
Section 4.3: Pseudogapping in Indonesian and French.....	219
Section 5: Comparison of the Proposed Analysis with Fortin's (2007) Analysis.....	223
Section 5.1: Fortin's (2007) LF Copy + Long Distance Agree Analysis.....	223
Section 5.2: Four Problems with Fortin's (2007) Analysis.....	227
Section 5.3: Derivationalism vs. Representationalism.....	229
Section 6: Conclusions.....	231
Section 6.1: Chapter Summary.....	231
Section 6.2: Implications of Indonesian Sluicing for the Theory of the Syntax-Phonology Interface.....	233

TABLE OF CONTENTS — *Continued*

CHAPTER 4: THE MORPHOSYNTAX OF BARE NOMINALS IN INDONESIAN AND JAVANESE: A RELATIVIZED PARAMETRIC THEORY OF NOMINAL DENOTATION.....	235
Section 1: Introduction.....	235
Section 2: Chierchia's (1998a, b) Nominal Mapping Parameter.....	240
Section 3: The Denotation and Morphosyntax of Bare Nominals in Indonesian and Javanese.....	243
Section 3.1: Bare Nominals in Indonesian.....	244
Section 3.2: Bare Nominals in Javanese.....	258
Section 4: A Relativized Parametric Theory of Nominal Denotation: From Indonesia to the World.	265
Section 4.1: A Relativized Parametric Theory of Nominal Denotation and Morphosyntax.....	266
Section 4.2: Deriving the Denotation and Morphosyntax of Bare Nominals across Languages.....	284
Section 5: New Typological Predictions of the Proposed Analysis.....	311
Section 5.1: Child English Morphosyntax.....	311
Section 5.2: Slavic Morphosyntax.....	317
Section 5.3: Chinese Morphosyntax.....	319
Section 5.4: Other Languages.....	322
Section 6: Chapter Summary: Theoretical Implications for the Syntax-Semantics Interface.....	323
 CHAPTER 5: WHETHER WH-IN-SITU MOVES OR NOT IN INDONESIAN: CHOICE FUNCTION AND INTERFACE ECONOMY.....	333
Section 1: Introduction.....	333
Section 2: <i>Wh</i> -in-Situ in Indonesian.....	336
Section 2.1: "Overt" Syntactic Movement?.....	338
Section 2.2: Covert Syntactic Movement?.....	343
Section 2.3: Unselective Binding?.....	356
Section 2.4: Section Summary.....	363
Section 3: <i>Wh</i> -in-Situ in Indonesian is Not an Interrogative Definite Description.....	365
Section 3.1: Saddy's (1991) Analysis of <i>Wh</i> -in-Situ as Interrogative Definite Description.....	365
Section 3.2: Problems with Saddy's (1991) Analysis.....	370
Section 4: Choice Function as an Optimal Interface Strategy at the Syntax-Semantics Interface.....	375
Section 4.1: Choice functions.....	377
Section 4.2: Deriving the Properties of <i>Wh</i> -in-Situ in Indonesian.....	386
Section 4.3: New Predictions: The NP vs. non-NP Asymmetry.....	393
Section 4.4: Cole and Hermon's (1998) Unselective Binding vs. Choice Function.....	395
Section 5: Conclusions.....	402

TABLE OF CONTENTS — *Continued*

CHAPTER 6: REDUPLICATION ASYMMETRIES IN INDONESIAN AND THEIR THEORETICAL IMPLICATIONS FOR THE LEXICON-SYNTAX INTERFACE.....	411
Section 1: Introduction.....	411
Section 2: Lexicalist vs. Non-Lexicalist Theories.....	417
Section 2.1: Lexicalist Theories.....	417
Section 2.2: Non-Lexicalist Theories.....	427
Section 3: Asymmetries between Nominal and Verbal Reduplications in Indonesian...	432
Section 4: Reduplication Asymmetries in Indonesian and lexicalist Theories.....	437
Section 4.1: Chomsky's (1970) Weak Lexicalist Hypothesis.....	437
Section 4.2: Anderson's (1982, 1992) Weak Lexicalist Theory.....	447
Section 4.3: Kiparsky's (1982a, b, c, 1985)/Mohan's (1986) Lexical Phonology.....	449
Section 4.4: Di Sciullo and Williams' (1987)/Williams' (2007) Strong Lexicalist Theory...	454
Section 4.5: The "Lexicon" as the Source of the Ordering Paradox.....	457
Section 5: A Distributed Morphology Approach to Reduplication Asymmetries in Indonesian...	458
Section 5.1: Nominal Reduplication.....	459
Section 5.2: Verbal Reduplication.....	462
Section 6: Conclusions.....	471
Section 6.1: Implications for the Proper Theory of the Lexicon-Syntax Interface.....	474
Section 6.2: Minimalist Interfaces.....	476
 CHAPTER 7: MINIMALIST INTERFACES.....	478
Section 7.1: Summary of the Dissertation.....	478
Section 7.2: Minimalist Interfaces.....	480
Section 7.3: Questions for Future Research and Conjectures about Linguistic Interfaces.....	491
 REFERENCES.....	503

LIST OF FIGURES

Figure 1: The Architecture of the Interfaces under the Thesis of the Minimalist Interface...21

LIST OF TABLES

Table 1:	The Parametric-Theory of P-Stranding at the Syntax-Phonology Interface.....	131
Table 2:	Pseudosluicing Diagnostics in Indonesian.....	166
Table 3:	A Relativized Parametric Theory of Nominal Denotation (preliminary version).....	283
Table 4:	A Relativized Parametric Theory of Nominal Denotation (final version).....	325
Table 5:	The Corpus Survey of Four Popular Newspapers in Indonesia..... (approx.160, 000 words)	433

LIST OF ABBREVIATIONS

Abs	=	Absolutive Case	Hyp	=	Hypothetical
Acc	=	Accusative Case	Instr	=	Instrumental
Act	=	Actor Topic	Intr	=	Intransitive
Av	=	Active voice	Loc	=	Locative
C	=	Complementizer	Link	=	Linker
Circum	=	Circumstantial	Neg	=	Negation
Cl	=	Clitic	Nom	=	Nominative Case
Class	=	Classifier	Ov	=	Object voice
Conj	=	Conjunctive Particle	Part	=	Participal
Cop	=	Copula	Past	=	Past Tense
Dat	=	Dative	Perf	=	Perfective Aspect
Det	=	Determiner	Pl	=	Plural
Emp	=	Emphatic Particle	Prog	=	Progressive
Erg	=	Ergative Case	Pv	=	Passive voice
Ez	=	Ezafe particle	Q	=	Question morpheme
f	=	Feminine	Red	=	Reduplication
1	=	First Person	Rel	=	Relative clause
Foc	=	Focus	Sg	=	Singular
Fut	=	Future Tense	3	=	Third person
Gen	=	Genitive	Top	=	Topic Marker
Hab	=	Habitual	Vz	=	Verbalizer

ABSTRACT

This dissertation is a theoretical investigation of the thesis of *Minimalist Interfaces*, namely, that syntax-external linguistic interfaces that interact with the core syntactic computation and language-independent sound and concept systems play a more critical role in manipulating syntactic objects to make them legible to those systems than is currently assumed in the recent minimalist inquiry. The core theme of this thesis lies in the idea that syntax is not entirely crash-proof but could make a variety of derivational mistakes; phonological and semantic linguistic interfaces conduct a handful of independent domain-specific operations to attempt to legitimize illicit syntactic objects, if any, for the purposes of legibility at the language-external sound and concept systems. Evidence is provided that the syntax-external components use whatever resources they can to repair certain “imperfections” created by syntax but only within the range of options made available by the universal principles of syntax in tandem with the language-specific parameter values. This dissertation explores some of the ramifications and empirical consequences of this thesis based on the comprehensive description of a sizable portion of the grammar of Indonesian and Javanese collected by my fieldwork with three native Indonesian and Javanese consultants. Phenomena discussed here include the distribution of active voice morphology, P-stranding under sluicing, the denotation and morphosyntax of bare

nominals, *wh-in-situ* questions, and reduplication asymmetries between nominal and verbal derivational affixes. These diverse ranges of phenomena in the two languages are analyzed in depth to provide converging evidence that the thesis of minimalist interface as defined above yields a deep understanding of the way the syntax interacts with the language-dependent interfaces responsible for phonological and semantic interpretation. The investigation conducted here, necessitates serious reconsideration of the commonly held view of linguistic interfaces as passive, merely ornamental components of natural language grammar ruled by the universal law of syntax.

CHAPTER 1 INTRODUCTION

1. Minimalist Interfaces

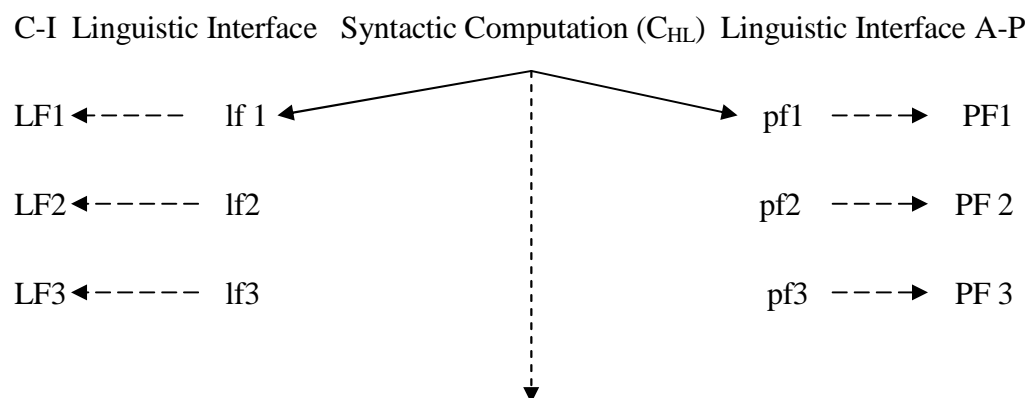
This dissertation is a theoretical investigation of a thesis I call *Minimalist Interfaces*. The goal is to see the extent to which this thesis serves as an adequate hypothesis of the correspondence between syntactic computation and the linguistic interfaces responsible for phonological and semantic interpretation.¹ The fundamental idea behind this hypothesis is that syntax-external linguistic components play a critical role in applying a handful of domain-specific operations to legitimize otherwise illicit objects created by the universal combinatorial process of Merge so that they become legible for the purposes of actual use in the language-independent articulatory-perceptual/AP and conceptual-intentional/CI systems (Chomsky 1995, 2000, 2001, 2004, 2005). The core intuition behind this thesis is that syntax is not entirely crash-proof, as argued for in Frampton and Guttman (1999, 2002), in that it makes certain derivational mistakes but syntax-external linguistic interfaces make use of whatever resources they can to attempt to make them legible for the purposes of the external phonetic and conceptual systems. This thesis also establishes that only *certain* mistakes

¹ I am grateful to Andrew Carnie (personal communication) for originally suggesting the keyword *minimalist interfaces* back in October 2007 as one of the potential overarching hypotheses that characterizes this whole dissertation. Thanks also go to Noam Chomsky (personal communication) for encouraging me to take seriously the role of linguistic interfaces and their relation with the language-independent concept and sound modules.

committed by syntax can be repaired at the interfaces. This idea thus leads to the claim that syntax-external components that interface both with syntax and the AP/CI systems can make use of their autonomous independent operations distinct from syntax to legitimize certain syntactically illegitimate objects but only within the range of options made available by the combination of the universal law of syntax with the language-particular parameter values.

When combined with the recent derivational theory of syntax and its correspondence with phonology and semantics (see Chomsky 1995, 2000, 2001, 2004, 2005, Epstein et al. 1998, Uriagereka 1999, and Grohmann 2003), the thesis of minimalist interface defined above yields the following architecture of the interface between syntax, syntax-external interfaces, and language-independent AP and CI systems. The small *lf* and *pf* should read as the chunk of syntactic structure that is transferred to syntax-external phonological and semantic interfaces. The big *LF* and *PF* should read as whatever representation that the C-I and A-P systems create based on the corresponding *lf* and *pf*.

Figure 1. The Architecture of the Interfaces under the Thesis of the Minimalist Interface



As noted above, this mode of dynamic interpretation of the relation between syntax and phonology/semantics is itself not a new idea but has been pursued in several different directions within the framework of the Minimalist Program. Where they differ is primarily in the exact size of material that undergoes transfer to the syntax-external interpretive components. The critical feature that distinguishes the derivational model of the minimalist interface proposed here from other instantiations lies in the idea that phonological and semantic interfaces are not merely interpretive but also actively participate in the linguistic computation as broadly conceived, in the sense that they use a handful of domain-specific operations such as deletion or choice functions to repair certain imperfections created by the functionally blind derivational process in the syntactic component. This idea, thus, meets the leading methodological minimalist desideratum of explaining linguistic

phenomena in solely in terms of interface conditions. Under the present thesis, then, syntax cannot be crash-proof in the sense of Frampton and Guttman (1999, 2002); the objects sent to the linguistic interfaces from syntax may or may not be legitimate by themselves, contrary to the minimalist assumption, as first expressed in Chomsky (1995), that syntax does not take any unnecessary or illegitimate derivational steps.

One primary goal of this dissertation, therefore, is to establish that linguistic interfaces do whatever operations they can to attempt to create legitimate objects out of illegitimate syntactic objects, if any, and send them off to the language-independent A-P and C-I systems in a way that they are usable for these systems, as shown in Figure 1. In other words, syntax creates objects that may or may not converge from the perspective of the AP and CI systems. Specifically, when syntax creates objects as its output that do not involve any syntactic violation and spells them out, the linguistic interfaces add minimal modifications to convert the syntactic object into a representation that is legible to the AP and CI systems. I show in chapters 2 and 4 that this scenario is instantiated in the proposed analysis of the distribution of active voice morphology and the denotation and morphosyntax of bare nouns in Indonesian and Javanese. However, when syntax creates objects that do involve certain syntactic violations, the linguistic interfaces conduct domain-specific operations such as deletion or choice functions to make them legible for the AP and CI systems. This does not

mean that this type of interface strategy is omnipotent, in the sense that all kinds of syntactic violation can be repaired at the interfaces so that the linguistic computation will always be able to create *lfs* and *pfs* that satisfy the demands of the AP and CI systems. For example, certain imperfections such as the failure of feature percolation and the lack of the diversity of communicative intent expressed by *wh*-questions can be repaired/supplemented by deletion and choice functions but other violations such as the failure of D-to-P incorporation or the category-sensitive in-situ option for *wh*-questions cannot. It is in this type of case that the minimalist interface thesis allows us to gain a deep understanding of the way syntax interacts with its neighboring linguistic interfaces and the language-external AP and CI systems. I show in chapters 3 and 5 that this scenario is instantiated in the proposed analysis of the P-stranding pattern and in-situ *wh*-questions in Indonesian.

There are several other aspects of the present thesis that bear emphasizing. First, the present thesis indicates that syntax is functionally blind; it does whatever its abstract computational processes such as Internal/External Merge, Agree, Spell-Out, Transfer allow it to do to construct complex objects in a recursive fashion based on a language-particular subset of the universal set of morphosyntactic features (such as T, *v*, V, C, etc.), without ever caring about the fate of the objects thus created, leaving the task of their interpretability/convergence entirely to the external sound-and meaning-related linguistic

components. Therefore, it is natural to expect that syntax creates certain objects that would be simply unusable from the perspective of the A-P and C-I systems. This view of interface-driven interpretability is a reasonable one in light of another consideration that what actually interfaces with the language-independent A-P and C-I systems is not syntax *per se* but the intermediate components that connect syntax and the systems, as shown in Figure 1. This view has also been argued for in recent work as in Boeckx (2007), who proposes to let linguistic interfaces determine the legitimacy of syntactic objects. This view, therefore, naturally leads us to the proposal made above, namely, that linguistic interfaces are equipped with domain-specific operations to legitimize syntactic objects to make them readable for the A-P and C-I modules. Second, the proposed architecture of interface summarized in Figure 1 makes it clear that there is no room for the *lexicon* as traditionally conceived of as a storage point for words and their formation processes; the traditional conception of the lexicon does not find its natural place under the most parsimonious version of the minimalist view of the linguistic computation, adopted in this dissertation, that what syntax interacts with is the sound and meaning component. Accordingly, the thesis of minimalist interface leads us to expect that (part of) the traditional roles of the lexicon in the lexicalist sense be played by the post-syntactic linguistic interfaces by such means as late insertion of phonological material, as recently argued for in Distributed Morphology (Halle and Marantz

1993, 1994; Harley and Noyer 1999; Embick and Noyer 2007). I show in chapter 6 that this minimalist view provides a natural account of certain asymmetries between nominal and verbal reduplication and the *ber*-prefixation pattern in Indonesian that would pose empirical/architectural problems for various versions of the lexicalist theory. Third, given that it is linguistic interfaces that directly “communicate” with the A-P and C-I systems, it is possible to think that the nature of domain-specific operations conducted at those interfaces are influenced by properties characteristic of AP and CI systems such as non-parsing/dephoneticization, logic, and set-formation. It is shown in this dissertation that this line of thinking yields a new analysis of the crosslinguistically peculiar P-stranding pattern under sluicing via deletion and of *wh*-in-situ in Indonesian via choice function.

The present dissertation discusses some of the empirical ramifications of the minimalist interface defined above on the basis of the comprehensive description and theoretical analysis of a sizable portion of the syntax, semantics, phonology and morphology of two under-represented Malayo-Polynesian languages from the Austronesian family, Indonesian/Bahasa Indonesia and Javanese. I examine a wide variety of areas in these languages where syntax interacts with phonology, morphology, and semantics. They include, but are not limited to, the interaction of syntactic movement with the distribution of active voice morphology, the crosslinguistically atypical P-stranding pattern under sluicing, the

denotation and morphosyntax of bare nominals, in-situ *wh*-questions, and nominal vs. verbal reduplication asymmetries. These apparently disparate phenomena in these two languages are analyzed in great depth to provide converging evidence for the single idea that syntax-external linguistic interfaces make use of whatever domain-specific resources they can to modify/remedy certain mistakes created by syntax, if any, but only within the logical space set up by the combination of the architectural design of syntactic computation and the language-particular values of independently motivated parameters. To mention one case, the analysis of the typology of P-stranding proposed in chapter 3 draws on two parameters concerning the percolation of the [+wh] feature of the *wh*-word onto the PP (Chomsky 1972; Stepanov 2001) and concerning the D-to-P incorporation in syntax (Law 1998, 2006; van Riemsdijk 1998) to account for the three-way contrast with respect to P-stranding under *wh*-questions and sluicing among English, French, and Indonesian. It is proposed that the phonological component can repair the failure of the [+wh] feature percolation by deleting the offending PP structure, but not the failure of the D-to-P incorporation, thereby providing an explanation for why possibilities of repairs at the phonological interface are restricted by language-particular parameter values.

Although the subtitle of this dissertation gives the impression that the database of our inquiry is limited to Indonesian and Javanese, this dissertation also contains a wealth of

examples and descriptions from a far wider range of genetically unrelated languages encompassing Indo-European, Austronesian, Altaic, and so on, as long as their investigation bears on the question of the extent to which the thesis of minimalist interfaces holds.

2. Overview of the Dissertation

Beyond the present introductory chapter, this dissertation is composed of 6 chapters. Chapters 2 and 3 concern the interface of syntax with phonology, Chapters 4 and 5 the interface of syntax with semantics, and Chapter 6 the interface of syntax with morphology/lexicon. Below is the concise summary of the content of each successive chapter, with particular emphasis on the relevance of the phenomena discussed therein to the thesis of minimalist interface.

Chapter 2 is a theoretical exploration of the interface between syntax and its neighboring phonological component within the derivational theory of syntax known as *Phase Theory* (Chomsky 2000, 2001, 2004, 2005, 2006) with a case study in the voice-movement interaction in Indonesian and Javanese. I start by reviewing the descriptive generalization, first made by Cole and Hermon (1998), that, in Malay/Indonesian, the movement of an NP across the active voice marker *meN-* results in the obligatory deletion of the active voice marker, and by

showing that this generalization also holds for Javanese in terms of the deletion of the nasal active voice prefix. Drawing on the core idea behind Kayne's (1989) analysis of participial agreement in French, namely, that syntactic movement of an argument affects the form of a verb within its extraction path, I propose that the obligatory deletion of the active voice morphology in Indonesian and Javanese is the reflex at the syntax-external phonological component of the Spec-Head D-feature checking relation that holds between the moved NP and its local *v* head at the *v*P phase. To the extent that this analysis is tenable, the current investigation provides important evidence for the role of the *v*P phase at the syntax-phonology interface. Though numerous types of evidence have been accumulated in recent generative research for successive cyclic movement through intermediate CPs, evidence supporting the comparable movement through intermediate *v*Ps has proven difficult to come by. In this regard, the contribution from under-represented languages such as Indonesian and Javanese is highly significant. I conclude this chapter by making it clear how the results in this chapter support the idea that the way the phonological component conducts the deletion of active voice morphology in Indonesian tells us a lot about the way syntactic derivation proceeds in tandem with its neighboring phonological interface. This idea received further support from examination of data from other languages like Irish, Kikuyu, Berber, Irish, Italian, and Turkish, as these languages employ other interface strategies such as the complementizer alternation,

the loss of tonal downstep, and the anti-agreement that serve to diagnose the phase-based local computation. This observation, then, leads to the conclusion, expected under the minimalist interface thesis, that the phonological interface is endowed with a handful of domain-specific operations, but can apply them only within the range of options set by principles of syntax.

Chapter 3 continues the exploration of the interface between syntax and the phonological component through detailed investigation of the P-stranding pattern and sluicing in Indonesian. In favor of his movement + TP deletion approach to sluicing constructions, Merchant (2001) establishes the so-called *P-Stranding Generalization* that P-stranding is permissible under sluicing only in those languages that independently allow this option under regular *wh*-movement. I start by reviewing the P-stranding pattern in several languages, such as Brazilian Portuguese, Mandarin Chinese, Polish, Serbo-Croatian, and Malagasy, that have been recently reported to contradict this generalization. I show, discussing recent work on sluicing and P-stranding in these languages, that none of these languages poses genuine counterevidence to the P-Stranding Generalization because the underlying structure of sluicing before deletion involves some other syntactic structure (cleft, pseudocleft, resumption) or phonological operation (P-omission at PF) than *wh*-movement of the English kind in these languages. One primary goal of this chapter is to establish that Indonesian is the first genuine counterexample to the relevant generalization

as a non-P-stranding language that nonetheless allows P-stranding under sluicing. I provide evidence based on the distribution of the question marker *-kah*, first discovered by Fortin (2007), and the obligatory lack of the complementizer *yang* with *wh*-questions with PPs that at least PP sluices in this language are based on regular *wh*-movement, rejecting alternative cleft (Cheng 1991) and headless relative clause analyses (Cole et al. to appear). This result, thus, splits languages into three types depending on the availability of P-stranding under *wh*-questions and sluicing: a) a handful of languages such as English, Danish, and Scandinavian languages that allow P-stranding under both contexts, many languages such as French, Spanish, and Italian that disallow P-stranding under either context, and languages such as Indonesian that allow P-stranding only under sluicing. The second goal of this chapter is, therefore, to provide a novel analysis for this three-way contrast across languages. Specifically, I propose that the relevant contrast receives a straightforward parametric explanation by introducing the following two parameters: a) whether the [+wh] feature CAN or MUST percolate from the nominal complement of a preposition onto the dominating PP, and b) whether D-to-P incorporation is attested in the language or not. The most important idea in the proposed analysis is that of *interface repair*, namely, the notion introduced above that syntax is allowed to make certain mistakes that can be repaired at the phonological interface. Details of the proposed analysis

show that certain derivational mistakes such as percolation failure, but not other mistakes such as failure to incorporate into P heads, can be undone by deletion at the syntax-external phonological interface. This result, therefore, provides substantial empirical support for the core idea of the minimalist interface thesis that syntax is not entirely crash-proof and that syntax-external linguistic interfaces apply domain-specific mechanisms such as deletion to make them legible and usable to the language-independent A-P and C-I systems. The same result also indicates, however, that syntax still has control over the fate of potentially illegitimate objects it sent out to its neighboring interfaces in that certain mistakes deeply rooted in syntax cannot be repaired by interface operations. Finally, I also provide one analysis regarding the question of why certain syntactic mistakes such as failure of the [+wh] feature percolation can be repaired by deletion, but not other mistakes such as the failure of D-to-P incorporation cannot be so repaired, by examining the input and output structures of the PF component. This analysis sheds a new light of the form and function of the syntax-external but linguistic-internal phonological component.

Chapter 4 turns to the investigation of the interface between syntax and its other neighboring linguistic component, semantics, with a case study in the denotation and morphosyntax of bare nominals in Indonesian and Javanese. Chierchia (1998a, b) recently proposed a restrictive parameter of the denotation and morphosyntax of bare nominals called the *Nominal Mapping*

Parameter, which states that there is a semantic parameter concerning whether a particular language allows its bare nominals to denote the name of a kind, the name of a property, or both in the mapping between syntax and semantics. One attractive feature of this hypothesis lies in the idea that the setting of this parameter exhaustively determines the morphosyntactic profile of nominals in a given language and that all languages should be characterized as falling within one of the three language types. My starting point in this chapter is to show, following Chung (2000), that Indonesian and Javanese do not fit into any one of these three languages under Chierchia's semantic typology by demonstrating that no combination of parameter values in Chierchia's system would accommodate the full range of morphosyntactic properties of bare nominals in these languages. This is an important result since Chierchia's (1998a, b) discussion concentrates on detailed comparison of relatively well-studied languages such as English, Italian, and Chinese. Following the standard assumption within the Principles-&-Parameters approach to linguistic variation (Borer 1984; Fukui 1986, 1995; Chomsky 1986a, 1995) that the setting of the values of a parameter is localized in the inventory of functional categories, I propose a relativized parametric theory of the denotation of bare nominals that draws on two independently motivated ideas: a) how high a language can allow its nominal functional superstructure to grow (Grimshaw 1991, 2005; Massam 2001; Guilfoyle and Noonan 1992; Vainikka 1993/1994) and b) what possible values the Num head can take in a given language

(Carson 2000; Chung 2000). I show that this analysis provides a principled explanation for the crosslinguistically variant semantic and morphosyntactic profile of bare nominals in not only in Indonesian and Javanese but also in other languages such as English, Italian, Japanese, Chinese, and Russian. I also argue that the proposed analysis makes correct predictions concerning the development/maturation of child language grammar with particular attention to the order of acquisition of functional categories in English-learning children documented by Guilfoyle and Noonan (1992), Fromkin et al. (2003), and Miller and Ervin-Tripp (1973). When combined with the Structure Building Hypothesis (Guilfoyle and Noonan 1992) and the Subset Principle (Wexler and Manzini 1987), the proposed analysis predicts that initial acquisitional stages of all languages should be like Javanese and Indonesian in the nominal denotation and morphosyntax of nominals because the latter represent the simplex nominal structure. I show that this prediction is indeed verified by utterances that are produced in the so-called *telegraphic stage*. To the extent that the current analysis holds, we have no need to make recourse to rigid one-to-one mapping principles as in Chierchia's Nominal Mapping Parameter; its effects directly follow from the complexity of nominal functional structures and the set of possible values for the Num head parametrically set in each language. This result, therefore, provides substance to the idea encoded in the minimalist interface, namely, that the syntax-external semantic

component can apply a restricted range of domain-specific operations to the output of syntax but only within the realm of syntactic derivation.

Chapter 5 continues the exploration of the syntax-semantics interface with detailed investigation of the syntax and semantics of *wh*-in-situ in Indonesian. In his seminal work on *wh*-questions in Indonesian, Saddy (1991) provides extensive arguments that there is something different about *wh*-in-situ in this language, whose interpretive and syntactic properties are quite unpredictable based on the recent study of corresponding *wh*-in-situ constructions in other Asian languages such as Japanese and Chinese. I review his major arguments to show why the two most popular analyses of *wh*-in-situ in terms of syntactic movement (Watanabe 1992, 2001; Huang 1982; Richards 2001) and unselective binding (Pesetsky 1987) fail in the face of *wh*-in-situ in Indonesian. At the same time, I make it clear that some of the crucial judgments reported by Saddy to establish this result could not be reproduced in my field work with native consultants and contradict the results reported in recent work on Indonesian and Malay in Cheng (1991), Cole and Hermon (1998, 2000), and others. Based on this result, I seek a new analysis of *wh*-in-situ informed by the guiding thesis of minimalist interface, which leads us to the claim that the very fact that a *wh*-phrase remains in situ has consequences for the semantic interface; the mechanism to license *wh*-in-situ does not lie in syntax *per se* but

rather in the semantic interface that connects it to the language-independent C-I system. Specifically, I propose that the relevant mechanism is a choice function in the sense of Reinhart (1992, 1995, 1997, 1998, 2006). I demonstrate that all apparently peculiar syntactic and semantic characteristics associated with *wh*-in-situ in Indonesian directly follow from this single mechanism at work at the linguistic interface. Finally, I compare the present analysis with the most recent version of the unselective binding approach to *wh*-in-situ in Malay/Indonesian proposed by Cole and Hermon (1998, 2000). I show that the latter has several shortcomings in face of the so-called “Donald Duck Problem” (Reinhart 1992) and the presence of the intermediate reading in long-distance *wh*-questions (Ruys 1992; Reinhart 1998, 2006) that is successfully resolved under the former from the very definition of choice function without any ad hoc stipulations on the syntax-semantics mapping. The proposed analysis, if correct, provides evidence that the semantic interface also is endowed with a handful of modular-specific operations such as choice functions to legitimize incomplete syntactic objects and send them off to the C-I system in a legible manner. I also briefly address the important question of why the semantics interface develops this particular semantic operation. I speculate that this state of affairs is naturally expected because what actually interfaces with the conceptual system is the linguistic interface, not syntax, and it is reasonable to expect that the

linguistic semantic interface “borrows” some mathematical properties characteristic of the conceptual system such as set formation, predicate logic, and so on.

Chapter 6 turns to the interface of the syntax with the morphology to provide part of the answer from Indonesian and Javanese facts to the hotly debated issue of the division of labor between syntactic computation and the so-called *lexicon* as a storage of words and their formation processes, as commonly conceived of in the generative enterprise. The empirical domain on which I base my investigation is a new observation that there is a curious asymmetry between nominal and verbal reduplication in Indonesian. A corpus study of four popular newspapers published in Indonesia reveals that verbal derivational affixes have a strong tendency to feed only stem reduplication whereas nominal derivational affixes allow either stem reduplication or stem-affix reduplication. I show that this new observation is also confirmed by the data I elicited with one native Indonesian consultant. I show that the stem-internal reduplication pattern as well as the observed asymmetry pose empirical/architectural paradoxes for several well-known variants of the so-called lexicalist theory as in Chomsky (1970), Anderson (1982, 1992), Kiparsky (1982a, b, c: 1985), Monahan (1986), and Di Sciullo and Williams (1987). Since the debate between lexicalist and non-lexicalist approaches to word formation has quite a long history sometimes coupled with heated rhetoric, I make it clear what specific aspects of these variants of the lexicalist theory are not tenable with respect

to the facts in Indonesian reduplication. Based on this result, I propose a morphosyntactic analysis of Indonesian reduplication within the uni-modular syntactic approach to word formation as in Distributed Morphology (Halle and Marantz 1993, 1994; Marantz 1997; Harley and Noyer 1999; Embick and Noyer 2007). I show that the observed facts receive a straightforward account under the post-syntactic bottom-up cyclic insertion of phonological features once we take seriously a hierarchical arrangement of morphosyntactic features and the underlying syntactic category of input stems for reduplication. The architecture of the syntax-morphology interface that emerges from this investigation is one where there is in fact no such interface in the strict sense, because the current analysis indicates that morphological structure is itself syntactic structure unless otherwise motivated (Harley and Noyer 1999; Embick and Noyer 2007). This conclusion, I argue, is optimal under the most restrictive view of the minimalist interface guideline that language must minimally interface with the A-P and C-I systems, but not with any other language-internal modules such as lexicon, unless empirical evidence suggests otherwise.

Chapter 7 summarizes the high points of the contents of my investigation in the previous chapters and draws some conclusions regarding the architecture of the faculty of human language, the nature of syntactic computation and its networking with language-independent sensory-motor and articulatory-perceptual systems.

CHAPTER 2 SUCCESSIVE CYCLICITY AND PHASE THEORY AT THE SYNTAX-PHONOLOGY INTERFACE: “VOICES” FROM INDONESIAN AND JAVANESE ¹

1. Introduction

This chapter examines the distribution of the active voice morphology in Indonesian and Javanese from the perspective of the syntax-phonology interface within the Minimalist Program. One of the central theses of the theory of generative grammar since its inception is the notion of *successive cyclicity*, namely, that syntactic movement occurs in a series of local steps on its way to the final landing site. Though theoretical implementations of this thesis have undergone several theoretical changes, e.g., *Subjacency* (Chomsky 1973), *Barriers* (Chomsky 1986b), *Shortest Move* (Chomsky 1993; 1995:ch. 3), and the *Minimal Link Condition* (Chomsky 1993; 1995: ch. 2), it has been generally considered as a centrally important discovery of generative investigation. This idea has been pushed forward further by the most recent derivational theory of syntax known as *Phase Theory* outlined in a series of work by Chomsky (Chomsky 2000, 2001, 2004, 2005), which takes it that syntactic derivation proceeds in tandem with semantic and phonological interpretation in a series of sub-chunks.

¹ Earlier versions of this chapter were presented at the 2007 Western Conference on Linguistics (WECOL 2007) held in San Diego (December 2007) and the 82nd Annual Meeting of the Linguistic Society of America (LSA) held in Chicago (January 2008). An abridged version of this chapter is to appear in Sato (in press a, b).

More specifically, Chomsky proposes that, once syntactic derivation constructs the ν P and CP structures headed by phase heads ν and C, their respective complements (VP and CP) undergo Spell-Out/Transfer to the syntax-external interpretive systems responsible for phonological and semantic interpretation in mid-derivation, contrary to the traditional T-model of syntactic derivation in which only the final output of the derivation is Spelled-Out once, as assumed in earlier architectures of the generative grammar such as Government-&-Binding Theory and so-called *Classical Minimalism* (Chomsky 1993: 1995: ch.2). In this strictly derivational model of syntax, elements such as direct object *wh*-phrases that are ultimately to undergo movement to the specifier of CP must first be moved into the specifier/edge of ν P, for otherwise they would undergo Transfer when the ν P was completed and be inaccessible to syntactic operations at the higher CP phase level. In this chapter, I discuss the syntactically governed distribution of the active voice marker in Indonesian and Javanese. I show that it provides strong support for the phase-based notion of successive cyclicity, in particular, for the role of the ν P phase from the perspective of the syntax-phonology interface.

It has been widely acknowledged in the literature on Indonesian as in Chung (1976), Saddy (1991), and Cole and Hermon (1998, 2000, 2005, in press), that, in Malay/Indonesian, the movement of an NP across a verb results in the deletion of the active voice morphology *meN*-

from the verb.² I show that this generalization, which I dub Cole and Hermon's Generalization, holds for Javanese as well. One question that has remained unresolved, however, is why this generalization holds for these two Javanic languages. I propose, drawing on Kayne's (1989) analysis of French participle agreement, that the obligatory deletion of the active voice morphology in these languages is the reflex at the PF interface of the Spec-Head D-feature checking relation between the moved NP and its local *v* head at the *v*P phase. Specifically, following the recent late-insertion theory of Distributed Morphology, I propose that the D-feature-driven movement of an NP across the *v* head causes a change in the feature content of the head, resulting in the failure of insertion at the post-syntactic phonological interface. The current analysis also correctly derives the fact that the movement of non-nominal phrases such as adjuncts and prepositional arguments does not cause obligatory active voice deletion, because these phrases lack D-features.

² I annotate the following allomorphs of the active voice marker in Indonesian collectively as *meN-* for simplicity's sake. (i) lists phonological conditions in which those various forms of *meN-* are realized.

(i) The Allomorphy of *meN-*

- a. *meN-* → *meng/menge* (if the stem starts with a, e, g, h, i, o, u)
- b. *meN-* → *mem* (if the stem starts with b, f, p, v)
- c. *meN-* → *men* (if the stem starts with c, d, j, t, z)
- d. *meN-* → *men/meny* (if the stem starts with s)
- e. *meN-* → *meny* (if the stem starts with k, l, m, n, v, w, y)

The proposed analysis of the distribution of the active voice morphology allows us to draw two conclusions for the phase theory of syntax and the syntax-phonology interface. First, as stated above, the phase-bound system of syntactic derivation predicts that syntactic movement occurs via the edge of CP and ν P until its final landing site. Although evidence abounds in the generative literature for the movement to the edge of CP, evidence has been hard to find for comparable movement via the edge of ν P; see Legate 2003 and the reply to Legate by den Dikken 2006). The syntactically governed distribution of the active voice morphology in Indonesian and Javanese is important in that it provides clear evidence from the syntax-phonology interface for the role of the ν P phase in the syntax. Second, the proposed analysis provides substance to the thesis of “minimalist interfaces” introduced in the introductory chapter. It shows that the syntax provides a parametrically defined curve that the syntax-external interpretive components must follow with whatever domain-specific resources they can avail themselves of to mirror the way syntax works as closely as possible. In other words, languages will do whatever they do within their available morphological and phonological resources to reflect the local step of movement via the edge of ν P and CP, as required by the phase-bound computation.

2. The Distribution of the Active Voice Morphology in Indonesian and Javanese

In this section, I provide a comprehensive description of the prefixes, *meN-* in Indonesian and *ng-* in Javanese. I concentrate on the distribution of these prefixes, which has been shown in recent research (Chung 1976; Saddy 1991; Cole and Hermon 1998, 2000, 2005, in press) to play a crucial role in diagnosing certain types of syntactic movement. I start in section 2.1 with the description of *meN-* in Indonesian and introduce Cole and Hermon's generalization that the movement of an NP cannot occur across an active verb unless the active voice *meN-* prefix is deleted. I show in section 2.2 that the same generalization also correctly governs the distribution of the nasal prefix *ng-* in Javanese. In section 2.3, I review Davies' (2003) evidence from Madurese, an Austronesian language spoken in the Madura Island in Indonesia, which shows that a similar generalization also holds for this language.

2.1. *Cole and Hermon's 1998 Generalization in Malay/Indonesian*

A large body of work has been produced in the literature on Indonesian on the exact nature of the prefix *meN-*. It is termed the "transitive/agentivity marker" by Chung (1976), the "agent trigger" by Englebretson (2003), the "[+active] feature" by Postman (2002), the "actor topic" by Guilfoyle et al. (1992), or the "active voice" by Sneddon (1996) and Voskuil (2000). For the purposes of this chapter, I assume that *meN-* is the marker of

active voice in Indonesian but is also affected by the syntactic transitivity of a verb that this affix is attached to. Evidence for this assumption comes from the following two facts. First, this prefix may appear in an active sentence but not in a passive sentence, as shown by the contrast between (1a) and (1b).

(1) a. Esti **mem**-baca/***di**-baca buku itu.

Esti **AV**-read/**PV**-read book that

‘Esti read that book.’

b. Buku itu ***mem**-baca/**di**-baca oleh Esti.

book that **AV**-read/**PV**-read by Esti.

‘That book was read by Esti.’

In (1a), an active sentence, the verb *baca* ‘read’ is prefixed by *meN*-, but not by *di*-, whereas the reverse situation holds in (1b), a passive counterpart to (1a). This contrast indicates that *meN*- signals active voice in Indonesian. Second, Fortin (in press) observes that, though *meN*- prefixation is optional with many transitive verbs in colloquial Indonesian, there is one case in which this prefixation is obligatory. The relevance case is

when a transitive verb is used in its intransitive use, as in English examples such as *John ate at noon*. This is illustrated in (2a, b).

- (2) a. Ali sedang mem-baca. b. * Ali sedang baca.
 Ali Prog AV-read Ali Prog read
 ‘Ali is reading.’ ‘Ali is reading.’ (Fortin in press)

In a normal transitive sentence as in (2a), *meN-* is optional. However, in its intransitive counterpart of (2a) as in (2b), this affix is obligatory. These two arguments, therefore, support my working assumption that *meN-* is an active voice marker whose realization is affected by syntactic transitivity (see Gil 2002 for a review of the historical development of this marker). To be more specific, since the diathesis here is connected to voice, as shown in (1a, b), and to the presence of overt objects, as shown in (2a, b), it is natural to associate it with *v*, which is the locus of external arguments (Hale and Keyser 1993, 2002; Chomsky 1995) and of accusative Case (Kratzer 1996).

As noted by recent work on the syntax of Indonesian as in Chung (1976), Saddy (1991), and Cole and Hermon (1998, 2000, 2005, in press), however, this characterization of *meN-* in terms of voice morphology and transitivity alone fails to explain certain interesting

restrictions on the obligatory deletion of the prefix caused by a particular type of movement. Cole and Hermon (1998) observe, drawing on data and observations made earlier by Chung and Saddy, that the distribution of *meN-* in Singapore Malay and Indonesian is governed by the generalization in (3). All the examples from Cole and Hermon (1998) are based on Malay but the same generalization they developed for Malay also holds for Indonesian; for example, they note that “the distribution of *meng-* in Indonesian...seems to be identical to what we have found in Singaporean Malay” (p. 230).

(3) 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.

(Cole and Hermon 1998: 233)

This generalization accommodates all existing cases of *meN-*deletion. Let us start with A'-movement. *Wh*-questions and relativization cause *meN-* deletion from all verbs that the moved NP/DP crosses on its extraction path toward its final landing site, as illustrated in (4a, b) and (5a, b).

(4) *Wh*-Questions

- a. Apa_i yang Bill (***mem**)-beritahu ibu-nya [yang Fatimah (***mem**)-beli *t_i*]?

what that Bill **AV**-tell mother-his that Fatimah **AV**-buy

‘What did Bill tell his mother that Fatimah bought?’

- b. Siapa_i yang Bill (***mem**)-beritahu ibu-nya [*t_i* (**mem**)-beli buku]?

who that Bill **AV**-tell mother-his **AV**-buy book

‘Who did Bill tell his mother bought a book?’ (Indonesian)

(Indonesian: modeled after the Malay examples from Cole and Hermon 1998: 231, 232)

(5) Relativization

- a. [Buku_i [OP_i yang [John (***mem**)-beli *t_i*]] itu] menarik.

book that John **AV**-buy the interesting

‘The book that John bought is interesting.’

- b. [Lelaki_i [OP_i yang [*t_i* (**mem**)-beli buku itu]]] adik saya.

man that **AV**-buy book that brother my

‘The man who bought that book is my brother.’

(Indonesian/Malay: slightly modified from Cole and Hermon 1998: 233)

In (4a), A'- movement occurs from the embedded object position to the matrix [Spec, CP]. The active voice prefix is obligatorily deleted from both the higher verb *beritahu* 'tell' and the lower verb *beli* 'buy' in conformity with Cole and Hermon's generalization. The same generalization also accounts for the mandatory deletion of *meN-* from the higher verb, but not from the lower verb, in (4b); *wh*-movement of the embedded subject *siapa* 'what' crosses only the matrix verb *beritahu* 'tell'. The same pattern holds for relativization, as shown in (5a, b). In (5a), operator movement of the object crosses the verb *beli* 'buy', causing the deletion of *meN-* from the verb. This deletion does not occur in (5b), because the null operator movement of the subject does not occur across the same verb.

Further support for Cole and Hermon's (1998) generalization comes from the movement options allowed in Malay/Indonesian and their effects on the active voice deletion. As first discovered by Saddy (1991), Indonesian has three ways of forming *wh*-questions: a) overt *wh*-movement, as illustrated in (4a, b) above, b) partial *wh*-movement, and c) *wh*-in-situ; see chapters 2 and 5 of the present dissertation for various analyses of Indonesian *wh*-questions. The three options are illustrated in (6a-c), respectively. (The examples in (6b, c) are from Cole and Hermon 1998: 233, 237.)

(6) Full *Wh*-Movement, Partial *Wh*-Movement, and *Wh*-in-situ in Indonesian/Malaya. Full *Wh*-Movement

apa_i Ali (***mem**)-beritahu kamu tadi [t_i yang Fatimah (***mem**)-baca t_i]?

what Ali **AV**-told you just now that Fatimah **AV**-read

‘What did Ali tell you just now that Fatimah was reading?’

b. Partial *Wh*-Movement

Ali (**mem**)-beritahu kamu tadi [apa_i yang Fatimah (***mem**)-baca t_i]?

Ali **AV**-told you just now what that Fatimah **AV**-read

‘What did Ali tell you just now that Fatimah was reading?’

c. *Wh*-in-situ

Ali (**mem**)-beri Fatimah apa?

Ali **AV**-give Fatimah what

‘What did Ali give Fatimah?’ (Indonesian/Malay)

In (6b), the movement of the embedded direct object *apa* ‘what’ targets the embedded, non-scopal [Spec, CP], though the interpretive outcome is the same as that of fully moved *wh*-questions as in (6a). The active voice deletion occurs from both the higher and lower verbs in (6a) whereas it occurs only from the lower verb in (6b). When the in-situ option is

utilized as in (6c), there is no *meN*-deletion observed from the verb *beri* ‘give’. All these patterns, thus, show that the form of a verb is affected by the options of *wh*-movement independently available in Indonesian.

Let us now turn to A-movement. As in A'-movement, passivization and object preposing in Indonesian causes active voice deletion from verbs that it crosses. Consider (7a, b) (= (1a, b)) and (8a, b).

(7) Passivization

- a. Esti (**mem**)-baca/*(**di**)-baca buku itu.

Esti **AV**-read/**PV**-read book that

‘Esti read that book.’

- b. Buku itu (***mem**)-baca/*(**di**)-baca oleh Esti.

book that **AV**-read/**PV**-read by Esti

‘That book was read by Esti.’

(8) Object Preposing

- a. Buku itu_i adik saya (***mem**)-beli *t_i*.

book that brother my **AV**-buy

‘My brother bought that book./That book was bought by my brother.’

- b. Ali_i saya (***men**)-cubit *t_i*.

Ali I **AV**-pinch

‘I pinched Ali./Ali was pinched by me.’

(Indonesian/Malay: Cole and Hermon 1998: 232)

In (7b), a passive counterpart to the active sentence in (7a), the active voice marker must be replaced by the passive voice marker *di-*. 8a, b) might appear to involve topicalization, a case of A'-movement. However, evidence based on the interaction of object preposition and Equi NP Deletion first presented in Chung (1976) and further expanded by Hopper (1983), Musgrave (2001), Aldridge (in press), and Cole and Hermon (2005, in press), suggests that object preposing in Indonesian is derived by A-movement of a logical object into [Spec, TP]. Consider (9a-d).³

³ The examples here are given here as they appear in Chung (1978: 46, 47) except slight modifications of the morphological glosses. The orthography used here reflects the one used prior to the spelling reform of 1972.

(9)a. Dia datang untuk ber-tjakap2 dengan Ali.

he come for Intr-talk with Ali

‘He came to talk with Ali.’

b.?* Saja mem-bawa surat itu untuk teman saja (dapat) (mem)-batja.

I AV-bring letter the for friend my can AV-read

‘I brought the letter for my friends to (be able to) read.’

c. Saja mem-bawa surat itu untuk (dapat) di-batja oleh teman saja.

I AV-bring letter the for can PV-read by friend my

‘I brought the letter to (be able to) be read by my friends.’

d. Saja mem-bawa surat itu untuk (dapat) kau batja.

I AV-bring letter the for can you read

‘I brought the letter to (be able to) be read by you.’ (Chung 1976: 46, 47)

Chung (1976) observes that Equi-NP deletion (subject control) targets only subject NPs, as the contrast between (9a) and (9b) shows. The example in (9c) shows that the derived subject can be PRO. With this background in mind, the fact that the preposed object in (9d) can be PRO indicates that Object Preposing involves movement of the logical object to [Spec, TP], a case of A-movement, like passivization. Therefore, the examples of object preposing as in (8a, b)

provide evidence that A-movement also affects the morphology of a verb that is contained within the extraction path of the movement.

Finally, as stated in Cole and Hermon's generalization in (3), what matters for the deletion of the active voice morphology is the movement of a nominal category. In other words, movement of non-nominal phrases such as adverbial adjuncts and prepositional arguments does not cause the active voice deletion from verbs it crosses, as in (10a, b).

(10) Movement of non-*Wh*-NPs

- a. Kenapa_i Mary (**mem**)-beli buku itu *t_i*?

why Mary **AV**-buy book that

'Why did Mary buy that book?'

- b. Kepada siapa_i Mary (**mem**)-beri buku *t_i*?

to who Mary **AV**-buy book

'To whom did Mary give a book?' (Cole and Hermon 1998: 231, 232)

In (10a, b), movement of phrases like *kenapa* 'why' and *kepada siapa* 'to whom' does not result in *meN*-deletion though the movement itself crosses the verb *beli* 'buy' / *beri* 'give' in both cases.

In sum, we have reviewed Cole and Hermon's (1998) generalization that the movement of an NP causes the obligatory deletion of the active voice marker *meN-* in Indonesian. One obvious question is, then, why is it that this generalization holds for Indonesian? Before I answer this question, I show in the next two subsections that Cole and Hermon's generalization also holds for at least other Javanic languages, Javanese and Madurese.

2.2. *The Distribution of the Nasal Prefix ng- in Javanese*

The distribution of the active voice morphology *ng-* in Javanese is also constrained by Cole and Hermon's generalization. The examples in (11a-d) illustrate the basic paradigm.

- (11) a. Basic Transitive Clause c. A'-movement (Relativization)
- Esti ***(nge)**-sun Fernando. Wong lanang_i sing Esti ***(nge)**-sun _{t_i} kuwi Fernando.
- Esti **AV**-kiss Fernando person male that Esti **AV**-kiss Cop Fernando
- 'Esti kissed Fernando.' 'The man that Esti kissed is Fernando.'
- b. A'-movement (*Wh*-Questions) d. A-movement (Passivization)
- Sapa_i Esti ***(nge)**-sun _{t_i}? Fernando_i ***(nge)**-sun (karo) Esti _{t_i}.
- who Esti **AV**-kiss Fernando **AV**-kiss by Esti
- 'Who did Esti kiss?' 'Fernando was kissed by Esti.'

In Javanese, the nasal prefix *ng-* is obligatorily attached to a transitive verb, as shown in (11a). This is in contrast with Indonesian *-meN*, which is optional, as we saw in the previous subsection. The examples in (11b, c) show that A'-movement of the objects in *wh*-questions and relativization leads to the obligatory deletion of the otherwise mandatory nasal prefix from the verb that it crosses. The example in (11d) further shows that A-movement of the object with passivization has the same morphological consequence for the form of the verb contained within the extraction path. Consider examples as in (12). This example is given here to show that long-distance extraction of an NP such as subject across a series of verbs also triggers *ng*-deletion from the verbs in the extraction path of the NP, though it does not trigger *ng*-deletion on the embedded verb, which the subject does not cross, exactly as in Malay/Indonesian.

(12) Long-Distance Extraction of Subject across Verbs in Javanese

[_{CP} Sapa_i sing John (***ng**)-ira [_{CP} *t*_i **nge**-sun Fernando]]?

who that John **AV**-think **AV**-kiss Fernando

‘Who does John think kissed Fernando?’

Thus, it is only the higher verb *ira* ‘think’ that has its nasal prefix obligatorily deleted. This example also shows that the morphology of a verb is affected by the movement of a phrase *across* it.

Finally, as in Indonesian, it is the movement of an NP that leads to the *ng*-deletion in Javanese. Thus, movement of non-nominal phrases such as adverbials (e.g., *nangapa* ‘why’) and prepositional arguments (e.g., *ning sapa* ‘to whom’) has no effect on the fate of the active voice marker in this language either, as shown in (13a, b).⁴

(13) Movement of Non-*Wh*-NPs

- | | |
|--|---|
| a. Nangapa _i Esti nge -sun Fernando <i>t_i</i> ? | b. Ning sapa _i Esti nge -irim packet <i>t_i</i> ? |
| why Esti AV -kiss Fernando | to whom Esti AV -send package |
| ‘Why did Esti kiss Fernando?’ | ‘To whom did Esti send a package?’ |

⁴ Heidi Harley (personal communication) asks whether Indonesian allows pseudopassivization like English, as in (i).

(i) This bed was slept in by Washington.

The prediction would be that the *meN*-deletion should occur in this construction because the movement of an NP crosses the (reanalyzed) verb. Unfortunately, this prediction is impossible to test for two reasons. First, Indonesian does not have pseudopassivization. Second, as we saw earlier, passivization introduces the passive suffix *di-*, which is in the complementary distribution with *meN-*. Similarly, P-stranding under *wh*-movement would be predicted to cause *meN*-deletion, but this prediction also cannot be tested since, as we will see in chapter 3, P-stranding is impossible in Indonesian.

The evidence presented here, therefore, shows that the Javanese voice-movement interaction is also governed by Cole and Hermon's generalization, originally established for Indonesian.

2.3. *A Brief Excursus on Madurese*

Evidence presented in Davies (2003) indicates that Madurese, a language spoken on the Madura Island of Indonesian, is also subject to the voice-movement alternation governed by Cole and Hermon's (1998) generalization. Consider the basic paradigm in Madurese, given in (14a-c).

(14) The Basic Voice System in Madurese

- a. Embi' juwa **ng**-ekke' Ali.

goat that **AV**-bite Ali

'The goat bit Ali.'

- b. Ali embi' juwa kekke'.

Ali goat that bite

'The goat bit Ali.'

c. Ali e-kekke' (bi') embi' juwa.

Ali OV-bite with goat that

'The goat bit Ali.'

(Davies 2003: 246)

In Madurese, the active voice morphology *ng-* appears on the verb when the agent or actor of a transitive sentence occurs as the subject, as shown in (14a). In passive examples as in (14b, c), the active voice marker must be eliminated; either the verb appears as a bare stem or the overt object voice prefix *e-* is attached to the verb. The appearance of the object voice marker in (14c) can be analyzed on a par with the voice-movement interaction observed in Tagalog and Malagasy (Rackowski and Richards 2005; Pearson 2001, 2005); see also section 3.4 for related discussion.

With this in mind, the contrast between (15a) and (15b) shows that the syntactic movement of a phrase across a verb results in the active voice deletion from that verb.

(15) *Wh*-Movement and the Voice-Movement Interaction in Madurese

a. Sapa_i se e-kera Ali (ja'/se) *t_i* melle motor?

who Rel OV-think Ali C/Rel AV.buy car

'Who did Ali think bought a car?'

b. *Sapa_i se **ngera** (ja'/se) t_i melle motor?

who Rel **AV.think** C/Rel **AV.buy** car

‘Who did Ali think bought a car?’

(Davies 2003: 246)

In these examples, the movement of *sapa* ‘who’ contains the higher verb within its extraction path, triggering the deletion of the active voice marker *ng-*, as in (15b), or its replacement with the object voice marker.

Further evidence that Cole and Hermon’s generalization is at work in Madurese comes from the correlation between the deletion of the active voice morphology and *wh*-movement options. As in Indonesian and Javanese, Madurese allows overt *wh*-movement, partial *wh*-movement, and *wh*-in-situ, as illustrated in (16a-d).

(16) *Wh*-questions in Madurese and the Voice-Movement Interaction in Madurese

a. Apa_i se **e-yaken-ne** Amir [_{CP} (ja') **e-bala-agi** Hasandha' Atin [**e-baca** Siti t_i]]?

what Rel **OV-sure-E** Amir C **OV-say-BV** Hasan to Atin **OV-read** Siti

‘What is Amir sure that Hasan told Atin that Siti read?’

- b. Amir yaken [_{CP} apa se e-bala-agi Hasan dha' Atin [_{CP} e-baca Siti *t_i*]]?
 Amir sure what Rel **OV**-read-BV Hasan to Atin **OV**-read Siti
 'What is Amir sure that Hasan told Atin that Siti read?'
- c. Amir yaken [_{CP} Hasan a-bala dha' Atin [_{CP} apa_i se e-baca Siti *t_i*]]?
 Amir sure Hasan **AV**-say to Atin what Rel **OV**-read Siti
 'What is Amir sure that Hasan told Atin that Siti read?'
- d. Amir yaken [_{CP} Hasan a-bala dha' Atin [_{CP} ja' Siti maca apa]]?
 Amir sure Hasan **AV**-say to Atin that Siti **AV**.read what
 'What is Amir sure that Hasan told Atin that Siti read?'

In the full *wh*-movement case shown in (16a), all the verbs must take object voice morphology. When the partial *wh*-movement targets the specifier of the higher intermediate CP as in (16b), the verbs that are contained within its extraction path all require object voice morphology. When this type of movement targets the specifier of the lower intermediate CP, as in (16c), only the most deeply embedded verb must be in object voice. Finally, in the case of the *wh*-in-situ as in (16d), all verbs can retain the active voice morphology. These examples, thus, clearly show that Cole and Hermon's (1998) generalization serves to adequately characterize the movement-voice interaction in Madurese. To complete the picture, examples as in (17) show that the movement of

non-nominal phrases such as prepositional objects does not result in the active-to-object voice change, as in Indonesian and Javanese.⁵

- (17) Dha' sapa Atin ng-erem paket?
 to whom Atin AV-send package
 'To whom did Atin send a package?'

Thus, based on the data available in Madurese, it seems that Cole and Hermon's generalization holds for Madurese. Due to the limited availability of the data from Madurese, however, I concentrate on facts from Indonesian and Javanese in the rest of this chapter, hoping that the analysis developed below will also hold for Madurese.

3. Successive Cyclicity and the Role of ν P Phases at the Syntax-Phonology Interface

We have seen in the previous section that Cole and Hermon's (1998) generalization adequately characterizes the voice-movement interaction in Indonesian and Javanese. As stated above, Cole and Hermon (1998) do not provide a theoretical explanation for why such a generalization holds for Javanic languages but instead merely speculates (p. 234)

⁵ Thanks to William Davies (personal communication) for providing me with the example in (17).

that “the treatment of *meng-* [*meN-* in this chapter—YS] is along the lines proposed by Chung for wh-agreement in Chamorro.”⁶ The purpose of this section is to show that this generalization provides strong morphosyntactic evidence for the role of *vP* phase at the syntax-phonology interface. Specifically, I propose that the obligatory deletion of active voice morphology from a verb that is crossed by the movement of an NP is a reflex in the syntax-external component of the Spec-Head D-Feature checking relation between the moved NP and its local *v* head at the *vP* phase level. This analysis also correctly captures the fact that non-nominal phrases such as adverbials and prepositional arguments does not cause deletion, in tandem with the independently motivated distinction between Internal Merge and External Merge, made recently by Chomsky (2004, 2005).

3.1. *Kayne’s (1989) Analysis of Participle Agreement in French*

The fundamental idea I pursue below to account for the voice-movement interaction in Indonesian and Javanese is part of the theory of “minimalist interface”, a central thesis developed through in this dissertation, namely, that syntax-external interpretive components do whatever they can to reflect the path that syntax carves within the parametrically defined set of

⁶ Cole and Hermon repeat a similar remark in Cole and Hermon (2005: 86): “the omission of *meN-* should be viewed as Wh-agreement, similar to that described by Chung 1982, 1994 for Chamorro.” However, I provide evidence in section 4.6 that the Case-based analysis is not adequate for the active voice deletion in Indonesian or Javanese.

options available to each language. In the present case, this means that (morpho-) phonology is subservient to the needs of syntax.

The idea that syntactic movement affects the morphology of verbs within its path is not a new idea. Kayne's (1989) analysis of participle agreement in French and other Romance languages, followed by Chomsky (1991: 1995; ch.2), is one of the most well-known attempts that represent this idea.⁷ Consider the examples in French given in (18a-c).

(18) French Participle Agreement

- a. Paul a *repeint*/**repeintes* les chaises.

Paul has repainted the chairs

'Paul has repainted the chairs.'

- b. Je me demande combien de tables Paul a *repeintes*.

I wonder how-many of tables Paul has repainted

'I wonder how-many of tables Paul has repainted.'

⁷ I thank Andrew Carnie (personal communication) for directing my attention to Kayne's (1989) work.

c. Je me demande combien de chaises il sera *repeint/*repeintes* cette année.

I wonder how-many of chairs Imp be.Fut painted this year

‘I wonder how-many of tables will be repainted this year.’


(Kayne 1989: 85, 86, 91)

French does not show overt agreement between a verbal participle and its selected direct object, as the example in (18a) illustrates. However, when the object undergoes *wh*-movement, as shown in (18b), there occurs overt agreement between the object and the participle. By contrast, this agreement does not occur in impersonal passive constructions, as shown in (18c).

Kayne argues that this pattern is naturally accounted for if there is an agreement relation established between the *wh*-phrase and Agr in the manner illustrated in (19) for (18b).

(19) The Structure of the Example in (18b)

[_{CP} combien de tables_i [_{TP} Paul ... [_{AgrOP} *t_i* [_{AgrO'} Agr [_{VP} *repeintes t_i*]]]]]


Spec-Head Agreement

Kayne himself assumes that the intermediate position above the VP for the purposes of agreement in (18b) is an adjoined position to the Agr projection. Chomsky (1991: 1995; ch.2) revises this original analysis and proposes that there is a Spec-Head Agreement relation between the moved *wh*-phrase and the participle that yields overt agreement morphology on the participle in (18b). Given that agreement is contingent on the Spec-Head Agreement relation between the moved NP and the AgrOP (see also Mahajan 1989; Koopman and Sportiche 1991 for further evidence for this claim), the Spec-Head Agreement relation in (19) is reflected morphologically on the participle either as the result of V-to-Agr raising or Agr-to-V lowering, a choice we can leave open here.

As noted above, this participle agreement is impossible in impersonal constructions as in (18c) above. Kayne argues that this is exactly what we expect under two independently motivated assumptions of the theory around the early 1990s: a) movement from an A' position to an A position is prohibited as a case of improper movement (see Fukui 1993 for an account of this prohibition in terms of chain uniformity) and b) LF expletive replacement is a form of A-movement (Chomsky 1986b). The presence of overt agreement means under Kayne's analysis that the movement of the *wh*-phrase *combien de chaises* 'how-many of chairs' has occurred to [Spec, AgrOP]. Kayne takes this movement to be an instance of A'-

movement.⁸ This movement is followed by further movement of the phrase into the position of the impersonal form *il* ([Spec, TP]) at LF. Kayne considers this movement a case of A-movement. This sequence of movement, thus, counts as improper movement and hence renders (18c) ungrammatical with agreement on a par with English examples such as (20).

(20) * [TP Mary_i seems [CP *t_i* that Mary loved *t_i*]].

However, (18c) is grammatical without agreement because the derivation of such an example does not involve A'-movement of the *wh*-phrase to [Spec, AgrOP], thereby circumventing the danger of improper movement.

In the following section, drawing on Kayne's analysis of French participle agreement, I propose that the voice-movement interaction in Indonesian and Javanese receives a straightforward account within the Phase Theory of Chomsky (2000, 2001, 2004, 2005).

⁸ Whether this assumption holds across languages is a matter of considerable debate, as Simin Karimi (personal communication) points out. Research as in Mahajan (1990) and Depréz (1989) argues, contrary to Kayne, that the specifier of the AgrOP is an A-position. I leave comprehensive discussion of what determines whether the specifier of AgrOP is an A- or A'-position for another occasion. The proposed analysis of *meN/ng*-deletion casts doubt on the A vs. A'-position.

3.2. *Phase Theory*

Let us review here several central assumptions of the Phase Theory as currently outlined in a series of Chomsky's work to lay the groundwork for the analysis presented in the next subsection. One central thesis of the derivational theory of syntax pursued within the Minimalist Program (Chomsky 2000, 2001, 2004, 2005, Epstein et al. 1998, Uriagereka 1999; Grohmann 2003) is that phonological and semantic information is transferred to the PF and LF interfaces 'online' in a piecemeal fashion (see the notion of the syntactic cycle in Bresnan 1971a, b, Jackendoff 1972, and Lasnik 1972 for an important earlier antecedent). The phase theory proposed by Chomsky is one particular version of this hypothesis. Chomsky proposes that this transfer occurs, not after every application of a structure building operation (as in Epstein et al. 1998), but instead at specific derivational cascades called *phases*. A phase is a mid-derivational object created by the syntactic computation that is headed by an instance of v or C.⁹ This theory significantly reduces computational complexity in that the derivation can forget about material after it has been transferred to the external systems. According to this theory, the complement of the phase-defining heads (i.e. VP or TP) is sent to the sound and meaning components for

⁹ Chomsky (2000, 2001, 2004) assumes that only those verbs that instantiate "full argument structure" (transitive and experiencer constructions) have strong phase heads, as indicated by v^* . In this chapter, however, I assume that every instance of v is a strong phase. See the following discussion on Legate (2003), who provides a number of phonological and semantic arguments that all types of verbs, including unaccusative and passives, constitute a strong phase.

phonological and semantic interpretation once a higher phase head is introduced into the derivation. More concretely, at the point where the ν P and CP structures have been built up, the complement domains of the ν and C, i.e. VP and TP, are transferred and interpreted at the interfaces. This model yields the so-called Phase Impenetrability Condition/PIC, which can be defined as in (21).^{10 11}

(21) Phase Impenetrability Condition (adopted from Chomsky 2001: 14)

In $[_{ZP} Z \dots [_{HP} \alpha [H YP]]$, where HP is a strong phase and ZP is the smallest strong phase, the domain of H is not accessible to operations at ZP; only H and its edge are accessible to such operations.

¹⁰ Chomsky (2000, 2004) proposes a slightly different formulation of the PIC, as in (i), adopted from Chomsky (2000:108).

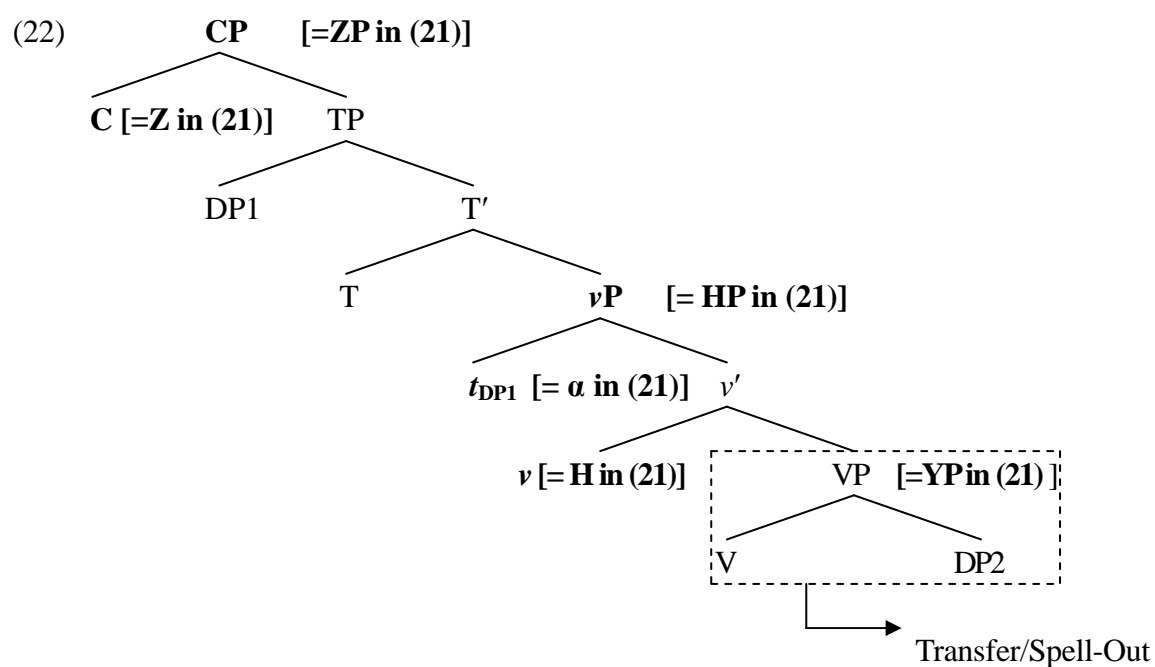
(i) Phase Impenetrability Condition (adopted from Chomsky 2000: 108)

In phase α with head H, only H and its edge are accessible to operations outside α .

I adopt the definition of the PIC given in (21) for the purposes of exposition in this chapter.

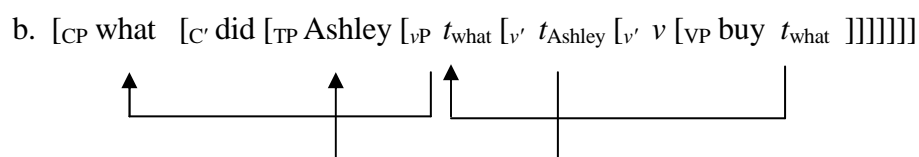
¹¹ As Heidi Harley (personal communication) points out, the PIC may mean that there is a functional processing benefit associated with overtly marking the phase boundaries within movement chains, as independently proposed by Givón 1979); see Lasnik 1999 for a critical discussion. Specifically, as long as a gap created by movement remains unfilled, phase edges must indicate this information, possibly motivating the parser to search further within the domain for the gap. I discuss this and a related question in chapter 8, when I propose to derive several architectural properties of the phase theory from interface conditions imposed by the A-P and C-I systems.

As stated above, Chomsky proposes that only ν P and CPs form a strong phase. The “edge” of the phase head H includes any specifiers of and any adjunct to H. Since CPs and ν Ps form a strong phase, the PIC states that the complement of the ν P phase cannot be a target for any operations at the CP phase, as shown in a schematic derivation such as (22). This is because, at the point the phase head C is introduced into the syntactic derivation, the VP domain of the lower phase head ν undergoes Transfer/Spell-Out to the interfaces for semantic and phonological interpretation.



In other words, the PIC requires that all elements base-generated within the complement of ν that are ultimately to move higher than the ν P phase must first move to the edge of ν ([Spec, ν P]) in the manner seen in (23b) for English sentences such as (23a).

(23) a. What did Ashley buy?



In this derivation, the C can access only the edge of the ν P due to the PIC. Thus, the direct object *what* passes through the escape hatch at the edge of ν P so that it may be accessible for further movement to [Spec, CP] at the CP phase level.

This phase-based computation leads us to a prediction that the local steps of movement, as illustrated in (23b), should somehow be reflected in the semantics and phonology of a language by the syntax-external components. Evidence of the first kind involves semantic facts concerning scope, binding, parasitic gaps, reconstruction, island effects, and so on. To take a recent case, Legate (2003) provides evidence for the notion of ν P phase based on binding and reconstruction, parasitic gaps, and antecedent-contained deletion (as well as the nuclear stress pattern of English). I reproduce Legate's argument concerning binding

here. The phase-based theory as explained above predicts that successive cyclic movement leaves copies in all intermediate ν Ps (as well as CPs). Drawing on the data and analyses presented by Lebeaux (1988) and Fox (1998), Legate shows that examples of A'-movement like (24a-c) that involve the interaction of binding and reconstruction effects can only be accounted for by the copy left by movement via the edge of ν P phases. Potential reconstruction sites are indicated by underlining in these examples.

(24) The Interaction of Binding and Reconstruction in *Wh*-Movement (A'-Movement)

- a. [Which of the papers that he_i gave Mary_j] did every student_i ✓ ask her_j to read * carefully?

(Legate 2003: 507)

- b.*[Which of the papers that he_i gave Mary_j] did she_j * ask every student_i to revise *?

(Fox 1998: 157)

On the one hand, Condition (C) of the Binding Theory requires that, in the example in (24a), the complex *wh*-phrase must not reconstruct to the thematic position of the verb *read* indicated by *. On the other hand, for the pronoun to receive a bound variable interpretation, it must be the case that the complex *wh*-phrase must reconstruct below *every student*. To satisfy both of these requirements, the *wh*-phrase must leave a copy

In the example in (25a), the copy left at the edge of the lower ν P phase by the successive cyclic movement of the complex *wh*-phrase meets two binding-theoretic requirements: *he* is correctly c-commanded by *everyone* and *Mary* is not in a position to be bound incorrectly by *her*. The example in (25b) is ungrammatical because there is no single position that meets both requirements. The same story can be told for the examples in (26a, b) that involve unaccusative verbs like *escape* meaning *forget* (Pesetsky 1995). In (26a), the copy left in the specifier of the lower ν P phase serves to provide the unique appropriate position for the purposes of Condition C and variable binding. A comparable position is not available in the example in (26b). Thus, the contrast between (25a)/(26a) and (25b)/(26b) suggests that passive and unaccusative verbs also constitute strong phases, contra Chomsky's (2000, 2001, 2004) claim that only transitive and experiencer verbs count as strong phases (cf. note 6).

Evidence of the second kind involves reflexes of local movement whose distribution can only be accounted for by positing successive cyclic movement, as in complementizer agreement in Irish (McCloskey 1979; 2001) and Chamorro (Chung 1982, 1994, 1998). I show in the next subsection that the distribution of the active voice morphology in Indonesian and Javanese adds to this type of evidence for the locality of movement required by the PIC.

3.3. A Phase-Theoretic Analysis of the Active Voice Deletion in Indonesian and Javanese

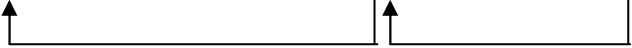
Let us now come back to the central point of this chapter, the distribution of the active voice morphology in Indonesian and Javanese, and see how the phase-based analysis correctly derives Cole and Hermon's (1998) generalization. Consider (11b) from Javanese, repeated here as (27a), for illustration. The derivation for the example is given in (27b).

(27)a. Sapa_i Esti (***nge**)-sun *t_i*? (=11b)

who Esti **AV**-kiss

'Who did Esti kiss?'

b. [CP Sapa ... [TP Esti [vP *t*_{Esti} [v' *t*_{sapa} [~~ng~~-v [VP sun *t*_{sapa}]]]]]]]



Spec-Head D-Feature Checking

In this derivation, the *wh*-phrase *sapa* 'who' undergoes successive cyclic movement through the edge of the vP phase into the edge of the CP phase in a manner required by the PIC reviewed above. The external argument *Esti* is base-generated in the specifier of vP and undergoes A-movement into [Spec, TP], as typically assumed for languages such as English.

When the *wh*-phrase stops by in the edge of *v*P phase, Spec-Head D-feature checking occurs between the moved phrase and its local little *v*, as shown in (27b). This syntactic agreement is realized in the syntax-external phonological component as the deletion of the nasal prefix *ng-* in Javanese. More specifically, this analysis claims that only the derivation in which *ng-* is not present on the series of transitive verbs crossed by the movement of an NP converges at the syntax-external interface. Note that this analysis is strikingly parallel to the analysis for French participle agreement developed by Kayne (1989). The fact that the movement of non-nominal phrases does not trigger *meN-/ng-* deletion also naturally falls out since those phrases do not have D-features to be checked against their local *v* heads. The same analysis holds for the movement-voice interaction in Javanese and Madurese.

It should be noted here that, to the extent that the current analysis is tenable, the facts examined in this chapter provide direct support from the syntax-phonology interface for the role of *v*P phase. This point bears emphasis for the following reason. For reasons detailed above, the phase-bound derivational system as in Chomsky (2000, 2001, 2004, 2005) predicts that movement occurs at the edge of *v*P and CP. Evidence abounds in the literature for successive cyclic movement via intermediate [Spec, CP]s. To mention but a few, complementizer alternation in Irish and Chamorro (McCloskey 1979, 2001; Chung 1982, 1994, 1998), the complementizer agreement in several dialects of Dutch, German,

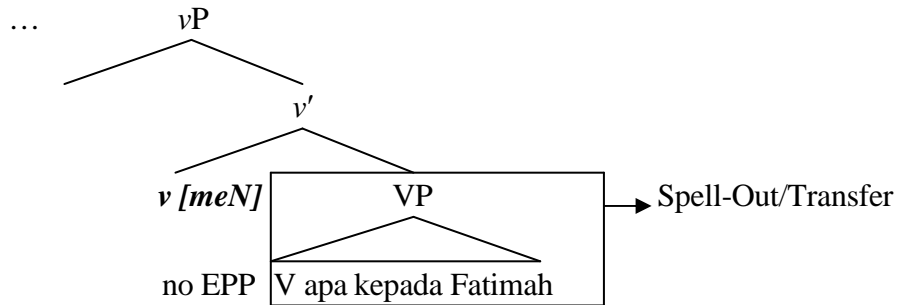
Frisian, and West Flemish (Haegeman 1992; Zwart 1993, 1997; Watanabe 2000), *wh*-copying in languages such as Frisian, Dutch, Afrikaans, German, and Child English (Felser 2004; Thornton 1995) as well as the existence of partial *wh*-movement in languages such as German, Hungarian, and Indonesian (McDaniel 1989; Horvath 1997; Saddy 1991; Cole and Hermon 1998, 2000; see also chapter 5) provide morphological evidence that *wh*-phrases move through the edge of CP phases. However, evidence has been somewhat hard to come by for comparable cyclic movement via the edge of vP, though the WH-agreement in Chamorro (Chung 1982; 1994, 1998), Innu-aimûn (Branigan and MacKenzie 2002), and Kilega (den Dikken 2001), and the obligatory loss of tonal downstep in Kikuyu that is triggered by *wh*-movement (Clements *et al.* 1983; Clements 1984; Sabel 2000) have been sometimes interpreted as possible arguments that *wh*-movement occurs through the edge of vP phases; see Radford (2004), for example. The pattern of the active voice deletion in Indonesian and Javanese discussed here, however, precisely fills this empirical gap and provide clear morpho-syntactic evidence that the movement of an NP stops at the edge of each vP phase that intervenes between the launching and landing site of the movement. In this regard, the contribution to linguistic theory from under-represented languages such as Indonesian and Javanese is highly significant.

Note further that the proposed analysis provides further support for Legate's (2003) view that passive *v*Ps are strong phases, contra Chomsky's (2000, 2001, 2004). In the derivation of passive and object preposing constructions, the active voice deletion is triggered by the Spec-Head D-feature checking requirement between the A-moved NP and its local little *v*. This means that A-movement also passes through the edge of *v*P.¹²

The proposed analysis of the active voice deletion in Indonesian and Javanese is related to another phase-theoretic account recently proposed by Aldridge (in press). Drawing on certain parallels between *meN*- clauses and Tagalog-type anti-passive constructions, Aldridge proposes that the feature bundle inserted in *v* and spelled-out as *meN*- cannot include an EPP feature, hence blocks objects from moving out of the VP due to the PIC. To illustrate her analysis, consider the schematic derivation for the example in (4b) in (28).

¹² Two notes are in order. First, following Guilfoyle et al. (1992), I assume that the passive voice marker *di-* in Indonesian is a clitic-like element that needs to be incorporated into a verb. This assumption receives morphological support from the fact that *di-* can be considered as an allomorph of the third person singular pronoun *dia* (Guilfoyle et al. 1992; 400). Second, it is impossible to construct a parallel argument for the phasehood of unaccusative verbs. This argument, in principle, could be made if Indonesian has transitive unaccusatives like *escape* in the sense of 'forget' (see (26a, b)). However, my informant work so far has not been successful in identifying the Indonesian analogue of verbs like *escape*. Unergative verb such as *nyanyi* 'sing' may occur with or without *meN*- in both its transitive and intransitive use whereas other verbs such as *dansa* 'dance' only occurs with the derivational affix *ber-*.

(28) C [= phase head]



Aldridge assumes that the v head spelled-out as *meN-* in Indonesian cannot have the EPP feature. Recall that the PIC dictates that the object that is to be ultimately moved into the higher CP phase must first move into the edge of the v . Since the v head in (28) does not carry the EPP feature, the movement of the object into [Spec, CP] violates the PIC. The example in (4b) is grammatical without *meN-* on the higher verb because the v head spelled-out this way can carry the EPP feature that triggers the movement of the direct object into the specifier of the vP in (28).

Two problems remain with Aldridge's analysis, however. First, it is not clear what prevents *meN-* from having an EPP feature. Under Chomsky's (2000, 2001, 2004, 2005) version of the EPP-assignment on phase heads, nothing blocks transitive v heads from being assigned an EPP feature that would allow movement of an VP-internal element to

the its edge.¹³ Aldridge does develop a diachronic explanation for the lack of EPP features on *v* heads in Indonesian by analyzing *meN-* as anti-passive morpheme that can be considered a historical residue of the ergative system of the Philippine-type languages as in Tagalog and Malagasy. However, this explanation could lead to the claim that Indonesian is an ergative language. This reasoning might be undermined by the argument made by Ndayiragije (2006) that Kirundi is a nominative-accusative language yet has antipassives but it is at least a reasonable move given the observation that the antipassive voice is found mostly in ergative languages. There are two arguments that Indonesian is morphosyntactically not an ergative language. First, as we will see in section 4.1, Chung (1976) and Cole and Hermon (2005) present evidence that direct object extraction is possible in Indonesian. Given that the direct extraction of non-subject argument is impossible in many of the ergative Austronesian languages such as Tagalog, Malagasy, and Formosa, Chung's discovery suggests that Indonesian is not an ergative language of the Tagalog-type. Second, Chung (in press) claims that the derivation of the Indonesian SVO word order is derived by simply raising the external argument to [Spec, TP], in much the same way as the derivation of the English SVO order. Several researchers in

¹³ Chomsky (2000, 2001, 2004) claims that the EPP-assignment on the phase edge is done only when movement has an effect on semantic interpretation at the interface. If the proposed analysis is correct, the EPP assignment on the phase edge is motivated by PF effects because the EPP-checking results in the *meN-* deletion at the PF interface.

Indonesian as in Cartier (1989) and Verhaar (1989) argued that Indonesian is morphosyntactically an ergative language on the grounds that both the regular *di*-passive and the so-called zero-passive/Passive Type Two are available in this language (cf. (8); see also section 4.1 for detailed discussion). This argument, Chung argues, is incompatible with the observation confirmed by Anderson (1976: 18-19), Manning (1996), and other researchers that morphosyntactic ergativity is found in verb-initial or verb-final languages, but not in verb-medial languages. Thus, to the extent that the Indonesian is neither a verb-initial or verb-final language, as Chung claims, then this language is not a morphosyntactically ergative language of the Tagalog-type, as proposed by Aldridge.

The second problem with Aldridge's analysis is that we would predict that non-nominal phrases such as prepositional and adjunct phrases moving from embedded to matrix clauses should trigger deletion on the matrix verb, since they must also move through intermediate phase edges as well in order to extract long distance, and hence must check EPP features on such phase heads. This prediction is false since, as we saw earlier, the movement of these phrases does not trigger *meN*-deletion in this context. This suggests that it is not an EPP feature but a D-feature that is responsible for this deletion.¹⁴

¹⁴ This argument was suggested to me by Heidi Harley (personal communication).

One important question that arises in this connection is the following.¹⁵ Under the present analysis, it is the D-feature checking relation between the specifier and the head of the ν P that triggers active voice deletion on the verb dominated by the ν P. If so, why is it that the merger of external arguments such as *Esti* in (11b) does not cause the deletion of the active voice marker? An answer to this question is available once we introduce the distinction between *External Merge* and *Internal Merge*. Chomsky (2004) proposes that the concatenative operation of Merge comes free in two forms; when X is merged with Y, X can be either external to Y (External Merge) or part of Y (Internal Merge). Chomsky (2004) (see also Richards 2007) argues that pretty much everything is driven by phase heads other than External Merge, which is triggered solely by selectional features. For example, Chomsky (2004: 111) maintains that “Argument structure is associated with external Merge (base structure), everything else with internal Merge (derived structure).” It is clear that whatever feature that causes *meN-/ng*-deletion does not belong to argument structural/ θ -theoretic features but instead forms some formal computational feature that is triggered by phase heads such as ν heads. Indeed, we have seen evidence above that the deletion is sensitive to Internal Merge of an NP across a transitive verb. Also recall from Kayne’s (1989) analysis that French

¹⁵ Thanks to Heidi Harley (personal communication) and Simin Karimi (personal communication) for raising this question and providing possible directions to take with respect to this question such as the External vs. Internal Merge distinction.

participle agreement is sensitive to Internal Merge, as the contrast between (18a) and (18b) illustrates. Then, the fact that the External Merge of an external argument does not cause active voice deletion in examples like (11a) follows if only Internal Merge has consequences such as the active voice deletion for the sound-related syntax-external interpretive component. This assumption is also in keeping with the “complementarity between θ -theory and formal feature checking” that has been expressed in one way or another in the minimalist literature (e.g., Chomsky 1995), namely, that all arguments must form non-trivial chains by movement to checking their formal features with an appropriate functional category such as T or ν (see López 2001, though, for an opposing argument from exceptional case-marking subjects).

The assumption that the θ -requirement and formal feature checking is dissociated this way is also conceptually natural in view of the following observation. The ν is considered both lexical and functional at the same time, lexical in that it selects an external argument in its specifier and functional in that it assigns accusative Case to the direct object. These two roles were distributed by AgrO and the V within the Agr-based Case Theory but are now served by the ν in the more recent Agr-less ν P-approach. This claim, in turn, has led to controversies on whether the A vs. A'-distinction can be established in syntax. Thus, we need some principled way to encode this dual behavior of the ν head. The distinction between Internal Merge and External Merge seems to be able to make the desired

distinction; the v head assigns configurational properties such as Case and Agreement to only an Internally Merged object but, at the same time, assigns a θ -theoretic interpretation to an Externally Merged object based on its configuration (Hale and Keyser 1993, 2002). Given that active voice deletion belongs to the former type of feature-driven computation, it stands to reason that the External Merge of an external argument as in (11a) does not have any consequence for the fate of the active voice morphology at the phonological interface.

It is in this regard that the proposed analysis is crucially different from the most recent analysis presented by Cole and Hermon (in press). Cole and Hermon propose an account of the distribution of the active voice morphology in Malay/Indonesian that assimilates the active voice deletion pattern to the more general Philippine-type voice agreement discussed in Rackowski and Richards (2005). They propose (p. 15) that “the presence of *meng-* indicates that the agent is the highest specifier of vP , and, hence, that object shift has not occurred.” Conversely, then, the lack of *meng-* indicates that the non-external argument is moved to the edge of vP phase to be accessible to movement into [Spec, CP]. Thus, the impossibility of the movement of a non-external argument across the *meN-* verb is naturally predicted because the presence of *meN-* means the lack of the object shift of a post-verbal argument into the edge of vP , thereby blocking its movement across *meN-* due

to the PIC. This line of approach, in fact, has been already proposed in a slightly different form by Guilfoyle et al. (1992: 385-387) within the Government-and-Binding Theory. Assuming that *meN*- and *di*- are Actor Topic and Theme Topic, respectively, Guilfoyle *et al.* argue that the impossibility of extraction of an object across *meN*- is a natural consequence of the general fact observed in many Austronesian languages such as Tagalog and Malagasy that only the NP which agrees with the verb which has the appropriate topic/voice morphology is extractable.

Cole and Hermon further note (p.15) that the agreement-based analysis makes a correct prediction that the extraction of an NP across the higher verb, not the lower verb, in a bi-clausal environment, as in (29) (=4a), triggers the *meN*-deletion only from the higher verb.

(29) Siapa_i Bill (***mem**)-beritahu ibu-nya [CP yang *t_i* (**men**)-cintai Fatimah]? (=4a)

who Bill **AV**-tell mother-his that **AV**-love Fatimah

‘Who does Bill tell his mother that loves Fatimah?’

According to Cole and Hermon, this deletion pattern is explained as follows: the deletion of the *meN*-(or the null prefix \emptyset_{meN} - in their terminology) signals the agreement between the higher *v* head and the CP clause out of which the movement has occurred.

Cole and Hermon's agreement-based analysis appears to make the exact same predictions concerning the distribution of the active voice morphology as the proposed D-feature checking analysis. Three considerations suggest, however, that the present analysis is to be preferred on both theoretical and empirical grounds. One difference between the two analyses lies in their view that *meN-* indicates agreement between the ν and the external argument base-generated in the specifier of the ν head. Under the complementarity between θ -theory and formal feature checking noted above, their analysis would amount to the claim that this complementarity is broken down in Indonesian. Unless evidence is presented for this view, however, it is at least theoretically desirable to eliminate this extra stipulation, as in the proposed D-feature checking analysis. Another difference lies in their treatment of the pattern of *meN*-deletion in long-distance extraction as illustrated in (29). As we have seen, Cole and Hermon claim that the null prefix occurs on the higher verb due to the agreement between the upstairs ν and its complement CP. However, they do not make it explicit in what formal feature they Agree with each other. Let us then consider what feature would be relevant for their analysis to go through. The possibility that the feature checked is Case is unlikely because verbs like *beritahu* 'tell' in (29), like their English counterparts, can only assign one accusative Case to its immediately following Goal DP. Another possibility that the feature checked is the nominal D-feature also cannot be maintained because CPs as in

(29) that occupy non-subject positions have been standardly assumed to lack D-features. Note that this kind of problem does not arise under the proposed analysis. The proposed analysis explicitly states that the higher *meN*- in (29) is elided due to the *D-feature* checking relation between the upstairs *v* and the successively moved NP. The final difference, related to the second, is the following.¹⁶ Cole and Hermon argue that it is the agreement of the complement CP with the matrix verb in (29) that triggers *meN*-deletion. Their analysis predicts that matrix *meN*-deletion should occur on the matrix verb in the configuration in (29) even when a non-DP argument is extracted from an embedded CP. This is because the syntactic category of a moving element should have nothing to affect the agreement between the verb and its complement CP. This prediction, however, is incorrect, as shown in (30), which minimally contrasts with (29), in that what moves in the former is a non-nominal (prepositional) expression, because *meN*- on the matrix verb does not undergo obligatory deletion in (30), contrary to what Cole and Hermon's analysis predicts.

(30) [_{PP} Kepada siapa]_i Bill (**mem**)-beritahu ibu-nya [_{CP} Fatimah (**mem**)-beri buku *t_i*]

to whom Bill AV-tell mother-his Fatimah AV-buy book

'To whom did Bill tell his mother that Fatimah gave a book?'

¹⁶ Many thanks to Heidi Harley (personal communication) for suggesting this argument to me.

On the other hand, the obligatory deletion of *meN-* in (30) naturally falls out from the present analysis because non-nominal expressions such as *kepada siapa* ‘to whom’ do not have D-features. On the above-mentioned grounds, I conclude that, although Cole and Hermon’s analysis does seem to make essentially the same set of predictions as the D-feature checking analysis concerning Indonesian facts, the latter analysis is preferred on both conceptual and empirical grounds.

3.4. *meN-/ng-Deletion as Failure of Vocabulary Insertion*¹⁷

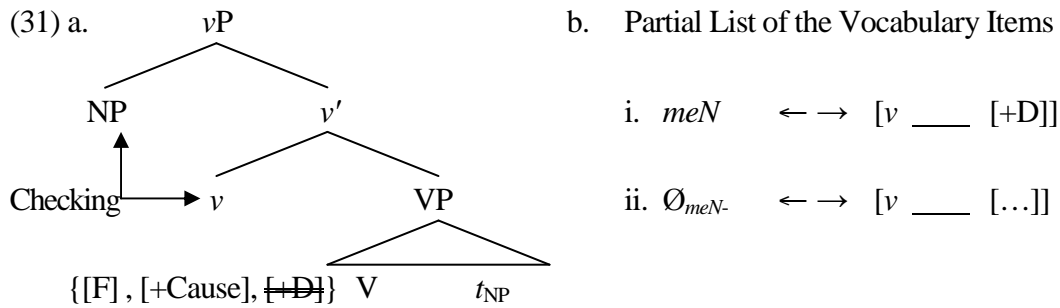
So far, I have remained vague on the precise nature of “deletion” involved in the present phase-theoretic analysis of the active voice deletion in Indonesian/Javanese. The issue of exactly what “deletion” means in contemporary syntactic theory has been far from settled. In this section, following Harley (2005), I propose that the active voice deletion is failure of the post-syntactic insertion of *meN-/ng-* at the PF interface within the morphosyntactic framework of Distributed Morphology.

It is claimed within the Distributed Morphology that phonological features are assigned to abstract morphemes (as well as roots under certain views) post-syntactically.

¹⁷ Many thanks to Heidi Harley (personal communication), Mark Baker (personal communication), and Andrew Carnie (personal communication) for encouraging me to formulate a rule for the deletion and very useful discussion. See detailed explication of theoretical claims of Distributed Morphology in chapter 6.

The mechanism to assign this feature is Vocabulary Insertion in the technical sense; “the Vocabulary is the list of the phonological exponents of the different abstract morphemes of the language, paired with conditions on insertion. Each such pairing of a phonological exponent with information about the grammatical (i.e. syntactic and morphological) context in which the exponent is inserted is called a vocabulary item.” (Embick and Noyer 2007: 297). A list of potential candidates compete for the same abstract morpheme position in the morphosyntactic derivation; which candidate is selected as the exponent of that position is determined by the Subset Principle proposed by Halle (1997).

I propose that the Spec-Head D-feature checking between the moved NP and its local v head changes the D-feature of the v and that this change in the feature content of the head drives *meN/ng-* out of the candidate set for insertion under the v head at PF. Suppose for the sake of argument that the v head has the bundle of c features consisting of [+F(unctional)], [+Cause/Do/Have], and [+D-(feature)] and that the *meN-* and its null variant *ØmeN-* is inserted under the condition specified in (31).



Under my proposal, the checking relation illustrated in (31a) deletes the [+D] feature of the *v* head. Since *meN*- can only be inserted under the *v* head with the corresponding feature, as shown in (31bi), *meN*-insertion is correctly blocked; then, the null counterpart \emptyset_{meN-} is inserted as an elsewhere case, as shown in (31bii). This proposal, thus, sheds new lights on the precise nature of active voice deletion in Indonesian/Javanese and of ellipsis processes, more generally.

3.5. Section Summary

To sum up this section, I have proposed, following the insight of Kayne's (1989) analysis of participle agreement in French, that Cole and Hermon's generalization is naturally derived by analyzing the active voice deletion as the PF reflex of a Spec-Head Agreement relation between the moved NP and its local *v* head at the *vP* phase. The movement of non-nominal phrases does not cause active voice deletion because they lack

D-features to be checked in this configuration. The data examined here, thus, provides empirical support for the phase-based implementation of successive cyclicity, in particular, the role of the ν P phase at the syntax-phonology interface.

4. Other Alternative Accounts

In this section, I compare the proposed phase-theoretic analysis with several existing and potential analyses of the *meN*-deletion in Indonesian and *ng*-deletion in Javanese, by extension:

a) the NP Accessibility Hierarchy and the Subjects-Only Restriction b) Soh's (1998) Relativized Minimality account, c) Fortin's (in press) antipassive account, d) Anti-Agreement-based accounts, and e) Case-based accounts. I show that each of these analyses encounters a particular set of empirical and conceptual problems that are successfully resolved in the present phase-theoretic approach.

4.1. *NP Accessibility and Subject-Only Restriction*

In this subsection, I review and reject a potential analysis of Cole and Hermon's generalization based on the so-called *Accessibility Hierarchy*.¹⁸ Based on a large-scale survey of many western Malayo-Polynesian languages, Keenan and Comrie (1977: 70)

¹⁸ Thanks to Paul Kroeger (personal communication) for suggesting this possible account.

(see also Keenan 1972) observe that only subjects can be relativized. This observation is also made in the one of the most comprehensive grammars of Indonesian in Snedon (1996), who notes (p. 286) that “a relative clause can contain any constituent occurring in an independent clause except the subject, which is identical to the head of the embedding noun phrase.” The accessibility-based analysis is further articulated in the minimalist framework by Nakamura (1994), which derives this subject-only restriction in terms of the global economy principle (“minimize chain links” in the sense of Chomsky and Lasnik 1993) that essentially forces non-subject arguments to be first promoted to a subject/topic position for the purposes of creating the shortest possible link in a set of alternative competing derivations. Under this analysis, examples that involve extraction of an apparent non-subject argument all involve passivization of this argument into [Spec, TP] prior to movement to [Spec, CP]; the active voice morphology, then, is deleted as a natural result of the passivization.

This analysis seems desirable in two respects. First, it fits the apparent mysterious restriction on the active voice deletion in Indonesian and Javanese into the independently motivated restriction or whatever account derives such a restriction. Second, this analysis appears to be particularly true for Indonesian because this language has what several Indonesianists have variously called “Passive Type Two,” “Object Preposing”, “Object

Voice,” “Subjective Passive”, or “Bare Passive” (Sneddon 1996; see also Chung 1976, Kana 1986; Guilfoyle et al. 1992; MacDonald and Dardjowidjojo 1967; Sie 1988; Cole and Hermon 2005, Arka and Manning 1998), in which a bare stem is immediately followed by a subject, preferably a pronominal subject, as shown in (32b) (recall the examples in (8a, b), whose derivation also instantiates the Passive Type Two).¹⁹

(32) Active and Passive Type Two Sentences

a. Aku sudah **mem**-beli buku itu.

I already **AV**-read book that

‘I already read that book.’

b. Buku itu sudah Ø-beli aku/ku.

book that already **OV**-buy me

‘I already read that book./That book was already read by me.’

If the subject-only restriction is correct for Indonesian and Javanese, then all the examples that involve *meN*-deletion may be analyzed as passivization of the non-subject argument

¹⁹ A construction similar to the Passive Type Two in Indonesian is found in Malagasy and Tagalog with the Theme-Topic marker. See Rackowski and Richards (2005) for Tagalog and Pearson (2001, 2005) for Malagasy.

into [Spec, TP], followed by A'-movement into [Spec, CP]. However, there are many arguments presented in the literature that this analysis is untenable. I review several of those arguments below.

4.1.1. Indonesian

The first argument against the accessibility-based analysis comes from its prediction that all A'-extractions should pass through [Spec, TP], which was shown to be false in Malay/Indonesian by Soh (1998) on the basis of the relative position of subjects with respect to aspectual markers, weak crossover effects, long-distance extraction and the *that*-trace asymmetry. To repeat one of her arguments based on weak crossover effects, it has been widely known that a) a pronoun must be c-commanded by a binder and its variable at the surface/derived structure to be construed as a bound variable, as shown by the contrast between (33a) and (33b) and that b) the trace created by A-movement provides a new binder for a pronoun, as shown in (33c) (Mahajan 1990: 24).

(33) Weak Crossover Effects in English

- a. Who_i t_i saw his_i mother?
- b.* Who_i did his_i mother see t_i?
- c. [Who_i t_i seems to his mother [t_i to have come]]. (Soh 1998: 300)

Then, the prediction of the passivization-based account is that extraction of a non-subject argument should not cause a weak crossover effect, as in (33c), but, as the ungrammaticality of (34b) shows, the bare passive counterpart to (34a), shows that this prediction is wrong; note that, in a ‘real’ passive of the English kind in Indonesian given in (34c), we don’t have weak crossover effects, a pattern that suggests that the movement of *siapa-kah* ‘who-Q’ is preceded by the prior movement to [Spec, TP].

(34) Weakcrossover effects in Malay/Indonesian

- a. Emak-nya_i sayang Ali_i.
 mother-his love Ali
 ‘His mother loves Ali.’

b.* Siapa_i-kah yang emak-nya_i sayang t_i ?

who-Q that mother-his love

‘Who does his mother love?’

c. Siapa_i-kah yang t_i di-sayangi emak-nya_i t_i ?

who-Q that OV-love mother-his

‘Who is loved by his mother?’

(Soh 1998: 300)

The presence of the weak crossover effect in the example in (34b) and the lack thereof in the example in (34c) indicate that the subject-only restriction cannot be the proper account for the active voice deletion in Indonesian.

The second argument against the accessibility-based account is that made by Cole and Hermon (2005); see also Musgrave (2001). Following Chung (1976), Cole and Hermon provide numerous arguments that the distributional restrictions on the Passive Type Two do not hold for object relativization in Indonesian, contrary to what the Accessibility-based analysis would lead us to predict. In addition to the two properties noted above, in the Passive Type Two, the negative marker *tidak* and auxiliaries must precede the agent. This is illustrated by the contrast between (35a) and (35b).

(35) Negation and Auxiliaries Precede the Agent in Passive Type Two

a. Buku ini tidak akan kami Ø-baca.

book this not will we OV-read

‘This book will not be read by us.’

b.* Buku ini kami tidak akan Ø-baca.

book this me not will OV-read

‘This book will not be read by us.’

(Cole and Hermon 2005: 62)

These characteristics of the Passive Type Two, thus, restrict the possible word orders to the [Neg Aux Agent-Verb]. Cole and Hermon show that this word order restriction does not hold for object relativization. If object relativization were done through Passive Type Two, followed by null operator movement of the derived subject, as the NP-Accessibility analysis would argue, then the agent must follow negatives and auxiliaries in cases of object relativization since the input configuration for this relativization should be Passive Type Two. This prediction is false because object relativization allows agents to precede negation and auxiliaries, as shown in (36a, b).

(36) Object Relativization in Indonesian

- a. [Buku [yang Budi tidak akan baca]] sangat menarik.

book that Budi will not read very interesting

‘The book that Budi will not read is very interesting.’

- b. Anak [yang Wati tidak pukuki]] itu men-angis.

child that Wati not hit that AV-cry

‘The child that Wati didn’t hit is crying.’ (Cole and Hermon 2005: 64)

This order is unacceptable in the Passive Type Two, as in (37a) and (38a). By contrast, the active counterparts to (37a) and (37a) in (39a, b) are grammatical. Note further that the relativization based on the Passive Type Two construction in (35a) is, of course, grammatical, as in (40).

(37) Passive Type Two and Active Sentences in Indonesian

- a.* Buku itu dia tidak akan baca. b. Buku itu tidak akan dia baca.

book that he not will read book that not will he read

‘The book will not be read by him.’ ‘The book will not be read by him.’

(Cole and Hermon 2005: 64)

(38) Passive Type Two and Active Sentences in Indonesian

a. * Anak itu kami tidak sedang jemput. b. Anak itu sedang kami jemput.

child that you not Prog pick up child that Prog you pick up

‘The child is not being picked up by us.’ ‘The child is not being picked up by us.’

(Cole and Hermon 2005: 64)

(39) Active Counterparts to the Examples in (36a) and (37a) in Indonesian

a. Dia tidak akan mem-baca buku itu.

he not will AV-read book that

‘He will not read that book.’

b. Kami tidak sedang mem-jemput anak itu.

you not Prog AV-pick up child that

‘We are not picking up that child.’

(Cole and Hermon 2005: 64)

(40) The Example of Object Relativization Based on the Passive Type Two

Buku [yang tidak akan kami Ø-baca] sangat menarik.

book that Neg will me read very interesting

‘The book that I won’t read is very interesting.’

The data examined above, therefore, provides evidence that non-subject relativization can occur without prior promotion/movement to [Spec, TP], hence argues strongly against the account of *meN*-deletion in terms of the NP-Accessibility. I refer the reader to Chung (1976), Cole and Hermon (2005) and Musgrave (2001) for numerous other arguments along these lines against such an analysis.

Finally, the Accessibility Hierarchy encodes the universal statement that all human languages allow at least relativization of a subject. Specifically, Keenan and Comrie (1977: 67) make such a statement in the form of the Hierarchy Constraint “A language must be able to relativize subjects.” However, the Accessibility-based account by itself predicts that not only local extraction but also long-distance extraction of a subject should be acceptable with or without *meN* on its extraction path because subject relativization is available in all languages. We have seen above that examples such as (4a, b), (5a, b), and (6a-c) clearly falsify this prediction. This indicates that the distribution of the active voice morphology in Indonesian is something beyond Keenan’s (1972)/Keenan and Comrie’s (1977) Accessibility Hierarchy.

4.1.2. Javanese

A similar argument against the NP Accessibility-based account for the active voice morphology in Javanese was made by Cole et al. (1999). Examples as in (41a, b) show that

relativization of an object results in the deletion of the active voice marker *ng-*. The analysis based on accessibility would explain this deletion as the result of the passivization that applied prior to operator movement.

(41) Object Relativization

- a. * Buku kuwi [DP [CP sing Budi ***m**-waca

book that that Budi AV-read

‘That book is the one that Budi read.’

- b. Buku kuwi [DP [CP sing Budi maca]]

book that that Budi read

‘That book is the one that Budi read.’

(Cole *et al.* 1999: 88)

This analysis, thus, makes a prediction that Javanese should also have a construction akin to Passive Type Two, as found in Indonesian, a prediction that is incorrect. As Cole *et al.* (1996: 92) observe, the bare passive construction does not occur in Javanese. Thus, the NP-accessibility analysis would wrongly predict that examples as in (42a, b) are ungrammatical, contrary to facts.

(42) Object Relativization in Javanese

- a. Bukune sing aku waca lucu. b. Montor sing Siti setir abang.
 book-ne that I read funny car that Siti drive read
 ‘The book that I read is funny.’ ‘The car that Siti drives is red.’

(Cole et al. 1999: 92)

The grammaticality of object relativization in (42a, b), thus, argues against the NP Accessibility-based account of the *ng*-deletion.

4.2. *Soh’s (1998) Relativized Minimality Account*

Soh (1998) provides a new account of the syntactically governed distribution of *meN*- in Malay/Indonesian in terms of Relativized Minimality (Rizzi 1990). According to this account, the active voice marker *meN*-, base-generated in the A'-position within the VP, blocks A'-movement of a phrase across the marker, in violation of the minimality constraint defined as in (43) and (44).

(43) Relativized Minimality

Antecedent Government: X antecedent-governs Y iff

(i) X and Y are coindexed

(ii) X c-commands Y

(iii) no barrier intervenes

(iv) Relativized Minimality is respected (Rizzi 1990: 6-7)

(44) Relativized Minimality: X α -governs Y only if there is no Z such that

(i) Z is a typical potential α -governor for Y,

(ii) Z c-commands Y and does not c-command X. (Rizzi 1990: 6-7)

To take (4a) for illustration, repeated here as (45), this example is ruled out with the active voice marker on the higher verb because *meN-* blocks A'-extraction of the *wh*-phrase *siapa* 'who' across it, in violation of the Relativized Minimality defined as in (44).

(45) *Siapa_i Bill (*mem)-beritahu ibu-nya [yang t_i (men)-cintai Fatimah]?*

who Bill AV-tell mother-his that AV-love Fatimah

'Who does Bill tell his mother that loves Fatimah?

We have seen in section 2 that *meN-* does not block movement of non-nominal phrases such as prepositional objects and locative adjuncts across the verb marked with it, as shown in (46) (recall also the examples in (10a, b) and (13a, b)).

(46) Movement of non-*Wh*-NPs

[_{PP} kepada siapa]_i Bill (**mem-**)beritahu ibu-nya [bahwa Esti (**mem-**)beri motor _i]_i?

to who Bill **AV**-tell mother-his that Esti **AV**-give motor bike

‘To whom did Bill tell his mother that Esti gave a motor bike?’

Soh (p. 304) claims that this NP-PP asymmetry is accounted for if Minimality is relativized not only in terms of position (A vs. A') but also category (NP or PP). Then, a position such as the base-generated position of *meN-* is restricted to NPs, thus *meN-* only blocks extraction of an NP, not PP, over the voice marker.

Soh's Relativized Minimality account is quite close to the proposed D-feature checking analysis of the active voice deletion. The two analyses essentially block movement of an NP across the active voice verb from independently motivated constraints on syntactic derivation. However, two problems remain unresolved in Soh's analysis. First, Soh assumes that *meN-* is in an A'-position. This analysis, thus, predicts that only A'-movement of an NP

across the verb with the active voice morphology is prohibited. We have seen, however, that A-movement such as object preposing also causes the active voice deletion of a verb that it crosses, as illustrated in (8a, b), repeated here.

(47) Object Preposing

- a. Buku itu_i adik saya (***mem**)-beli t_i.

book that brother my **AV**-buy

‘My brother bought that book./That book was bought by my brother.’

- b. Ali_i saya (***men**)-cubit t_i.

Ali I **AV**-pinch

‘I pinched Ali./Ali was pinched by me.’

Thus, Soh’s analysis falsely predicts that examples like (47b) should be grammatical with *meN*- marked on the verb that the movement has crossed because the movement involved is an A-movement that thereby should not be blocked by an A’-element like *meN*-. On the other hand, as we have shown in section 3, the phase-theoretic D-feature checking account correctly predicts that whatever movement is involved across an active verb results in the deletion of *meN*- across it. Another problem with Soh’s analysis concerns the base-position

of *meN-*. She assumes (p. 297) that “*meN-* occupies a position above the verb and below the internal subject position,’ but this assumption is difficult to sustain. Within her Relativized Minimality account, *meN-* should be base-generated in an A’-position to block A’-movement of an NP across the active voice marker; the blocking effect *would* not arise if it were base-generated under V because a head position is not normally characterized as an A vs. A’ position (cf. Li 1990). Updating Soh’s remark above within the more recent split *vP* model, the relevant position would arguably correspond to a lower specifier of the *vP*. However, the assumption that voice markers such as *meN-* are base-generated in a specifier of a functional head (be it *v* or Voice) is not a commonly held assumption and requires independent empirical motivation.²⁰ Based on these grounds, I conclude that the present phase-theoretic analysis is empirically superior to Soh’s Relativized Minimality account.

4.3. Voskuil’s (2000) Pro-Based Account

Voskuil (2000) attempts to account for the distribution of the active voice morphology in Indonesian from the interaction of several constraints and principles. Voskuil assumes that

²⁰ True, this assumption is not utterly unreasonable, as long as Fortin’s (in press) argument that *meN-* in Indonesian is an object clitic holds. However, I provide evidence in section 4.4 that this treatment of *meN-* is problematic.

meN- licenses *pro* in its object position, as stated in (48), an assumption that he claims to be independently supported by facts concerning left dislocation, topicalization and binding (Baker 1996).²¹

(48) *pro* is licensed if governed by a verb prefixed with *meN*- or *di*-. (Voskuil 2000: 206)

He further adopts the assumption in (49), which, combined with the previous one, yields the result that the object of the verb marked with *meN*- is always occupied by *pro*.

(49) Pronoun Principle

An empty category is interpreted as *pro* if morphological licensing conditions for *pro* are met, irrespective of derivational history. (Voskuil 2000: 206)

Working with this set of assumptions, he proposes that the A'-extraction of an NP across the *meN*-verb is blocked by the constraint on variables stated in (50). This constraint essentially states that, when syntactic movement does not violate island constraints, the operator-variable relation must be established by movement leaving a trace instead of *pro*.

²¹ However, this assumption is incorrect because direct objects *can* occur with *meN*-, as many of the examples in this chapter have shown.

(50) Constraints on Variables

A variable-operator pair at LF must be derived through movement, unless movement is prohibited by conditions pertaining to the structural distance between the variable and the operator. (Voskuil 2000: 206)

According to this proposal, the extraction of an NP across a verb with *meN-* in a monoclausal context is blocked in the following way. The presence of *meN-* on the verb requires that the *pro* is generated in the object position of the verb. However, since the syntactic movement involved in such a derivation does not violate any island constraints, the operator-variable relation must be established by movement rather than a resumptive-like strategy in Indonesian, leaving a trace, not a *pro*, in the direct object position. The clash of these two requirements gives rise to ungrammaticality. When *meN-* is absent, however, no such clash arises, because the tail of the movement does not have to be a *pro*. The NP-only property may well be derived naturally because the only antecedent for *pro* is an NP operator.

Voskuil's analysis does seem to account for the restriction on the movement of an NP across the *meN*-verb but it is not clear at all whether the three constraints he devised to derive this restriction are independently motivated in Indonesian. Granted that these constraints are

indeed motivated, there are two serious problems with his account. One problem is that, as is clear from his generalization in (51), Voskuil's analysis is designed to block only A'-extraction of an NP across the *meN*-verb.

- (51) *meN*- exclusively blocks A-bar movement of the direct object, *di*- likewise blocks A-bar movement of the genitive agent. (Voskuil 2000: 206)

Voskuil's analysis, therefore, predicts that A-extraction of an NP across such a verb should be possible acceptable, contrary to facts. The other problem comes from one of his assumptions, made elsewhere in Voskuil (2000: sec.3), that LF movement also triggers *meN*-deletion. Saddy (1991) and Cole and Hermon (1998, 2000) provide evidence that LF movement in Indonesian is subject to subadjacency constraints; see chapter 5 of this dissertation for detailed discussion on *wh*-movement in Indonesian. An illustrative example from Indonesian to make this point is in (52).

- (52) * Kamu kira (bahwa)_{[DP} cerita bahwa siapa yang *t_i* **men**-geritik Jon itu] di-jual.
 you think that story that who that AV-criticize John the Pass-sell
 'Who do you think that the story that criticized Jon was sold.' (Saddy 1991: 195)

Both Saddy and Cole and Hermon show that LF movement does not cause *meN*-deletion from the verbs, as shown in examples like (6b), repeated here as (53).

(53) Ali (**mem**)-beritahu kamu tadi [apa_i yang Fatimah (***mem**)-baca *t_i*]?
 Ali **AV**-told you just now what that Fatimah **AV**-read

‘What did Ali tell you just now that Fatimah was reading?’

Voskuil’s assumption above, thus, incorrectly predicts that the higher verb *beritahu* ‘tell’ should have its active voice deleted. Therefore, the ban on A-movement of an NP across *meN*-verbs as well as the interaction of *meN*-deletion with partial *wh*-movement suggest that his analysis does not adequately capture the distribution of the active voice morphology in Indonesian and in Javanese, by extension.

Note that the fact that the “covert” movement does not cause *meN*-/*ng*-deletion in Indonesian/Javanese provides support for the D-feature checking analysis of this deletion. Under the recent Agree-based framework within the Minimalist Program (see Chomsky 2000, 2001, 2004, 2005), there is no LF movement as conceived of in the traditional T-model of grammar. Chomsky proposes instead that overt movement is the byproduct of Agree + EPP. The data in (53) shows that Agree is dissociated from EPP-feature checking,

for otherwise we would make the wrong prediction that the covert movement involved in this example should be able to cause *meN*-deletion in the matrix clause. Thus, since Aldridge's (in press) EPP-analysis has been shown to be problematic, the active voice deletion pattern observed in (53) provides indirect support that what is involved in the active voice deletion in Indonesian/Javanese is D-feature checking.²²

4.4. Fortin's (in press) Anti-Passive Account

Fortin (in press) attempts to derive the syntactically governed distribution of the active voice marker from the assumption that *meN*- is an object-clitic antipassivizing morpheme (Baker 1988); also recall Aldridge's (in press) treatment of *meN*-as anti-passive morpheme. Fortin's argument for the analysis of *meN*- as the object clitic, not merely the active voice/transitivity marker, comes from her observation that *meN*- is obligatory when verbs like *baca* 'read' are used intransitively, namely, without any overt syntactic arguments, as shown by the contrast between (2a) and (2b), repeated here as (54a) and (54b), respectively. The contrast here naturally follows if the prefix *meN*- itself is base-generated as the internal argument of the verb *baca* 'read' and its absence leads to the violation of the Theta-Criterion.

²² My thanks to Heidi Harley (personal communication) for suggesting this line of argument for the proposed analysis.

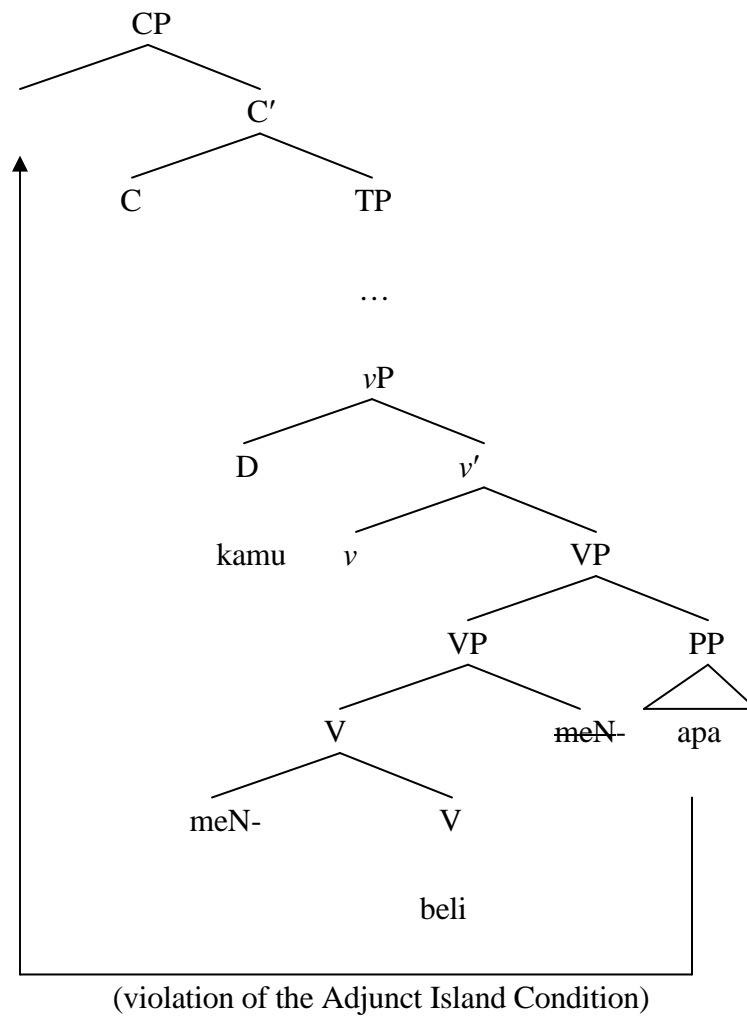
- (54) a. Ali sedang mem-baca. b. * Ali sedang baca.
 Ali Prog AV-read Ali Prog read
 ‘Ali is reading.’ ‘Ali is reading.’

We have seen in the beginning of section 2, however, that this contrast is equally expected under the analysis of *meN*- as the active voice whose appearance is affected by syntactic transitivity. Thus, this contrast itself is mute on whether the morpheme is an example of object-clitic.

According to Fortin's analysis, for clauses with *meN-*, the prefix originates in the θ -position for the internal argument and overly moves to the verb. As is the case of bona fide antipassive constructions generally, such clauses allow an optionally specified object. In this case, the specified object is an oblique adjunct in a VP-adjoined position. In clauses without *meN-*, by contrast, the object NP is a direct object that is base-generated into the θ -position for the internal argument. The two derivations for the sentence in (55), then, will be as in (56) and (57).

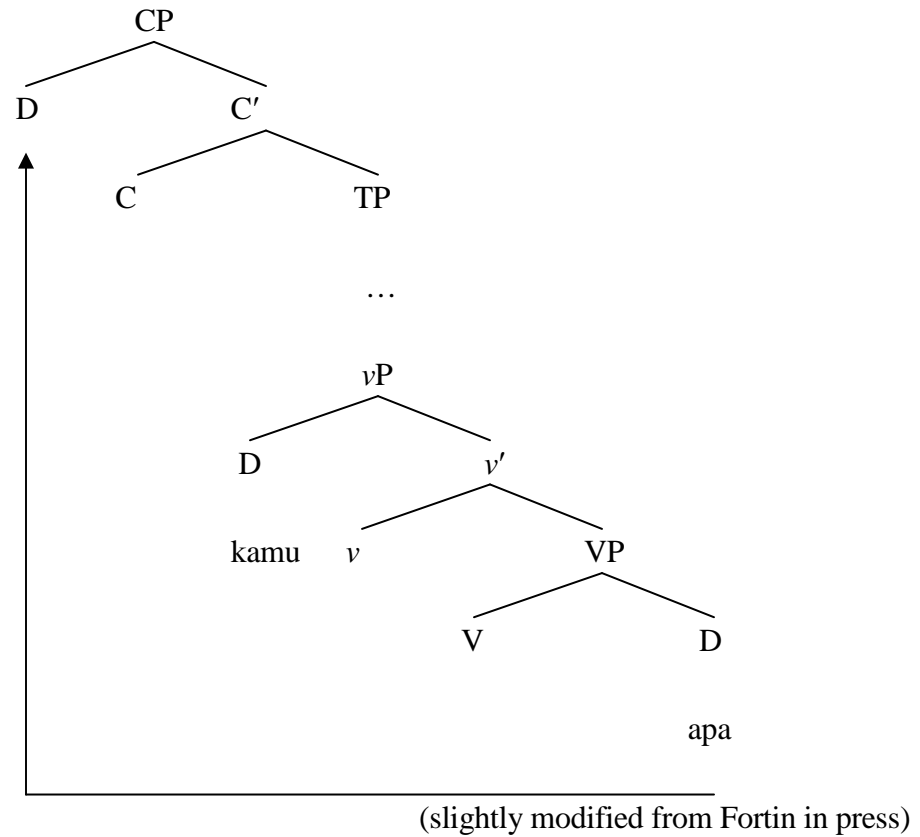
- (55) Apa_i yang kamu (***mem**)-beli *t*_i?
 what that you **AV**-buy
 ‘What did you buy?’ (Fortin in press)

(56) The Derivation of the Example in (54) with *-meN*



(slightly modified from Fortin in press)

(57) The Derivation of the Example in (55) without *-meN*



Since *meN-* is present as in the derivation in (56), the prefix is base-generated in the object position of the verb *beli* ‘buy’ and later moves to the V position. The specified object, in turn, is base-generated as an oblique adjunct in the VP-adjoined position. If *meN-* is absent in the same clause as in the derivation in (57), *apa* ‘what’ occupies the object position of the verb. Fortin argues that the inability of an NP to move across a verb with *meN-* in the derivation in (56) is derived by the Adjunct Island Condition

because the movement of the specified object involves extraction from within an adjunct phrase. By contrast, there is no problem with the derivation of the *meN*-less clause shown in (57) with respect to the movement of *apa* ‘what’ from the thematic position to the specifier of CP.

Fortin’s antipassive account derives the incompatibility of the movement of an NP with the active voice marker from the independently motivated syntactic constraint, the Adjunct Condition. Three considerations indicate, however, that this analysis is inadequate. Firstly, Fortin simply assumes that, in the derivation for *meN*-clauses in (56), Case on the adjunct NP is checked against a null preposition, akin to the null preposition that checks the Case of the agent NP in *di*-passives. Examples of this type of passive are given in (58).

(58) Piring itu sudah **di**-cuci oleh/Ø Pak Ali.

dish that already **PV**-wash by Mister Ali

‘These dishes were already washed by Pak Ali.’

(slightly modified from Fortin in press)

The crucial assumption for Fortin’s analysis to work is then that the oblique adjunct *wh*-phrase in *meN*-questions would have to have a preposition. However, it is not clear

whether we have independent evidence that the *wh*-phrase is contained within the PP in the derivation of *meN*-clauses as in (56). At any rate, evidence from Heavy NP Shift in Indonesian as in (59a), modeled after the Malay examples given in Soh (1998: 302), indicates that the assumption that object DPs in *meN*-marked phrases are PPs seems hard to sustain.

(59) Heavy NP Shift in Indonesian

- a. Saya **meny**-epak [_{NP} bola yang dia campak kepada saya dengan kuat] **kemarin**.

I **AV**-kick ball that s/he throw to me with strength yesterday

‘I kicked the ball that s/he threw to me very hard yesterday.’

- b. Saya **meny**-epak_{t_i} **kemarin** [_{NP} bola yang dia campak kepada saya dengan kuat]_i.

I **AV**-kick yesterday ball that s/he throw to me with strength

‘I kicked yesterday the ball that s/he threw to me very hard.’

Under the standard analysis of Heavy NP Shift as a case of A'-movement, the example in (59a) involves rightward movement of the phonologically heavy NP into a *v*P-adjoined position. That the rightward movement has occurred can be seen from the relative order of the heavy NP with respect to the adverb *kemarin* ‘yesterday’. Fortin’s analysis incorrectly

predicts that this sentence should be ungrammatical with *meN-* because the A'-movement of the heavy NP here would also be blocked by the Adjunct Condition on a par with *meN-* sentences, as in (56). However, this problem does not arise under the proposed phase-theoretic analysis. The heavy NP moves rightward into the adjoined position, not the specifier position of the ν P. This movement does not cause the mandatory active voice deletion because it does not bring the moved NP into the Spec-Head D-feature checking relation in [Spec, ν P].

Secondly, a question remains as to why the whole PP, not the specified object NP, cannot move into the specifier of CP in the derivation in (56). If this movement were possible, then *meN-*deletion would not be obligatorily caused in this derivation. However, (10b) and (46) show that a PP can be fronted in Indonesian. Fortin's analysis, thus, does not seem to go unless this available movement option in Indonesian is blocked by independent principles of syntax in (56).

Finally, Fortin treats *meN-* in Indonesian as the antipassivizing object clitic. In other words, *meN-* clitic-doubles the adjoined PP which gives the full specification of the object. However, this treatment of *meN-* seems problematic because *meN-*+ object NP constructions do not show properties characteristic of bona-fide double-clitic constructions, as found in Romance languages. To mention one case, Franco and Landa (2006) show that

quantifiers and *wh*-words cannot be clitic-doubled due to their inherent quantificational force. This is shown in Spanish examples in (60a, b); see also Suñer (1988), Franco (1993), Cinque (1990), and Rizzi (1997) for much relevant discussion.

(60)a. Juan (*lo_i) quiere a todo el mundo_i. b. ¿A quién_i (*lo_i) viste?

Juan Cl love-3s to everybody to whom Cl saw-2

‘Juan loves everybody.’

‘Who did you see?’

(Franco and Landa 2006: 37)

Thus, Fortin’s treatment of *meN*- predicts that quantifiers and *wh*-expressions should also not co-occur with the prefix. This prediction, however, is falsified by examples of *wh*-in-situ as in (6c), repeated here as (61a), and examples as in (61b) that have quantifiers in direct object positions.

(61) a. Ali (**mem**)-beri Fatimah apa? b. Ali (**mem**)-beri setiap orang.

Ali **AV**-give Fatimah what

Ali **AV**-love every one

‘What did Ali give Fatimah?’

‘Ali loves everyone.’

The three considerations above, therefore, suggest that Fortin's antipassive analysis of the active voice deletion in Indonesian is difficult to sustain.

4.5. *Anti-Agreement-Based Accounts*

Another potential analysis of the *meN-/ng-* deletion in Indonesian and Javanese is to claim that this deletion instantiates a type of agreement known as *anti-agreement* since Ouhalla (1993, 2005), whereby (local) extraction of the subject results in impoverished morphology.²³ Some examples from Berber are given here for illustration.

(62) Anti-agreement in Berber

- | | |
|---|--|
| a. man tamghart ay yzrin Mohand? | b.*man tamghart ay t-zra Mohand? |
| which woman C see (Part) Mohand | whichwoman C 3fs-saw Mohand |
| ‘Which woman saw Mohand?’ | ‘Which woman saw Mohand?’ |

(Ouhalla 1993: 479)

In these examples, the subject *man tamghart* ‘which woman’ is moved into the specifier of CP. When this extraction happens, the agreement gives rise to ungrammaticality, as

²³ I thank Andrew Carnie (personal communication) and Catherine Fortin (personal communication) for suggesting this potential analysis and relevant work cited here.

illustrated in (62b); the verb instead must take its morphologically impoverished form, as in (62a), which traditional grammarians dub the *participle*. Ouhalla (1993) derives anti-agreement from the interaction of the licensing condition on *pro* with the A'-disjointness requirement on the distribution of pronominal elements. It has been often held that *pro* is licensed in the subject position in a language with rich agreement, as evidenced by *pro*-drop languages such as Italian and French. He follows this analysis and assumes that *pro* is licensed only by rich subject agreement. The A'-disjoint requirement, which he adopts from Aoun and Li (1990) and defines as in (63), states that pronominal elements, including *pro*, must be free within the minimal CP that contains them.

(63) The A'-disjointness Requirement

A pronoun must be free in the smallest Complete Functional Complex (CFC) which contains it.

(Ouhalla 1993: 490)

Ouhalla claims that anti-agreement is the result of the strategy to avoid violating the A'-disjointness Requirement. In (62b), regular rich agreement is reflected, which licenses a *pro* in the subject position. This pronominal argument, however, violates the disjointness requirement, giving rise to the ungrammaticality in (62b). The example in (62a) is

grammatical, by contrast, because the impoverished morphology of the verb does not license a *pro* in the subject position, which is occupied by the trace of the *wh*-movement.

Although Ouhalla's (1993) analysis has the potential of deriving the active voice deletion in Indonesian and Javanese as another manifestation of the anti-agreement effect, there are several problems in this potential attempt, to the extent that Ouhalla's analysis is tenable. First, the anti-agreement effect in a language correlates with the availability of rich agreement in that language. However, Indonesian and Javanese are both agreement-less languages in the same sense that languages Chinese and Japanese are. Second, Ouhalla observes that anti-agreement occurs only in the case of subject extraction. I have provided evidence in section 2 that *meN-/ng-* deletion is triggered (in fact, more typically) by non-subject NPs. Also recall that, as discussed in section 4.1, an NP can be extracted from non-subject positions into the specifier of CP without first passing through the subject position [Spec, TP] in Indonesian and Javanese. These two differences indicate that the analysis of active voice deletion as an instantiation of the anti-agreement pattern of the Berber kind is problematic.

4.6. *Case-Based Accounts*

Let us finally consider a potential Case-based account of the restriction on the movement of an NP across an active voice verb marked with *meN-* and *ng-*. This analysis would say that

the moved NP enters into a Spec-Head Case-checking relation with its local ν head, which results in the deletion of the active voice. This analysis correctly predicts that the extraction of the external argument from a clause headed by *meN-/ng-* verbs does not trigger the active voice deletion because that argument checks its nominative Case against T. This analysis also naturally explains the NP-only property of extraction because non-nominal phrases such as adverbials and prepositional arguments do not enter into the Spec-Head Case-feature agreement with their local ν heads.

The Case-based approach is also desirable in one important respect. Other Austronesian languages such as Chamorro, Palauan, Tagalog, and Malagasy exhibit a phenomenon in the form of WH-agreement, the mood alternation, and case agreement, which is structurally quite similar to the active voice deletion in Indonesian and Javanese. Indeed, existing analyses of these phenomena as in Chung (1982, 1994, 1998), Georgopoulos (1985, 1991), and Rackowski and Richards (2005) propose a Case-based explanation of these phenomena. Therefore, it seems that the Case-based analysis not only correctly derive Cole and Hermon's generalization but also allows a unified explanation of the Austronesian syntax-phonology interaction.

However, there are two reasons to believe that the Case-based analysis is incorrect for the *meN-/ng-* deletion in Indonesian and Javanese. First, if the Case-based analysis could be

extended to Indonesian and Javanese, it would predict that a single NP should receive multiple Accusative Cases on its way toward the final landing site in multi-clausal environments since it triggers *meN*-deletion on all intervening verbs on its movement path. This is a rather undesirable outcome. No languages have been found in which a single NP receives more than two structural Cases. Bejar and Masam (1999) document several instances of multiple Case checking in languages such as Niuean, Norwegian, and Latin and develop a late-insertion analysis of Case assignment constrained by markedness, so it is possible that a single NP could receive multiple cases.²⁴ However, all the cases discussed there are restricted to those in which a single NP receives either two different values for structural Case [Nominative and Accusative] or one structural Case [Nominative/Accusative] and one inherent Case [Oblique]. However, if the *meN/ng-* deletion were the reflex of the accusative Case-checking in Indonesian and Javanese, then a single A-bar moved phrase in a multi-clausal environment would receive a series of the same values for Case, in principle, the number of Cases that corresponds to the number of transitive verbs present in multi-clausal environments. However, this type of multiple case checking has always been assumed to be impossible in natural language, and as long as a better alternative account is possible, should be dispreferred.

²⁴ As suggested by Simin Karimi (personal communication). See also Boeckx and Hornstein (2006) for an analysis of the multiple case agreement in Icelandic control constructions.

Second, this multiple Case-checking account is suspect on theoretical grounds. With the exception of cases as discussed in Bejar and Masam (1999), it seems correct to think that a single DP receiving a single Case is the unmarked pattern in natural language syntax. The modern theory of Case assignment/checking/evaluation in the Government-and-Binding approach (Chomsky 1981, 1986a, b) or the Minimalist Program (Chomsky 1995, 2000, 2001, 2004, 2005) is indeed so constructed as to exclude a single DP from receiving more than one Case. Within the Minimalist Program outlined in Chomsky (1995), for example, the alleged Case-driven movement, as applied to long-distance extraction of *wh*-phrases in Indonesian and Javanese, would violate the Last Resort Condition/Greed, which effectively blocks movement from a Case position to another Case position, as shown in examples such as (64b).

(64) a. John_i seems to Mary [_{TP} *t_i* to be ill].

b.* John_i seems to Mary [_{CP} that *t_i* is ill].

This theoretical viewpoint, therefore, casts further doubts on the Case-based analysis of *meN-/ng-* deletion. Therefore, I conclude that the Case-based analysis of the distribution of the active voice deletion is not tenable for Indonesian or Javanese.

5. Conclusions

Let us summarize the present chapter. I have started by reviewing Cole and Hermon's (1998) generalization that, in Malay/Indonesian, the movement of an NP across a verb triggers the active voice deletion from the verb. I have shown that this generalization also adequately characterizes the distribution of the active voice morphology in Javanese and Madurese. I have argued, following Kayne's (1989) analysis of French participle agreement, that this generalization is naturally explained if the obligatory deletion of the active voice morphology is a reflex at the syntax-external sound component of the syntactic Spec-Head D-Feature Checking between the moved NP and its local ν head. I have shown that the external argument does not cause this deletion by introducing the crucial distinction between External Merge and Internal Merge recently proposed by Chomsky (2004, 2005). The current analysis naturally predicts that the movement of a non-nominal phrase such as adverbials and prepositional arguments does not trigger the deletion due to their lack of nominal D-features.

I have compared the proposed analysis with several existing/potential alternative accounts of the same phenomena as in Keenan (1972)/Keenan and Comrie's (1977) NP-Accessibility account, Soh's (1998) Relativized Minimality account, Fortin's (in press) antipassive account, the anti-agreement-based accounts, and the Case-based account. I have shown that each of these analyses has a particular set of empirical and conceptual problems that are

naturally resolved under the current phase-theoretic analysis. To the extent that this analysis is correct, the syntactically governed distribution of the active voice morphology in Indonesian and Javanese provides strong support for the role of the ν P phase at the syntax-phonology interface. This result is important because evidence has been deemed hard to find for the successive-cyclic movement via the edge of ν P.

5.1. *The Derivational Nature of the Syntax-Phonology Interface*

In this subsection, I discuss a theoretical implication of the proposed analysis of the distribution of the active voice morphology in Indonesian and Javanese for the syntax-phonology interface. The proposed analysis provides clear evidence that the syntax-phonology interaction is fundamentally derivational, meaning that the interpretation of syntactic objects at the phonological component occurs in a piece-meal fashion rather than by examining the final output of the syntactic derivation.

The idea that syntax interacts with phonology in a dynamic manner has proven a fruitful line of inquiry since the seminal work of Bresnan (1971a, b), who shows that nuclear stress assignment is cyclic. See also Legate (2003), Arregi (2003), Kahnemuyipour (2004) and Adger (2007) for recent phase-theoretic implementations of Bresnan's analysis which was couched within the Standard Theory. I have also shown elsewhere (Sato in

press c; see also Carnie in preparation) that the domain for certain phonological rule applications such as Taiwanese tone sandhi, Giyak lenition, Kinyambo high tone deletion, and Irish/Welsh consonant mutation, are isomorphic to domains demarcated by a derivational model of syntax like Phase Theory. The proposed account of the active voice deletion in Indonesian and Javanese, therefore, provides further evidence that the syntactic computation is tightly intertwined with phonology in a derivational manner.

5.2. Minimalist Interfaces

I conclude by noting what the results achieved in this chapter tell us about the thesis of “minimalist interfaces” outlined in the introductory chapter. Part of the idea behind this thesis is stating that the syntax-external interpretive components follow a parametrically defined path curved by the universal syntactic computation and assign whatever interpretation it can to a syntactic object sent to the components. The results in this chapter give substance for this notion of minimalist interfaces. Indonesian and Javanese happen to signal successive cyclic movement via the deletion of the active voice on the *v* head. This, of course, does not have to be the only way to indicate the universal locality principle in syntactic computation. Indeed, many other languages have different ways of accomplishing the same task of successive cyclic movement via the two phase edges at the interface. Irish (McCloskey 1979, 2001; Duffield 1995) expresses the locality of movement via the

edge of CP and ν P phases through complementizer agreement and the $ag \rightarrow a^L$ rule in progressives, respectively.²⁵ Chamorro (Chung 1982, 1994, 1998), Palauan (Georgopoulos 1985, 1991), Tagalog (Rackowski and Richards 2005), and Malagasy (Pearson 2001, 2005) exhibit movement to the ν P edge via Case agreement, a special mood alternation, and voice agreement, respectively. French marks the movement of an NP via [Spec, ν P] by participial agreement (Kayne 1989). Berber (Ouhalla 1993, 2005) marks the intermediate site of the movement in [Spec, CP] by the use of special impoverished agreement morphology. Kikuyu (Clements 1984; Clements *et al.* 1983) signals the movement via the ν P edge in the form of the loss of tonal downstep from a verb crossed by the movement. Innu-Aimûn (Branigan and MacKenzie 2002), and Kilega (den Dikken 2001) show evidence for the local movement through intermediate ν Ps in the form of Wh-agreement on the verbs, as in Chamorro. Hausa and Moore express the occurrence of successive cyclic movement via the ν P edge through the two different forms of modals and affixes. Afrikaans (Du Plessis 1977) marks successive cyclic movement through [Spec, CP] by prepositional copying in intermediate specifiers of CPs. Spanish (Torrego 1984) and French (Kayne and Pollock 1979) provide evidence for intermediate [Spec, CP]s via obligatory stylistic inversion in clauses out of which movement has occurred. *Wh*-copying in languages such as Frisian, Dutch, Afrikaans, German, and Child English (Felser 2004; Thornton 1995) as well as the existence of partial *wh*-

²⁵ I thank Andrew Carnie (personal communication) for bringing my attention to the $ag \rightarrow a^L$ rule.

movement in languages like German, Hungarian, and Indonesian (McDaniel 1989; Horvath 1997; Saddy 1991; Cole and Hermon 1998, 2000) provide morphological evidence that *wh*-phrases move through the edge of CP phases. Even languages with impoverished morphology such as English, which have no morphological ways to signal the phase-based locality of syntactic movement, still have alternative methods of reflecting the locality of movement in the other semantic interface via reconstruction, parasitic gaps, and antecedent-contained deletion (Lebeaux 1988; Fox 1998; Legate 2003).

All these phenomena are but several of the options within the range set by Universal Grammar the phonological interface can choose to reflect the fundamental principle of locality in syntactic computation. This is highly reminiscent of the remark below, made by Chomsky (1995), citing Otto Jespersen.

Jespersen held further that it is only “with respect to syntax” that we expect “that there must be something in common to all human speech”; there can be a “universal (or general) grammar,” hence a perhaps a far-reaching account of the initial stage of the language faculty in this domain, though “no one ever dreamed of a universal morphology.” (Chomsky 1995: 3)

The existence of the apparently diverse, outwardly different surface manifestations that nonetheless signal the single core property of syntax can be construed as powerful support for the idea of “minimalist interfaces” that the working of the sound-related interface component is fundamentally interpretive and subservient to the universal law of syntax in that it has a handful of domain-specific operations, some of them discussed above, to reflect the locality of syntactic derivation. I discuss further ramifications of this notion of minimalist interface in chapter 7.

CHAPTER 3 PREPOSITION STRANDING UNDER SLUICING AND REPAIR

BY DELETION: WHY IS INDONESIAN (NOT) SO SPECIAL¹

1. Introduction

In this chapter, I explore the issue of the syntax-phonology interface with a case study in sluicing in Indonesian to see the extent to which the thesis of minimalist interface proposed in the first chapter holds. I demonstrate that detailed investigation of the sluicing construction in this language lends further credence to the idea embedded within the proposed thesis that the syntax-external phonological component avails itself of whatever domain-operation it can to remedy certain “imperfections” created by the functionally blind computational process within the curve parametrically defined by the syntax.

My starting observation in this chapter is that sluicing in Indonesian behaves differently from other languages with respect to P-stranding. In favor of his movement and TP deletion approach to sluicing constructions, Merchant (2001) proposes, based on his survey of 24 languages, the P(reposition)-Stranding Generalization, namely, that P-stranding under sluicing is possible only in those languages that independently allow P-stranding under

¹ The core observation in this chapter was presented at the Linguistic Colloquium Datablitz held at the University of Arizona (April 2007) and later published in Sato (2007). An earlier draft of this chapter was presented at the 34th Annual Meeting of the Berkeley Linguistics Society (BLS 34) held at the University of California, Berkeley (February 2008) and will appear in Sato (forthcoming a).

regular *wh*-questions. I show that Indonesian presents itself as a clear counterexample to this cross-linguistic generalization to lay the groundwork for the investigation of the syntax-phonology interface. One of the most important findings in this chapter lies in my demonstration that Indonesian constitutes the first bona fide counterexample to the generalization stated above. Although several languages such as Brazilian Portuguese (Almeida and Yoshida 2007), Polish (Szczegielmiak 2006), Mandarin Chinese (Wang 2006), Malagasy (Potsdam 2003), and Serbo Croatian (Stepanović 2008) that appear to contract the P-Stranding Generalization, I point out that none of these languages poses a genuine counterexample to the generalization, unlike Indonesian. In light of this observation, I propose a new analysis of the typology of P-stranding under sluicing which takes seriously the role of prepositions across languages with respect to the syntax-phonology interface. Specifically, I argue that the apparently atypical P-stranding pattern observed in Indonesian naturally falls out from independently motivated assumptions concerning the percolation of the [+wh] feature onto the PP, D-to-P incorporation, and the minimalist constraint on movement/attraction. The core idea behind the proposed analysis is that certain violations created by syntactic derivation can be remedied by syntax-external operations such as deletion at the phonological component, an idea that goes back to Ross' (1969) global evaluation of the syntactic derivation and has been developed in recent minimalist inquiries

as in Merchant (2001), Boeckx and Lasnik (2006), Lasnik (1999, 2001, 2005, 2007), and others. The proposed analysis is summarized in Table 1.²

Table 1: The Parametric-Theory of P-Stranding at the Syntax-Phonology Interface

Parameters Languages	[+wh] feature percolation from the nominal to PP	D-to-P incorporation in the syntax
English	OPTIONAL	NO
Indonesian	OBLIGATORY	NO
French	OBLIGATORY	YES

I demonstrate that the proposed analysis correctly captures the three-way contrast between English, French, and Indonesian with respect to the P-stranding pattern under sluicing and *wh*-movement. I also show that the proposed analysis also makes a correction prediction that not only TP deletion but deletion of smaller constituents saves the otherwise illicit syntactic configuration through examination of pseudogapping in Indonesian and French.

The core observation in this chapter that Indonesian presents a challenge to the P-Stranding Generalization has been independently made by Fortin (2007a), around the

² I discuss the fourth logically possible parameter setting, namely in section 4.2, where I argue that the non-existence of languages with this setting is fully predicted due to the incompatibility of obligatory D-to-P incorporation with optional *wh*-feature percolation.

same time I started working on sluicing in this language, a preliminary result of which was reported in Sato (2007). We have not been aware of each other's work until quite recently, when her dissertation (Fortin 2007b) became available to me through our e-mail correspondence. As we will see below, this chapter owes a great deal to her dissertation, and reaches the same observation that Indonesian is a real counterexample to the P-Stranding Generalization. However, the analysis developed in this chapter is quite different from her analysis of the P-stranding pattern, each motivated by two distinct sets of theoretical assumptions within the minimalist framework. Nonetheless, I hereby would like to acknowledge that her work represents the first comprehensive description and analysis of Indonesian sluicing and that the current chapter also draws on her data and analysis.

The organization of this chapter is as follows. In the next section, I provide an overview of the analysis of sluicing constructions proposed by Merchant (2001) as the result of the syntactic movement of the *wh*-remnant, followed by TP deletion at the PF component, and introduce one of his core generalizations in support of this analysis, the P-Stranding Generalization, namely, that a language allows P-stranding/omission under sluicing if and only if it allows P-stranding under *wh*-movement. Several languages have been reported in the literature that appear to be counterexamples to this generalization in that P-stranding is

impossible in sluicing but permitted under *wh*-questions in those languages. These languages include Brazilian Portuguese (Almeida and Yoshida 2006), Mandarin Chinese (Wang 2006), Polish (Szczegielniak 2006), Serbo-Croatian (Stjepanovic 2007), and Malagasy (Potsdam 2003). I review the derivation of sluicing constructions in these languages with reference to the works cited here and conclude that none of these languages refutes the P-Stranding Generalization because *wh*-questions in these languages involve alternative sources of syntactic derivation such as clefting (Brazilian Portuguese and Polish), resumption (Mandarin Chinese), pseudoclefting (Malagasy), and P-drop at PF (Serbo-Croatian), that do not involve *wh*-movement of the English kind.

In section 3, I turn to a detailed investigation of the syntax of sluicing constructions in Indonesian. The sole purpose of this section is to establish that sluicing in this language presents the first real counterexample to the P-Stranding Generalization. I achieve this goal as follows. I first review Fortin's (2007b) results of applying some of the operational tests developed by Merchant (2001) to distinguish genuine sluicing (derived by *wh*-movement + TP deletion) and pseudosluicing (derived by clefting + ellipsis), which suggest that Indonesian sluicing behaves more like English sluicing rather than like English clefting. I conclude that this result, though indicative of the nature of sluicing in Indonesian, is not conclusive, as the operational tests were developed on the basis of English and related

languages. Thus, I provide language-internal evidence based on the distribution of the question marker *-kah* (Fortin 2007b) and the obligatory absence of the complementizer *yang* with PP remnants that show that sluicing in this language is derived by true *wh*-movement plus TP deletion, rejecting alternative analyses of *wh*-questions as reduced clefts/headless relative clauses, as proposed by Cheng (1991) and Cole et al. (to appear).

Having established that Indonesian sluicing presents itself as the first genuine counterexample to the P-Stranding Generalization, in section 4, I propose a new analysis of the typology of P-stranding patterns that draws on independently motivated assumptions concerning the percolation of the [+wh] feature onto the PP (Chomsky 1972; Stepanov 2001) and the D-to-P incorporation in Romance (Law 1998, 2006; van Riemsdijk 1998). The core idea pursued here is that certain violations such as the failure of feature percolation can be ameliorated as long as offending configurations created in syntax can be nullified via deletion at the syntax-external phonological component. I show that the proposed analysis straightforwardly derives the three-way contrast between English, French, and Indonesian with respect to P-stranding under *wh*-questions and sluicing. The proposed analysis further predicts that the size of the syntactic object deleted at PF does not matter for the purposes of amelioration at the interface as long as the offending

configuration is eliminated at this level. I show that this prediction is indeed borne out by Indonesian and French pseudogapping constructions.

In section 5, I compare the proposed interface-based approach to Indonesian sluicing with the most recent analysis of it by Fortin (2007b) in terms of LF Copying and Long-Distance Agree. I show that, although her analysis correctly captures the apparently atypical P-stranding pattern in Indonesian, it encounters serious empirical and conceptual problems with regards to the case checking of the *wh*-remnant, the restriction on [+wh] feature percolation, and case-matching effects. I also briefly address a number of issues concerning the derivation and representation of syntax and its interface with the PF component.

In section 6, I provide the summary of this chapter and point out several important implications of the proposed analysis for the proper theory of syntax-phonology correspondence. Among them, I point out that the current investigation provides strong confirmation to the idea in the minimalist interface that the syntax-external phonological component does whatever it can to save otherwise illegitimate violations created by universal syntactic mechanisms but only does so within the restricted range of options parametrically available in a particular language.

2. Sluicing Constructions

Sluicing is one of the most investigated constructions in recent syntactic literature. It was identified in seminal work by Ross (1969). “Sluicing” refers to the elided part of sentences such as *Mary bought something but I don’t know what []*. Research since Ross (1969) has concentrated on identifying the external and internal syntax of the elided part as well as syntactic/semantic conditions that license it. The purpose of this section is two-fold. First, I review the most recent analysis of this construction proposed by Merchant (2001) and introduce one of his core form-identity generalizations concerning P-stranding under sluicing known as *the P-Stranding Generalization*. This generalization states that P-stranding/omission in a language is possible only if that language independently allows P-stranding under regular *wh*-movement. Second, I review several languages that potentially undermine this generalization and show that none of these languages is a genuine counterexample to this generalization because alternative syntactic and/or phonological strategies for P-stranding are available to derive sluicing constructions other than syntactic *wh*-movement.

2.1. *Merchant's (2001) Analysis of Sluicing and the P-Stranding Generalization*

Drawing on the data and analysis presented in Ross (1969), Merchant (2001) argues that sluicing constructions as in (1a) are the product of syntactic *wh*-movement of the remnant *who*, followed by the deletion/dephoneticization of the TP constituent, as shown in (1b).

- (1) a. Somebody just left. - Guess who.
 b. Somebody just left. - Guess [_{CP} *who*_i ... [_{TP} ~~*t*_i just left~~]]

Merchant adduces a wide variety of syntactic and morphological effects such as case-matching, number agreement, and so on, many of them mentioned in Ross (1969), to support his movement plus TP deletion analysis. One of the most convincing arguments in its favor comes from what he calls the Preposition-Stranding Generalization stated as in (2).

(2) Preposition-Stranding Generalization (Merchant 2001: 92)

A language *L* will allow preposition stranding under sluicing iff *L* allows preposition stranding under regular *wh*-movement.

This generalization states that, if a preposition can be omitted under sluicing in a particular language, that language should independently allow P-stranding under regular *wh*-movement. The logic behind this generalization is clear. Under Merchant's analysis, sluicing is derived by regular *wh*-movement plus TP deletion. Therefore, the availability of P-stranding under sluicing means that the same option should be independently available under regular *wh*-movement. Merchant (2001) surveys the P-stranding pattern both under *wh*-movement and sluicing in 24 languages to show that this generalization is crosslinguistically correct for these languages. As is well-known, English allows P-stranding both under *wh*-movement and sluicing, as shown in examples as in (3a, c), in conformity with the P-Stranding Generalization. Note that the preposition can also be pied-piped along into the specifier of the CP, as illustrated in (3b).

- (3)a. Who was he talking with?
- b. With whom was he talking?
- c. Peter was talking with someone, but I don't know (with) who.

(Merchant 2001: 92)

This pattern falls out from Merchant's analysis because the preposition-less variant of the sluicing in (3b) is derived when the preposition *with* is stranded within the TP, as in (3a). On the other hand, Romance languages such as French are strongly non-P-stranding languages, as shown by the contrast between (4a) and (4b). Thus, French does not allow omission of the preposition under sluicing, as shown in (4c), a pattern that is also predicted by Merchant's analysis. (The examples in (4a, c) are from Merchant 2001: 98.)³

(4)a. * Qui est-ce qu' elle l'a offert à?

who Q she it-has offered to

'Whom has she offered it to?'

b. À qui l'a-t-elle offert ?

to whom it-has-she offered

'To whom has she offered it?'

c. Anne l'a offert à quelqu'un, mais je ne sais pas *(à) qui.

Anne it-has offered to someone but I Neg know not to whom

'Anne has offered it to someone, but I don't know (to) whom.'

³ Some speakers readily accept the P-less sluice in (4c). Thanks to Myriam Bouveret (personal communication) for pointing this out.

The type of language that is not predicted by Merchant's generalization, therefore, have syntactic *wh*-movement, disallow P-stranding under this context, but nonetheless allow P-stranding/omission under sluicing. In section 3, I provide arguments based on the examples as in (5a-c) that Indonesian is precisely of this type.

(5)a. * Siapa yang kamu ber-dansa dengan?

who that you Vz-dance with

'Whom did you dance with?'

b. Dengan siapa kamu ber-dansa?

with who you Vz-dance

'With whom did you dance?'

c. Saya ingat Hasan ber-dansa dengan seseorang, tapi saya tidak

I remember Hasan Vz-dance with someone but I Neg

tahu (dengan) siapa.

know (with) who

'I remember Hasan danced with someone, but I don't know (with) whom.'

The contrast between (5a) and (5b) shows that Indonesian does not allow P-stranding under *wh*-questions, as in French. Surprisingly, however, the grammaticality of the P-less sluice in (5c) illustrates that the preposition *can* be deleted under sluicing. To the extent that the underlying syntactic source for sluices as in (5c) involves syntactic *wh*-movement, the observed pattern of P-stranding in (5a-c) presents a counterexample to the P-Stranding Generalization.

Before going into detailed discussion of the precise derivation of sluicing constructions in Indonesian, however, I review several other languages that have been reported in the literature to disallow P-stranding under *wh*-questions but allow this option under sluicing. I show that this P-stranding pattern in those languages does not undermine the P-Stranding Generalization because *wh*-questions in those languages have alternative derivational sources such as clefting (Brazilian Portuguese, Polish), pseudoclefting (Malagasy), resumption (Mandarin Chinese), or the post-syntactic P-drop (Serbo-Croatian).

2.2. *P-Stranding under Sluicing in Languages without P-Stranding*

In this subsection, I discuss the P-stranding pattern in Brazilian Portuguese, Polish, Mandarin Chinese, Malagasy, and Serbo-Croatian as examples of languages that appear to contradict the P-Stranding Generalization. Several works published on this issue show that none of

these languages undermines the relevant generalization because sluicing in these languages involve some syntactic or phonological operation other than regular *wh*-movement, as in English. I also discuss whether any of the analyses developed in the literature for these languages can be transported to the Indonesian P-stranding pattern and show that none of them works for Indonesian. The survey here is far from comprehensive, as many other languages such as Persian, Russian, and Finnish also seem to behave in a way not predicted by the generalization.⁴ It is by now widely acknowledged that the generalization does not hold across the board either across languages or within a single language.

2.2.1. Brazilian Portuguese

Brazilian Portuguese is perhaps the most well-known language that seems to directly contradict the P-Stranding Generalization. Almeida and Yoshida (2007) points out example as in (6a-c).

⁴ I thank Simin Karimi (personal communication) and Mans Hulden (personal communication) for bringing Persian and Finnish to my attention. See Szczegielniak (2006) and Hartman (2005) for some data concerning German and Finnish. However, whether or not the languages mentioned here are true counterexamples to the P-Stranding Generalization is unclear and requires detailed work that goes beyond the scope of this chapter.

(6)a. Com quem_i que a Maria dançou t_i ?

with who that the Maria danced

‘With whom did Maria dance?’

b.* Quem_i que a Maria dançou com t_i ?

who that the Maria danced with

‘Who did Maria dance with?’

c. A Maria dançou com alguém, mas eu não lembro quem_i a Maria dançou com t_i.

the Maria danced with someone but I Neg remember who the Maria danced with

‘Maria danced with someone, but I don’t remember who.’

(Almeida and Yoshida 2007: 350)

Applying those tests applicable to Brazilian Portuguese that were developed by Merchant (2001) to distinguish between genuine sluicing and elliptical clefts (i.e. prosody, aggressively non-D-linked *wh*-phrases, ‘mention-some’ modification, and ‘else’ modification, Almeida and Yoshida present arguments to show that the example in (6c) is not an instance of elliptical cleft (as argued for in Erteschik-Shir 1977) but a genuine sluicing. In a recent rebuttal, however, Rodriguez et al. (2007) argue that the P-stranding pattern observed in Brazilian Portuguese does not constitute a genuine counterexample to the P-Stranding Generalization by showing

that there are two distinct sources of sluicing in Romance, one *wh*-movement plus TP deletion as in Merchant's (2001) theory, and the other clefting plus TP deletion, as in Merchant's (1998) analysis of Japanese-type "sluicing" that does not involve P-stranding in the first place.⁵ Their cleft analysis is illustrated here with the derivation of the P-less sluice in Brazilian Portuguese in (7a) that is given in (7b).

- (7) O João falou com uma menina...
- the João talked with a girl
- a. mas eu não sei qual é a menina com a qual João falou
- but I Neg know which is the girl with the which João talked
- b. mas eu não sei [CP qual [IP é [DP a menina [RC com a qual João falou]]]
- butI Neg know which is the girl with the which João talked
- Lit. 'João talked with a girl, but I don't know which is the girl with which João talked.'

(Rodrigues et al. 2007: 4)

The derivation in (7b) involves a cleft containing a specificational copular sentence; the copular verb *é* is followed by a DP that contains a restrictive relative clause structure. Since the

⁵ I thank Heidi Harley (personal communication) for directing my attention to Rodrigues et al. (2007).

pivot of the cleft derivation here is not introduced by the preposition, deletion of all materials including the verb and the predicate has the same effect as P-stranding, as commonly understood. Rodrigues et al. provide evidence internal to Brazilian Portuguese and Spanish that the tests Almeida and Yoshida used against a pseudosubjunctive analysis for this language are not applicable to these languages due to the bleached semantics of clefts, the inadequacy of the D-linking test, and so on, which I will not review here.

Rodrigues et al.'s (to appear) cleft analysis for Spanish and Brazilian Portuguese is quite similar to the headless relative clause analysis proposed by Cole, et al. (to appear) for *wh*-questions in Malay/Indonesian. In section 3, I provide evidence that their cleft analysis is not correct for Indonesian.

2.2.2. Polish

Merchant provides data as in (8a, b) that Polish conforms to his P-Stranding Generalization.

- (8)a. * Kim rozmawiała Anna z?
 who spoke Anna with
 ‘Who did Anna speak with?’

b. * Anna rozmawiała z kimś ale nie wiem kim.

Anna spoke with someone but not I.know who

‘Anna spoke with someone, but I don’t know who.’ (Merchant 2001: 96)

Szczegielniak (2006) argues, however, that the facts in Polish are more complicated than what Merchant reports.⁶ Specifically, he provides examples as in (9a-c) to show that the omission of the preposition is possible with sluicing only with D-linked *wh*-phrases.

(9)a. * Którym₁ Anna tańczyła z t₁ mężczyzną.?

which Anna danced with man

‘Which man did Ann dance with?’

b. * [Którym mężczyzną]₁ Anna tańczyła z t₁?

which man Anna danced with

‘Which man did Ann dance with?’

⁶ Thanks to Andrew Carnie (personal communication) for directing my attention to Szczegielniak (2006).

c. Anna tańczyła z jednym mężczyzną, ale nie wiem (z) którym.

Anna danced with one man but Neg know with which

‘Ann danced with one man, but I don’t know which.’

(Szczegielniak 2006: 3)

The ungrammaticality of the examples in (9a, b) under *wh*-questions shows that Polish does not allow stranding of prepositions such as *z* ‘with’. The fact that the same preposition can be deleted in the sluicing counterpart as in (9c), therefore, poses a problem for Merchant’s generalization. Szczegielniak proposes, however, that, in cases of sluicing with D-linked *wh*-phrases, a non-*wh*-movement alternative is available for languages such as Polish and Russian in the form of a cleft construction. The underlying cleft structure for the P-less sluice in (9c) is given in (10a). Note that the deleted portion of (9c) is itself grammatical, as shown by the cleft *wh*-question in (10b).

(10)a. Anna tańczyła z jednym mężczyzną ale nie wiem [którym]₂ ~~to~~

Anna danced with one man but not know which it

~~[z t_2 ——— mężczyzną]₁ ——— (ona) ——— tańczyła t_1]~~

with man (she) danced

b. którym₂ to [z t_2 mężczyzną]₁ (ona) tańczyła t_1 .

which it with man (she) danced

‘With which man was it that she danced?’

(slightly modified from Szczegielniak 2006: 3)

In the derivation in (10a), the PP containing the *wh*-phrase undergoes movement, which is followed by clefting of the *wh*-phrase itself. Crucially, then, this derivation does not involve the P-stranding configuration in the sense that it does not extract the entire DP from the complement position of the preposition. This analysis also correctly accounts for why P-less sluices are ungrammatical with non-D-linked *wh*-phrases as in (8c); there is no possible cleft source for such *wh*-phrases, as shown by the possible but non-convergent derivation given in (11a). Note that the deleted portion of these derivations is itself ungrammatical in Polish, as shown in (11b).

(11)a. * Anna tańczyła z jednym mężczyzną ale nie wiem kim₂ tø

Ann danced with one man but Neg know who it

[z — t₂]₁ — ona — tańczyła — t₁.

with she danced

‘Ann danced with one man but I don’t know who it was that she danced with.’

b. * Kim₂ to [z t₂]₁ ona tańczyła t₁.

who it with she danced

‘Who was it that she danced with?’

(Szczegielniak 2006: 4)

Based on these data, Szczegielniak concludes that the apparently problematic P-omission pattern under sluicing naturally makes sense because only D-linked *wh*-phrases have cleft sources that do not involve extraction of them from the PP. Szczegielniak extends this analysis to German sluicing which exhibits the same behavior with Polish in terms of P-stranding under sluicing.

Like Rodriguez et al’s (2006) analysis for sluicing in Brazilian Portuguese, the clefting analysis could also account for the P-stranding facts in Indonesian, which is superficially similar to that in Polish in terms of P-stranding under *wh*-questions and sluicing. This consideration is important since several researchers in Indonesian as in Cheng (1991) and Cole

et al. (to appear) have independently proposed that *wh*-questions in this language are a form of cleft or headless relative clause. I show in section 3.2, however, that the cleft analysis is not the right approach to sluicing in Indonesian.

2.2.3. Mandarin Chinese

Wang (2006) reports that Mandarin Chinese presents a counterexample to the P-Stranding Generalization as a non-P-stranding language that nonetheless allows omission of the preposition under sluicing. Consider examples in (12a, b).

(12)a.* (shi) [na-ge ren]_i Lisi gen t_i zai shuohua?

Foc/Cop which-Class person Lisi with Prog talk

‘Which one is Lisi talking with?’

b. Lisi gen mou-ge ren qu wan, dan wo bu zhidao shi (gen) shei.

Lisi with certain-Class person go play but I Neg know Foc/Cop with who

‘Lisi has a trip with a certain person, but I don’t know who.’

(Wang 2006: 9, 10)

(12a) shows that Mandarin Chinese does not allow P-stranding under *wh*-questions, but (12b) shows that the P-less sluice is grammatical, contrary to the P-Stranding Generalization. Based on her observation that all languages that show this P-stranding pattern are reported to make an extensive use of a resumptive pronoun strategy, she argues that (12b) is grammatical because it involves the generation of a resumptive pronoun that follows regular *wh*-movement. This is evidenced by the acceptability of the resumptive strategy in *wh*-questions and sluicing constructions, as shown in (13a) and (13b), respectively.

- (13)a. na-ge ren_i Lisi hen zihuan (ta-*t_i*)?
 which-Class person Lisi very like him
- b. keshi wo bu zhidao na-ge ren_i <_{TP} Lisi gen ta-*t_i* qu kan dianying>
 but I Neg know which-Class person Lisi with him go see movies
 Lit. ‘... but I don’t know which person (did) Lisi go to the movies with him.’

(Wang 2006: 10, 11)

To the extent that her resumptive analysis of sluicing in Mandarin Chinese holds, the P-stranding pattern in this language does not contradict the P-Stranding Generalization, as the syntax of this construction does not involve regular movement of a question operator

into [Spec, CP]. The non-movement analysis of sluicing in Mandarin Chinese is also independently motivated by the fact that this language is a so-called *wh*-in-situ language. I do not examine the validity of Wang's resumptive analysis for Mandarin Chinese here as the sole purpose here is to see why Mandarin Chinese does not provide a conclusive counterexample to the P-Stranding Generalization. However, I do wish to note here that the resumptive analysis cannot be extended to accommodate the Indonesian P-stranding pattern. If Wang's analysis were extendable to Indonesian, we predict that this language also would make use of the resumptive pronoun strategy as in Mandarin Chinese. This prediction is clearly incorrect, since Indonesian does not use resumptive pronouns even under contexts where they would ameliorate island violations (Sells 1984), as shown in examples in (14a-c) provided by Fortin (2007b: 71).

(14)a.* Apa yang Ali jadi terlalu gemuk [CP karena dia makan *apa/pro*]?
 what that Ali be too chubby because he/she eat what

‘What did Ali get fat because he ate?’

b. * Apa yang Ali jadi terlalu gemuk [CP karena itu dimakan+nya]?
 what that Ali be too chubby because it Pas-eat+he/she

‘What did Ali get fat because it was eaten by him?’

- c. * Apa yang Ali jadi terlalu gemuk [CP karena dia makan+nya]?
 what that Ali be too chubby because he/she eat+he/she
 ‘What did Ali get fat because he ate it?’

(14a) shows that *wh*-movement in Indonesian shows the standard adjunct island effect; see chapter 5 for details on the syntax and semantics of *wh*-questions. (14b, c) are attempts to insert a resumptive pronoun via the clitic pronoun *-nya*. The fact that there is no way to save the adjunct island violation by resumptive suggests that the resumptive analysis motivated by Wang on Mandarin Chinese data cannot be the right approach to the superficially similar P-stranding pattern in Indonesian.

2.2.4. Malagasy

Malagasy was the first language reported in the literature whose P-stranding pattern under sluicing poses a potential counterexample to the P-Stranding Generalization. Potsdam (2003) shows that P-stranding is prohibited under *wh*-movement but is allowed under sluicing, as (15a, b) illustrate.

(15)a.* inona ny trano no mitoetra amina i Raso?

what the house Foc live. Act in Raso

‘Which house does Raso live in?’

b. mitoetra amin’ ny trano i Raso fa adinoko hoe inona ilay

live.Act in the house Raso but forget.1sg HOE what that

trano ~~no~~ ~~mitoetra~~ ~~amin’~~ ~~*t*_{which house}~~ ~~i~~ ~~Raso~~

house Foc live.Act in Raso

‘Raso lives in a house but I forgot which house.’ (Potsdam 2003: 299)

If syntactic isomorphism is the relevant condition governing the sluicing construction in Malagasy, as in Merchant’s original generalization, we wrongly predict that the P-less sluicing in (15b) should be ungrammatical. Potsdam claims instead that the correct derivation for (15b) before deletion is as in (16), in which the circumstantial voice is used to promote the object of the preposition to subject position before it undergoes *wh*-movement.

- (16) *mitoetra amin' ny trano i Raso fa adinoko hoe inona ilay trano*
 live.Act in the house Raso but forget.1sg HOE what that house
 (no itoeran' i Raso *t_{which house}*)
 Foc live.Circ Raso
 'Raso lives in a house but I forgot which house (is lived in by Raso)'

(Postdam 2003: 300)

In his later work, Potsdam (2007) provides extensive arguments that, in Malagasy, what is at stake is the semantic identity condition: an elided constituent must be semantically parallel to its antecedent but does not have to be syntactically parallel. Specifically, after establishing that *wh*-questions in Malagasy are pseudoclefts, Potsdam provides grammatical examples of sluicing in this language where the underlying structure for a sluice is a pseudocleft but the antecedent full-fledged is not. If this analysis holds, the P-stranding pattern observed above does not constitute a counterexample to the P-Stranding Generalization, as originally stated in Merchant (2001).

As noted above, a similar analysis has been independently proposed for *wh*-questions in Malay by Cole et al. (to appear). I show in section 3, however, that this analysis cannot be correct for sluicing in Indonesian.

2.2.5. Serbo-Croatian

Merchant (2001) mentions Serbo-Croatian as another non-P-stranding language that disallows P-stranding under sluicing. Stepanović (2008) provides examples to show that this language actually does not seem to obey the P-stranding generalization but ultimately concludes that the omission of the preposition under sluicing is due to some conditions at work at PF rather than stranding in the sense of extraction of the nominal complement from PPs.

The original examples that Merchant reports are given in (17a-c). (The examples in (17a, c) are from Merchant 2001: 97; the example in (17b) from Stepanović 2008: 179)

(17)a.* Kim je govorila Ana sa?

whom.Instr is spoken Ana with

‘Whom did Ana speak with?’

b. Sa kim je Ana govorila?

with whom is Ana spoken

‘With whom did Ana speak?’

c. Ana je govorila sa ekim, ali ne znam *(sa) kim.

Ana is spoken with someone.Instr but Neg I.know with whom.Instr

‘Ana spoke with someone, but I don’t know who with.’

Stepanović starts by pointing out that the non-strandability of *sa* ‘with’ in (17c) is due to an idiosyncratic property of the *wh*-phrase *kim* ‘who’. She illustrates this point with (18a-c).

- (18)a. Marko se ponosi Marijom /nekom djevojkom /nečim.
 Marko Refl takes.pride Marija.Instr some girl.Instr something.Instr
 ‘Marko is proud of Marija/some girl/something.’
- b. *(sa) kim se Marko ponosi?
 with whom.Instr Refl Marko takes.pride
 ‘Who is Marko proud of?’
- c. (?*Sa) kojom djevojkom/(?*Sa) čim se Marko ponosi?
 with which girl.Instr/ with what.Instr Refl Marko takes.pride
 ‘Which girl/What is Marko proud of?’

(Stepanović 2008: 180)

The example in (18a) shows that the verb *ponositi se* ‘to take pride’ selects the instrumental direct object. The fact that the animate, instrumental *wh*-phrase *kim* ‘who’ must occur following the preposition *sa* ‘with’ is due to independent factors idiosyncratic to this single *wh*-phrase since it is not required by the selectional properties of the relevant

verb, as shown in (18b). Indeed, when we have other *wh*-phrases such as *kojom djevojkom* ‘which girl’ and *čim* ‘what’, they cannot occur with the same preposition, as shown in (18c). This consideration, thus, suggests that it is hasty to draw conclusions on the status of Serbo-Croatian with respect to the P-Stranding Generalization. Stepanović then provides examples as in (19a-c) showing that Serbo-Croatian in fact presents a P-stranding pattern that superficially contradicts the generalization

(19)a. * Čega je Petar glasao protive?

what.Gen is Petar voted against

‘What did Petar vote against?’

b. Protiv čega je Petar glasao?

against what.Gen is Petar voted

‘Against what did Petar vote?’

c. Petar je glasao protiv nečega, ali ne znam (protiv) čega.

Petar is voted against something but Neg I.know against what

‘Petar voted against something, but I don’t know what.’

(Stepanović 2008: 181)

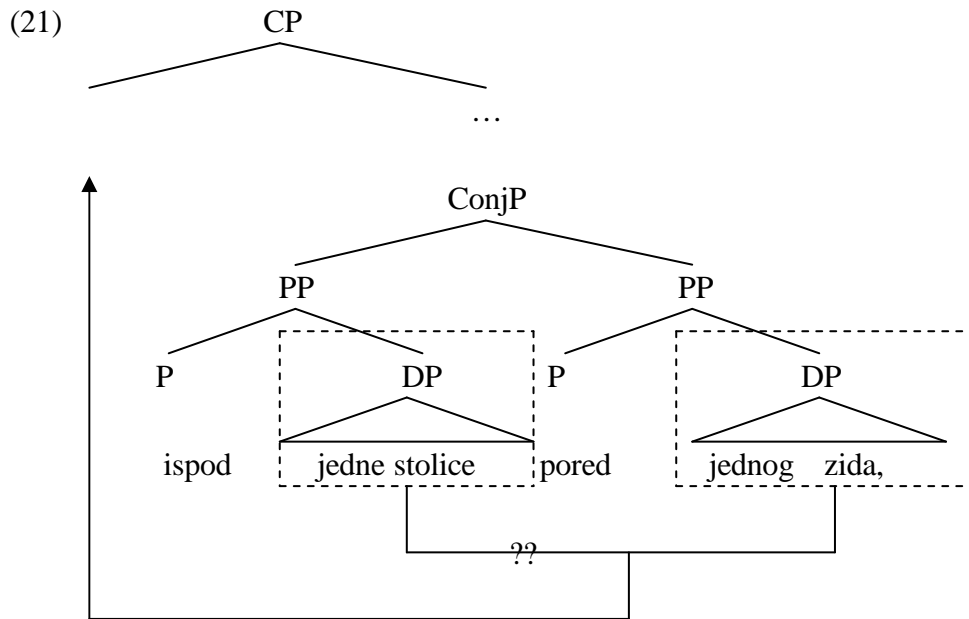
The contrast between (19a) and (19b) shows that the preposition *protive* ‘against’ cannot be stranded under *wh*-questions but the grammaticality of the P-less sluice in (19c), then, indicates that Serbo-Croatian is a counterexample to the P-Stranding Generalization.

Stepanović argues that this conclusion is not warranted since there is evidence that P-omission under sluicing is not due to the stranding of the preposition followed by TP deletion but some phonological operations at PF, rejecting potential alternatives such as base-generation (van Riemsdijk 1978) or cleft strategies on empirical grounds. Assuming that her demonstration is correct in that the underlying source of sluicing in Serbo-Croatian is sluicing (movement of a *wh*-phrase into [Spec, CP] followed by TP deletion), I concentrate here on evidence presented by her that the P-omission under sluicing cannot be due to P-stranding in the sense of extraction of the nominal complement of the P out of PPs as in Merchant (2001). Consider the following example of sluicing that involves coordinated PPs.

- (20) Petar je sakrio igračku ispod jedne stolice i pored jednog zida,
 petar is hidden toy under one chair.Gen and beside one wall
 ali ne znam (ispod) koje stolice i (pored) kojeg zida.
 but Neg I.know under which chair.Gen and beside which wall.Gen
 ‘Petar hide the toy under a chair and beside a wall, but I don’t know which chair and which wall.’

(Stepanović 2008: 183)

In this example, the antecedent PP in the full-fledged clause is a coordinated phrase with two PP conjuncts, each headed by a different preposition. The remnant in the sluicing clause also consists in a coordinated phrase that contains two *wh*-phrases without prepositions. If P-omission in this example were due to the stranding of the preposition as in Merchant’s (2001) deletion theory, then no derivation would yield the structure that feeds P-omission under TP deletion “because under no current theory of movement can the coordinated remnant phrase move as a constituent, while stranding the two Ps.” (p. 183). What she intends by this remark can be illustrated in the following schematic derivation for the example in (20): see Stepanović (2008: 186, 187) for arguments based on the number of places denoted by (20) that it involves PP coordination, not CP coordination.



This derivation involves extraction of two different *wh*-phrases into the specifier of CP but such a derivation is not possible under the current theory of extraction even under the Across-the-Board Movement because it should involve extraction of the same *wh*-phrase. The P-less sluice in (20), by contrast, naturally follows if the P-stranding is not due to the *wh*-extraction of the complement of the PP out of PPs but rather to some other operations such as P-omission at work at PF. Stepanović's (2008) analysis of the P-stranding in Serbo-Croatian, therefore, indicates that the examples in (18a-c) do not constitute a genuine counterexample to Merchant's Generalization. Stepanović (p. 188) speculates that P-stranding in this language "is a post-syntactic phenomenon, occurring possibly at PF" but leaves the precise implementation of this conclusion for future research. It is questionable, however, whether this PF P-omission

analysis works for Indonesian. It is known that Indonesian has the P-drop option, as shown in (22a).⁷ This observation, thus, gives the impression that Stepanovic's analysis might work for Indonesian. However, my native language consultants report that this option is not available for all prepositions. Thus, the free omission of prepositions such as *dengan* 'with' in the antecedent clause results in ungrammaticality, as shown in (22b).

(22)a. Saya ingat Hasan bicara (tentang)sesuatu, tapi saya tidak tahu apa.

I remember Hasan talk about something but I Neg know what

'I remember Hasan talked about something, but I don't know what.'

b. Saya ingat Hasan ber-dansa (*dengan)seseorang, tapi saya tidak tahu (dengan) siapa.

I remember Hasan Vz-dance with someone but I Neg know with who

'I remember Hasan danced with someone, but I don't know who.'

This contrast between (22a) and (22b), therefore, indicates that the P-drop analysis is not general enough to accommodate the consistent P-omission pattern observed in Indonesian sluicing.

⁷ Thanks to an anonymous reviewer for *Snippets* for pointing this out and Heidi Harley (personal communication) for useful discussion on problems with this analytical possibility.

2.2.6. “Noises” in Italian, Greek, and Polish

It has been noticed in several works that appeared as replies to Merchant (2001) that the P-Stranding Generalization holds neither across the board within a single language nor across languages. Merchant (2001) himself points out that there is a huge divergence among his consultants concerning the acceptability of the preposition-less sluice in languages such as Italian and Hebrew that uniformly prohibit P-stranding under *wh*-questions.⁸ In fact, Merchant gives judgments of the P-less variant of sluicing constructions in only 12 out of 18 non-P-stranding languages as totally ungrammatical; his reported judgments in the rest of the languages indicate that TP deletion improves at least a bit an otherwise illegal P-stranding violation under *wh*-questions. See also Tanaka (2007), who observes that the judgments reported by Merchant (2001) lack general agreement among speakers of languages such as Greek and Polish.

2.3. *Section Summary*

I have reviewed Merchant’s analysis of sluicing as the product of *wh*-movement, followed by TP deletion at PF. I have introduced the P-Stranding Generalization. I have examined several languages whose P-stranding pattern under *wh*-questions and sluicing constructions apparently contradict this generalization and have shown that none of these languages actually runs counter

⁸ Thanks to Adam Ussishkin (personal communication) for valuable discussion on Hebrew sluicing.

to it as the derivation of sluicing in these languages involves alternative syntactic or phonological means other than syntactic *wh*-movement. Based on the current survey, it is safe to conclude that no language seems to pose *genuine* counterevidence for the generalization. I have also argued that these alternative analyses cannot be applied to the superficially similar P-stranding pattern in Indonesian. It is against this background that the P-stranding pattern under *wh*-questions and sluicing in Indonesian, as illustrated in (5a-c), becomes very important. In the next section, I provide a variety of syntactic and morphological arguments to establish that Indonesian is *the first genuine counterexample* to the P-Stranding Generalization.

3. The Syntax of Sluicing in Indonesian

This section examines the syntax of sluicing in Indonesian. I start by reviewing Fortin's (2007b) results concerning Indonesian sluicing with respect to tests that Merchant developed to distinguish genuine sluicing and pseudosluicing. Though the overall result does indicate that Indonesian sluicing is derived by *wh*-movement, it is not conclusive, since the tests were based primarily on English, and there is independent evidence that some of the tests do not diagnose the syntax of sluicing in Indonesian in the same way as they do in English. I turn to Indonesian-internal arguments based on the distribution of the question morpheme *-kah* discovered by Fortin (2007b) and the lack of the complementizer *yang* in questions with non-nominal *wh*-phrases that

sluicing with PP remnants is derived via syntactic *wh*-movement followed by TP deletion, as in English. In the course of this discussion, I also reject two analyses of Indonesian *wh*-questions as reduced clefts (Cheng 1991) and headless relative clauses (Cole et al. to appear). This result confirms that Indonesian presents the first genuine challenge to the P-Stranding Generalization.

3.1. *Indonesian Sluicing = Pseudosluicing?*

Merchant (2001) argues that there are two types of sluicing. Genuine sluicing is derived by the *wh*-movement of a remnant in syntax followed by TP deletion. Pseudosluicing is derived by clefting of the *wh*-pivot (which involves *wh*-movement from the cleft predicate position, not from the complement position of a P), followed by deletion of the copula and the expletive subject. The two types of constructions are illustrated in (23a, b).

- (23)a. Pat was speaking to someone, but I don't know [_{CP} who ~~Pat was talking to~~ who].
- b. Pat was speaking to someone, but I don't know [_{CP} who ~~it was~~ that just left].

Merchant emphasizes the importance of the distinction between genuine sluicing and pseudosluicing as recent work on in-situ languages such as Japanese as in Nishiyama et al. (1996) and Kizu (2005) presents evidence that sluicing constructions in Japanese are elliptical clefts.

Merchant argues that English has genuine sluices by developing 10 operational tests to distinguish between pseudosluicing and genuine sluicing. The purpose of this section is to review Fortin's (2007b) discussion on Indonesian sluicing in this regard. Fortin applies those tests that are applicable to Indonesian *wh*-clefts and *wh*-questions and concludes that "the totality of the data appears to indicate that Indonesian sluices are elliptical *wh*-questions, and not elliptical clefts." (pp. 198-199). Her reported results are given in Table 2 (p. 206).

Table 2: Pseudosluicing Diagnostics in Indonesian

Diagnostics	Attested in sluices?		Attested in <i>wh</i> -clefts?		Attested in <i>wh</i> -questions?	
	English	Indonesian	English	Indonesian	English	Indonesian
adjuncts	√	√	X	X	√	√
implicit arguments	√	√	X	X	√	√
'mention-some'	√	√	X	X	√	√
'mention-all'	X	X	√	X	√	√
'else'	√	X	X	X	√	√
attributive adjs	√	√	X	X	X	√

The first diagnostic concerns the distribution of adjuncts. Merchant (p. 121) observes that an adjunct cannot be the pivot of a cleft (24a), but can be the remnant of a sluice (24b). The non-elliptical *wh*-question patterns with the sluice in this regard, as shown in (24c).⁹

- (24)a. * Pat is crying, but I don't know why it is.
- b. Pat is crying, but I don't know why.
- c. Pat is crying, but I don't know why Pat is crying. (Fortin 2007b: 199)

Fortin observes that Indonesian clefts are not possible with a *wh*-adjunct but the corresponding sluices are grammatical, as shown in (25a, b). (25c) shows that the *wh*-question patterns with (25b), indicating that Indonesian sluicing is based on *wh*-questions.

- (25)a.* Ali memperbaiki sepeda + nya, tapi saya tidak tahu bagaimana(kah) itu.

Ali AV-fix bike+3sg but I Neg know how-Q that

‘Ali fixed his bike, but I don't know how it was.’

⁹ One question, raised by Heidi Harley (personal communication), is how what Zagana (2007) analyze as quasi-PP arguments, such as instrumental, benefactive and locational expressions, behave with respect to this test. Because Zagana shows that this type of expression passes lots of tests for argumenthood, the prediction would be that they should be able to occur as the pivot of the cleft construction. I leave this important question for my future research.

- b. Ali memperbaiki sepeda + nya, tapi saya tidak tahu bagaimana (kah).

Ali Av-fix bike + 3sg but I Neg know how-Q

‘Ali fixed his bike, but I don’t know how.’

- c. Ali memperbaiki sepeda + nya, tapi saya tidak tahu bagaimana

Ali Av-fix bike + 3sg but I Neg know how

dia memperbaiki+nya.

3sg Av-fix-3sg

‘Ali fixed his bike, but I don’t know how he fixed it.’ (Fortin 2007b: 199, 200)

The second diagnostic concerns the distribution of sprouted implicit arguments. Chung et al. (1995) propose that there are two types of sluicing constructions in English. The first type, exemplified in (26a), has an overt correlate in the antecedent clause that corresponds to the *wh*-remnant in the sluice. The other type, exemplified in (26b), illustrates the sluicing construction in which the *wh*-remnant is licensed in the sluice without any overt correlate in the antecedent clause based on the argument structure of the verb in the antecedent clause. Chung et al. propose the LF structure building operation of “sprouting” to accommodate this second pattern of sluicing whereby the extra empty category can be constructed within the TP recycled from the antecedent TP.

- (26) a. She bought **something** the other day, but I don't remember **what**.
- b. They served the guests, but I don't know **what**. (Merchant 2001: 121)

Merchant (p. 121) and Fortin observe that sprouted implicit arguments are fine in sluicing and *wh*-questions but not in clefts, as the contrast between (27a) and (27b, c) shows.

- (27)a. * Robin was reading, but I don't know what it was.
- b. Robin was reading, but I don't know what.
- c. Robin was reading, but I don't know what Robin was reading. (Fortin 2007b: 200)

Fortin shows that this pattern is replicated in Indonesian, as shown in (28a-c). This is another indication that sluicing in Indonesian is based on *wh*-questions, not *wh*-clefts.

- (28)a. * Ali sedang memasak, tapi saya tidak tahu masakan apa(kah) itu.
- Ali Prog Av-cook but I Neg know dish what(-Q) that
- 'Ali is cooking, but I don't know what dish it is.'

- b. Ali sedang memasak, tapi saya tidak tahu masakan apa.
 Ali Prog Av-cook but I Neg know dish what
 ‘Ali is cooking, but I don’t know what dish.’
- c. Ali sedang memasak, tapi saya tidak tahu masakan apa(kah)
 Ali Prog Av-cook but I Neg know dish what-Q
 yang Ali masak.
 that Ali cook
 ‘Ali is cooking, but I don’t know what dish he is cooking.’ (Fortin 2007b: 201)

The next three tests concern different types of modification. The first of these is ‘mention-some’ modification, such as *for example*. Merchant (p. 122) and Fortin observe that this modification is possible with sluices and non-elliptical *wh*-questions, but impossible with *wh*-clefts, as the contrast between (29a) and (29b, c) shows. Fortin observes that Indonesian patterns with English in this regard, as illustrated in (30a-c).

(29) A: You should talk to someone in the legal department about that.

a. B: * Can you tell me who it is, for example?

b. B: Can you tell me who, for example?

c. B: Can you tell me who I should talk to, for example?

((12a, b) from Merchant 2001: 122, (12c) from Fortin 2007b: 201)

(30) A: Kamu harus makan lebih banyak sayur-mayur.

you should eat more many vegetable-Red

‘You should eat more (different kinds of) vegetables.’

a. B. # Misalnya, apa-kah itu? ‘for example, what is it?’

for example what-Q that

b. B: Misalnya, apa? ‘for example, what?’

for example what

c. B: Misalnya, apa yang harus saya makan?

for example what that should I eat

‘For example, what should I eat?’ (Fortin 2007b: 202)

The second of the three modification tests is a “mention-all” modification. Merchant (p. 122)

observes that this modification is fine with the cleft but not with the sluice. This is because *wh*-

clefts in English have the so-called exhaustivity requirement (Kiss 1998; Groenendijk and Stokhof 1997). This is illustrated in (31a, b). Note that *all* is compatible with the non-elliptical *wh*-question, as shown in (31c).

- (31)a. A bunch of students were protesting, and the FBI is trying to find out who all it was.
- b. * A bunch of students were protesting, and the FBI is trying to find out who all.
- c. A bunch of students were protesting, and the FBI is trying to find out who all was protesting.

((14a, b) from Merchant 2001: 122; (14c) from Fortin 2007b: 203)

This result shows that there is no telling from the *all*-modification test what the source of the sluice in Indonesian is. This is because, if *all* is compatible with the cleft and the full-fledged *wh*-question but not with the sluice, then there is an additional constraint at work in the sluice.

The Indonesian facts are also not clear in this regard. According to Fortin (p. 203), three of her four consultants report that *saja*, the Indonesian equivalent of *all*, can modify neither sluices nor clefts, though it can modify *wh*-elements in a full-fledged *wh*-question, as in (32a-c).

(32) A: Ada banyak tamu yang mendatangi pesta+ku.

exist many guest that Av-come-Loc party+1sg

a. B: *Tolong kasih tahu siapa saja-kah itu. 'Please, tell me who all it was.'

help give know who all-Q that

b. B: *Tolong kasih tahu siapa saja. 'Please, tell me who all.'

help give know who all

c. B: Tolong kasih tahu siapa saja-kah mereka. 'Please, tell me who all they were.'

help give know who all-Q they (Fortin 2007b: 203)

The contrast between (32c) and (32a, b) may be amenable to a semantic analysis independently of the validity of the 'mention-*all*' modification as a probe into the syntax of sluicing in Indonesian. As Rodriguez et al. (2007:10) also note for Brazilian Portuguese, Indonesian clefts do not exhibit the exhaustivity requirement due to some construction-specific bleached semantics. Evidence that the relevant requirement does not hold in Brazilian Portuguese or Indonesian comes from the fact that clefts in these languages allow negative quantifiers to serve as pivots of the cleft, a pattern that is impossible in English and Spanish. The contrast is illustrated in (33-d). (33b, d) are from Rodriguez et al. (2007: 10).¹⁰

¹⁰ As pointed out to me by Heidi Harley (personal communication), a possibility remains that *ninguém* 'nobody' in (33a) is a case of negative concord under a negated clause, not a true negative quantifier. If so, (33a) should be

(33)a. Não foi ninguém que bateu na porta. (Brazilian Portuguese)

not was nobody that knocked on.the door.

‘It was nobody that knocked on the door.’

b. Tak ada orang yang mengetuk pintu. (Indonesian)

Neg exist person that Av-knock door

‘It was nobody that knocked on the door.’

c. * No fue nadie que golpeó en la puerta. (Spanish)

not was nobody that knocked on the door

‘It was nobody who knocked on the door.’

d. * It was nobody who knocked on the door. (English)

The contrast between (33a, b) and (33c, d), therefore, indicates that the exhaustivity requirement is not a universal property of *wh*-clefts. If this conclusion is tenable, then the

grammatical only under the specific interpretation (“Somebody knocked on the door but no one of importance”) because the English equivalent *It was not anybody who knocked on the door* only allows this interpretation. I could not test this claim due to the unavailability of native speakers of Brazilian Portuguese. Interestingly, however, according to my three consultants, the Indonesian example in (33b) only allows a true indefinite reading (i.e. “Nobody knocked on the door”). Thus, the negative quantifier test could still serve to show that at least Indonesian clefts do not require exhaustivity. Nonetheless, it is a separate question whether or not the negative quantifier test proves that Indonesian clefts *forbid* exhaustivity, unlike English clefts, for otherwise (32a) would be acceptable. The crucial test case, then, would be one where a cleft in Indonesian is bad in a context which necessarily requires exhaustivity. My consultant work so far has been unsuccessful in identifying this case, so I must leave testing this prediction as an important task to be undertaken in my future research.

unacceptability of (36a) might follow from the contradictory semantic requirements imposed by *saja* ‘all’ and the non-exhaustive bleached semantics of clefts in Indonesian. (32c) is not surprising because the plural denotation of the subject of the embedded *wh*-question is compatible with the exhaustivity requirement.

The last test is *else*-modification. Expressions such as *else* cannot co-occur with the *wh*-cleft but can co-occur with sluicing and *wh*-questions in English, as shown in (34a, b). ((34a, b) from Merchant (2001: 122); (34c) from Fortin (2007b: 204).)

- (34)a. * Harry was there, but I don’t know who else it was.
- b. Harry was there, but I don’t know who else.
- c. Harry was there, but I don’t know who else was there.

The Indonesian facts do not come out in the same way as in English. *lagi* ‘again’ receives an interpretation akin to English *else* when modifying a *wh*-phrase. Fortin observes that *lagi* cannot modify the *wh*-pivot of a cleft (35a). However, it also cannot modify the *wh*-remnant of a sluice (35b), even though it can modify the *wh*-phrase in a full-fledged *wh*-question (35c).

(35) Ali datang ke pesta+ku

Ali come to party+1sg

a. ... * tapi saya tidak ingat siapa lagi(kah) itu.

but I Neg remember who else that

‘...but I don’t remember who else it was.’

b. ... * tapi saya tidak ingat siapa lagi.

but I Neg remember who else

‘...but I don’t remember who else.’

c. ... tapi saya tidak ingat siapa lagi(kah) yang datang.

but I Neg remember who else-Q that come

‘..but I don’t remember who else came.’ (Fortin 2007b: 204)

This result indicates that the *lagi*-modification test gives contradictory results to the other tests. Since the preponderance of evidence goes the other way in Indonesian, other factors seem to be at work in ruling out (35a).

The final test concerns adjectival modification. Merchant (p. 127) observes that extraction of an attributive adjective in both *wh*-clefts and non-elliptical *wh*-questions gives rise to

ungrammaticality, in contrast to sluicing, which allows such an extraction. This contrast is illustrated in (36a-c) from Fortin (2007b: 205); cf. Merchant (2001: 127).

- (36)a. * I heard that Pat met a nice guy, but I don't know how nice it is.
- b. I heard that Pat met a nice guy, but I don't know how nice.
- c. * I heard that pat met a nice guy, but I don't know how nice Pat met a guy.

The same pattern characterizes the extraction of an attributive adjective in Indonesian, as shown by the contrast between (37a, c) and (37b). ((20a, b) are from Fortin (2007b: 206).)

- (37)a. * Saya mendengar Siti menikah orang yang kaya, tapi saya
- I Av-hear Siti Av-marry person that rich but I
- tidak tahu [se-kaya apa itu].
- Neg know one-rich what that
- 'I heard Siti married a rich man, but I don't know how rich it is.'

b. Saya mendengar Siti menikah orang yang kaya,

I Av-hear Siti Av-marry person that rich

tapi saya tidak tahu se-kaya apa.

but I Neg know one-rich what

‘I heard Siti married a rich man, but I don’t know how rich.’

c. * Saya mendengar Siti menikahi orang yang kaya, tapi saya

I Av-hear Siti Av-marry person that rich but I

tidak tahu se-berapa kaya-kah itu sehinggadiala orang.

Neg know one-how rich-Q that far she person

‘I heard that Siti married a rich man, but I don’t know how rich she married a man’

This test requires careful elicitation. The point of this adjective modification test is whether the example in (37a) is acceptable to Indonesian speakers when the embedded question of the second clause asks *the degree of niceness* (i.e. *I don’t know how rich it is ~~that he is~~*). This consideration bears emphasizing because the topic of the sentence in (37a) is clearly a human, so speakers would find it so easy to access this topic, not the degree of niceness (which is hardly accessible even in English). With this consideration in mind, all of my three consultants uniformly reported that (37a, c) are bad while (37b) is fine under the “degree-of-

nicess” reading. This result, therefore, show that the sluice can be distinguished from the cleft in Indonesian, just as in English.

Let us summarize here what the results reported in Table 2 tell about Indonesian sluicing. Some of the tests (i.e. adjunct/implicit remnants and ‘mention-some’ modification) are suggestive that Indonesian sluicing is based on the non-elliptical *wh*-question, as in English. Some other tests (i.e. ‘mention-*all/else*’ modification) do not yield clear results in Indonesian, as they do in English. This is because there is an independent factor related to the arguably obligatory lack of the exhaustivity requirement of Indonesian *wh*-clefts on the pivot, which would be a prerequisite for all these tests to apply. The last test concerning extraction of the attributive adjective, however, at least serves to distinguish *wh*-clefts and sluices in Indonesian, as in English. Therefore, the reasonable conclusion is that although the results here are suggestive that sluices have different syntax from *wh*-clefts Indonesian, they are still conclusive as to the question of what is the underlying syntax of sluicing constructions in Indonesian. This result is hardly surprising, however, since many of Merchant’s tests are based on the syntax and semantics of *wh*-questions, sluicing, and *wh*-clefts in English which does not necessarily hold for Indonesian. For this reason, I turn in the next section to syntactic and morphological arguments internal to Indonesian that

conclusively determine whether Indonesian sluices are products of elliptical clefts or regular *wh*-movement plus TP deletion.

3.2. *Cheng (1991): A Reduced Cleft Analysis*

Cheng (1991) proposes the *Clausal-Typing Hypothesis*, namely, that the interrogative force of a statement must be marked either as a Q-particle in the scopal C or via the movement of a *wh*-operator into the specifier of the same C at S-structure. When combined with the Economy of Derivation (Chomsky 1995), which prohibits superfluous steps in syntactic derivation, the Clausal-Typing Hypothesis predicts that if a language has a Q-particle in its lexical inventory, that language should use it for all types of *wh*-questions, thereby excluding the need for a *wh*-phrase to undergo syntactic movement into the specifier of CP on the grounds of the Economy of Derivation and yielding the *wh*-in-situ option across the board. If a language lacks a Q-particle in its lexical inventory, on the other hand, then that language should use syntactic *wh*-movement for the purpose of clause-typing. As shown in chapter 2 of this dissertation (see also chapter 5), however, Indonesian has three ways to form *wh*-questions (overt movement, partial movement, and in-situ) and has a Q-particle, *-kah*. This state of affair would appear to contradict the prediction of Cheng's Clause-Typing Hypothesis because, as the *wh*-in-situ option indicates, Indonesian has a Q-particle as a

clause-typing morpheme and this option should block overt *wh*-movement into the scopal specifier of CP. Accordingly, Cheng (1991) argues that what appears to be an overt *wh*-movement construction in Indonesian is actually a reduced cleft (where the expletive subject and copula are missing), not a genuine *wh*-question, as in English *wh*-questions. According to this analysis, the sentence in (38a) would have the structure in (38b).

- (38)a. *Apa_i yang kamu beli t_i ?*
 what that you buy
 ‘What did you buy?’
- b. [_{CP1} *apa_i* [_{CP2} *Opi* *yang* [_{TP} *kamu* *beli* *t_i*]]

In the structure in (38b), the *wh*-phrase *apa* ‘what’ is base-generated in the specifier of CP1. The null operator undergoes TP-internal movement from the object of the verb into the specifier of CP2. If this analysis is tenable, the Indonesian P-stranding pattern does not contradict the P-Stranding Generalization.

However, there is evidence presented by Fortin (2007b) that the derivation of sluicing involves true *wh*-questions, not elliptical clefts. Fortin observes that the question particle *-kah*

can co-occur with the *wh*-pivot of the cleft construction but not with the *wh*-remnant of the sluicing construction. This is illustrated by the contrast between (39a) and (39b).

(39) Ada seseorang yang menelpon tadi ...

exist someone that AV-phone just now

a. coba tebak siapa-(**kah**) itu!

try guess who-Q Dem

‘try to guess who it was!’

b. coba tebak siapa-(***kah**)!

try guess who-Q

‘try to guess who!’

c. coba tebak siapa(***kah**) yang menelpon tadi!

try guess who-Q that AV-phone just now

‘try to guess who just called now!’

(Fortin 2007b: 207, 208)

Cheng argues that the ostensible cases of over *wh*-movement in Indonesian are all reduced clefts. Then, if the example in (39b) were derived by the cleft construction in (39a), as it were under Cheng’s cleft analysis, the example in (39b) should be able to allow the

question marker *-kah* to occur with the *wh*-phrase *siapa* ‘who’. Thus, the distribution of *-kah* separates “true” clefts from genuine sluices and *wh*-questions. Note that this pattern follows straightforwardly if Indonesian has true *wh*-movement of the English type. Cheng’s reduced cleft analysis, on the other hand, would need some extra stipulation to capture the distribution of the question morpheme illustrated in (39a-c).¹¹ Based on this consideration, I conclude that Indonesian has true *wh*-movement.

3.3. *Cole et al. (To appear): A Headless Relative Clause Analysis*

Cole et al. (to appear) argue for a different analysis of *wh*-questions in Malay (and Indonesian, by extension). They propose that *wh*-questions with the complementizer *yang* and those without have two different syntactic derivations: the former types of questions involve short focus movement of the *wh*-phrase from the post-copula position to the specifier of the matrix CP while the latter type of questions involve successive cyclic movement of the *wh*-phrase from its base position to the specifier of CP as in English *wh*-questions: see also Davies’ (2003, 2005) proleptic analysis of *wh*-questions in Madurese that deny the existence of long-distance movement. Consider (40a) and its derivation in (40b).

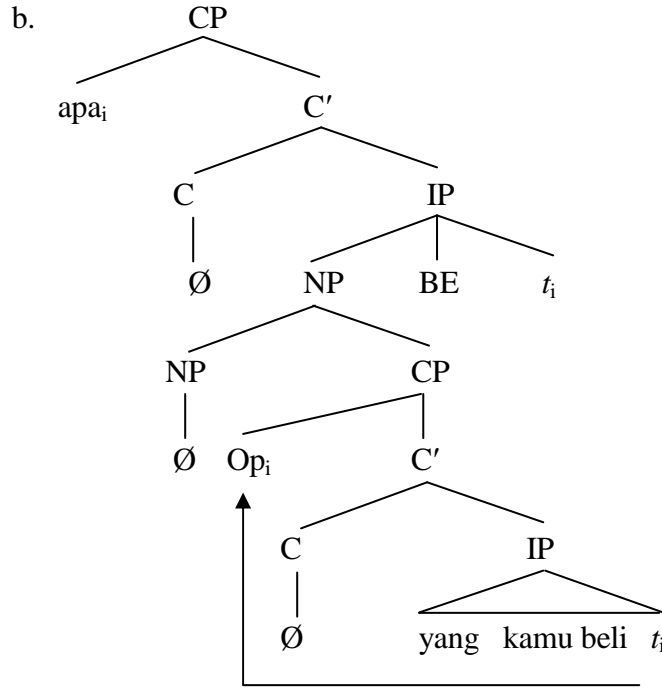
¹¹ For example, Cheng’s analysis could be saved by saying that genuine clefts could reduce *-kah* as well as the copula, but this would still leave unexplained why *kah- must* be reduced in the sluice in (39b) if the sluice were derived through clefting by the further deletion of the expletive subject.

(40)a. Apa_i yang kamu beli t_i ?

what that you buy

‘What did you buy?’

(Malay/Indonesian)



(modified from Cole et al. to appear: 4)

In this derivation, there is TP-internal movement of the null interrogative operator from the TP internal position to the specifier of the embedded CP. The *wh*-phrase *apa* ‘what’ undergoes focus movement from the position following the null copula (BE) to the specifier of the matrix CP. One of their arguments for this headless relative clause analysis of *wh*-questions with *yang* comes from the categorical restriction on interrogative elements that can be fronted in this type of question. Consider (41a-e) and (42a-e).

(41)a. *Apa* yang di-perbaiki Ali?

what that Pv-fix Ali

‘What was fixed by Ali?’

b. *Siapa* yang melihat kau?

who that see you

‘Who saw you?’

c.?? *Di mana* yang kau tinggal?

at where that you live

‘Where do you live?’

d.?? *Bagaimana* yang Ali memperbaiki kereta itu?

how that Ali fix car that

‘How did Ali fix that car?’

e. ?? *Kenapa* yang Ali di-pecat?

why that Ali Pass-fire

‘Why is it that Ali was fired?’

(Indonesian; modeled after the Malay examples from Cole et al. to appear: 6, 7)

(42)a.?? Yang aku tinggal (ialah) *di K.L.*

that I stay is at K.L.

‘The place that I live is in K.L.’

b.?? Yang Ali perbaiki kereta itu (ialah) *dengan alat-nya.*

that Ali fix car that is with tool-his

‘The way that Ali fixed that car is with his tool.’

c.?? Yang Ali di-pecat (ialah) *kerana dia cuai.*

that Ali Pv-fire is because he careless

‘Why Ali was fired is because he was careless.’

d. Yang aku makan *nasi goreng* (-lah).

that I eat rice fried-focus

‘Fried rice is what I am eating.’

e. Yang kau nampak *Siti*(-lah).

that you see Siti-Foc

‘Siti is what you see.’

(Indonesian; modeled after the Malay examples from Cole et al. to appear: 9)

It is clear from the examples in (41a-e) that, when *wh*-questions are formed with *yang*, only questions with nominal *wh*-phrases such as *apa* ‘what’ and *siapa* ‘who’ are well-formed. This categorical restriction would remain mysterious under the common analysis of *wh*-questions in languages such as English as fronting of an interrogative phrase into the specifier of the matrix CP, because no such restrictions would be imposed on the kind of elements to be fronted. This observation, by contrast, directly follows if the underlying structure of *yang*-questions is clefting, because the same restriction is independently observed in cleft constructions, as shown by the examples in (42a-e). Their analysis, thus, might mean that the underlying syntax of sluicing would be a headless relative clause, an idea that has also been argued for by Potsdam (2003, 2007) for Malagasy sluicing, as we have seen in the previous section.

However, the headless relative clause analysis would only work for sluicing examples with nominal *wh*-phrases. This point is clearly emphasized by Cole et. al’s (to appear: 26) conclusion that “questions without *yang* involve potentially long distance movement of the WH word itself.’ Consider examples in (43a, b).

(43)a. *Bila_i Maryam pikir [yang Ali akan datang ke sini t_i]?*

when Maryam think that Ali will come to here

‘When does Miriam think that Ali will come here?’

b. *Kenapa_i Siti kata [yang Fatimah beli ikan itu t_i]?*

why Siti say that Fatimah buy fish that

‘Why did Siti say that Fatimah bought that fish?’

(Indonesian; modeled after the Malay examples from Cole et al. to appear: 27)

The examples in (43a, b) show that *wh*-questions with non-nominal *wh*-elements such as *bila* ‘when’ and *kenapa* ‘why’. For Cole et al. (to appear), the lack of *yang* means that the questions are derived by regular *wh*-movement of an interrogative phrase into the specifier of CP. Then, the obligatory absence of *yang* in (41c) indicates that the *wh*-question with the fronted PP cannot be analyzed as the headless relative clause because there is an independent restriction that the nominal head of such a clause must be nominal *wh*-phrases such as *apa* ‘what’ and *siapa* ‘who’. Thus, I conclude that at least the PP *wh*-question is derived by regular *wh*-movement, as in English.¹²

¹² Note that the argument for the *wh*-movement analysis of sluices here is the Indonesian analogue of Merchant’s (2001) adjunct test reviewed in section 2.1.

3.4. *Sluicing in Indonesian ≠ Pseudosluicing: Why Indonesian Is So Special*

In this section, I have presented evidence based on the distribution of the question particle *-kah* and the obligatory absence of *yang* in *wh*-questions in Indonesian that the derivational source for sluicing in Indonesian (at least with PP remnants) cannot be clefting as in Cheng (1991) or headless relative clause as in Cole et al. (to appear). The argument made by Fortin (2007b) from the question particle shows that the sluicing construction patterns with a *wh*-question but not with a cleft counterpart. The argument from the lack of *yang* shows that the derivation of sluicing with the PP remnant involves regular *wh*-movement of the PP. This result, therefore, suggests that the Indonesian sluicing with the PP remnant is derived by regular *wh*-movement. With this observation in place, consider again the examples in (5a-c), repeated here as (44a-c).

(44)a. * Siapa yang kamu ber-dansa dengan?

who that you Vz-dance with

‘Whom did you dance with?’

b. Dengan siapa kamu ber-dansa?

with who you Vz-dance

‘With whom did you dance?’

- c. Saya ingat Hasan ber-dansa dengan seseorang, tapi saya tidak
 I remember Hasan Vz-dance with someone but I Neg
 tahu (dengan) siapa.
 know (with) who
 ‘I remember Hasan danced with someone, but I don’t know (with) whom.’

We have seen in section 3 that several other languages that behave superficially similar to this pattern of P-stranding are not genuine counterexamples to the same generalization. I have shown in this section, however, that at least the Indonesian slucing with PP remnants is derived by regular *wh*-movement of the remnant followed by TP deletion. Therefore, I conclude that the P-stranding pattern in Indonesian presents *the first genuine counterexample* to the P-Stranding Generalization.

4. Salvation by Deletion at the Syntax-Phonology Interface

In this section, I propose a novel analysis of the (un-)availability of the P-stranding across languages that draws on independently motivated assumptions. I show that the proposed analysis provides a straightforward explanation for the three-way contrast between English, French, and

Indonesian with respect to P-stranding under *wh*-questions and sluicing. In particular, the apparently atypical pattern of Indonesian is fully expected under the proposed analysis.

What is crucial below is the idea behind the thesis of minimalist interface, introduced in chapter 1, that certain imperfections created by the syntactic derivation can be ameliorated by syntax-external operations such as deletion of the offending part of the derivation. This idea of “salvation by deletion” goes back to Ross’ (1969) seminal analysis of the ameliorating effect of deletion on subjacency-violating movements in global terms and has been resurrected in recent minimalist research on the syntax-phonology interface as in Merchant (2001), Lasnik (1999, 2001, 2005, 2007), and Boeckx and Lasnik (2006). The proposed analysis, thus, further substantiates the claim that syntax-external components do whatever they can to save an otherwise illicit syntactic object within the narrow range of options permitted by the interaction of universal principles and parametrically defined options available in a particular language.

4.1. *Feature Percolation, Minimality, D-to-P Incorporation, and Interface Repair*

The proposed analysis adopts three independently motivated assumptions concerning the percolation of the [+wh] feature borne by a nominal element onto its dominating PP, D-to-

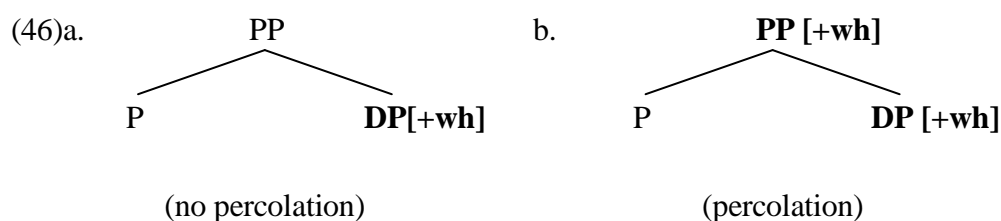
P coalescence, and, most importantly, the notion of interface repair. Let us first discuss each of these assumptions and how they are independently motivated.

4.1.1. Whether the [+wh] Feature Percolates or Not

The proposed analysis claims that there is a parameter concerning the percolation of the [+wh] feature of the interrogative D onto its dominating PP. This idea was first proposed by Chomsky (1972) to answer a criticism raised by Postal (1972). Postal observes that, if movement is successive-cyclic in the sense defined in chapter 2 of this dissertation, it would predict that the preposition should be able to be stranded in any one of the specifiers of intermediate CPs. The ungrammatical examples in (45d, e), however, indicate that this prediction is incorrect.

- (45)a. I believe Mary thinks Joan talked to someone.
- b. **Who** do you believe Mary thinks Joan talked *to*?
- c. **To whom** do you believe Mary thinks Joan talked?
- d. * **Who/Whom** do you believe *to* Mary thinks Joan talked?
- e. * **Who/whom** do you believe Mary thinks *to* Joan talked? (Postal 1972: 213)

The relevant generalization here is that prepositions in English must either be stranded in situ or be pied-pied into the specifier of the matrix CP. Chomsky argues that this generalization naturally falls out if we assume that the [+wh] feature of the *wh*-expression can percolate onto its dominating PP in English, in the manner seen in (46a, b); see Stepanov (2001) and Medeiros (2006a) for implementations of this idea within the phase-based framework to raising and pied-piping with inversion, respectively.



Rephrasing Chomsky's original analysis that employs the A-over-A Principle within the modern minimalist framework, when the [+wh] feature does not percolate as in (46a), the closest element from the perspective of the interrogative C is the DP. This option, thus, yields the stranded preposition structure as illustrated in (45b). When the [+wh] feature does percolate as in (46b), however, it is the PP now marked with that feature that is moved/attracted by the interrogative C. This option, thus, yields the pied-piping structure as illustrated in (45c). Notice that, under this feature-based analysis, there is no way in which the preposition can be stranded

in intermediate sites because the decision as to whether the relevant feature is percolated onto PP or not is made when the derivation constructs the PP, as shown in (46a, b): once it percolates, the shortest attract/movement requirement demands that the PP must move entirely as it is the closest interrogative element. If it doesn't, the same requirement demands that the *wh*-phrase itself must be carried onto the specifier of the matrix CP. Thus, the feature percolation operation receives independent empirical motivation.¹³

The proposed analysis also provides a natural explanation for one well-known fact about P-stranding. It has been widely acknowledged since the single most important work on P-stranding by van Riemsdijk (1978) that it is a crosslinguistically very marked option that is observed only in a handful of languages such as English, Frisian, Swedish, Norwegian, and other Scandinavian languages. It is virtually unattested in Romance languages and in other branches of Germanic, for example. Riemsdijk argues that the lack of this option in Romance can be accommodated if we assume that these languages have the syntactically represented marked option of projecting an escape hatch in the specifier of PPs that *wh*-

¹³ Alternatively, as suggested to me by Andrew Carnie (personal communication), “percolation” could be recast as whether or not the P head has an unvalued *wh*-feature which agrees with the interrogative DP or not. According to this alternative, in non-P-stranding languages like French and Indonesian, the P head has an unvalued *wh*-feature which agrees with the *wh*-DP. Thus, the closest constituent to be attracted is the whole PP. In contrast, in P-stranding languages, the P-head does not have this feature, hence the DP must move into [Spec, CP], leaving the P behind. I think that this alternative would need some special mechanism to explain why English also allows the pied-piping option because the DP would be the closest element to the C. Therefore, I assume the feature percolation analysis here.

phrases can go through to end up in the specifier of CP. I will not review many empirical problems with this analysis; see Hornstein and Weinberg (1981), Kayne (1981), Stowell (1981, 1982), Herslund (1984), Takami (1988), Abels (2003), Law (1998, 2006), Salles (1997), and sources cited therein. I do want to point out, however, that this analysis is not groundable within the modern theory of phrase structure such as the Bare Phrase Structure Theory (Fukui 1986, 1995; Speas 1990; Chomsky 1995) as nothing in this theory can block the projection of specifier as it is automatically created by the internal merge of the NP into [Spec, PP].¹⁴ One could technically block the projection of such a position within the Phase Theory by not assigning EPP-features to drive the required movement but this analysis is obviously circular unless independent evidence is provided for it.

By contrast, as Lasnik (2005) observes, Chomsky's feature percolation analysis provides a natural account of the unavailability of the P-stranding in Romance languages that is free from the technical problem noted above. Let us suppose that there is a parameter with respect to the optionality of the feature percolation; the [+wh] feature a) *can* percolate in English but *must* percolate in Romance and other languages including Indonesian that do not allow P-stranding under *wh*-questions. Under this parametric analysis, the latter type of languages do not allow P-stranding under *wh*-movement because the closest element to be

¹⁴ I thank Heidi Harley (personal communication) and Simin Karimi (personal communication) for pointing this out.

attracted by the interrogative C is always the PP, as shown in (46b), hence the *wh*-phrase would never be attracted by the same head on the ground of shortest movement; English and other languages mentioned above allow P-stranding, however, because these languages have the parametric option of not percolating the feature onto PPs. This parametric view of feature percolation is also in line with the standard assumption that the values of a parameter must be learnable from the visible cues and localized to the properties of lexical items; see also chapter 4 for relevant discussion. This consideration, therefore, provides further independent motivation for the feature percolation analysis.¹⁵

4.1.2. D-P Coalescence as D-to-P Incorporation

It is well-known that, in Romance languages, a preposition sometimes coalesces with the following determiner element into a suppletive form. Consider (47) from French and (48) from Italian. A list of examples of other examples of D-P coalescence in these languages is also given below each example.

¹⁵ See also our discussion of Law's (1998, 2006) analysis of the impossibility of P-stranding in Romance shortly below. I argue in section 4.2 that both Law's analysis and the minimality constraint on movement are independently necessary to derive the three-way contrast between English, French, and Indonesian with respect to P-stranding.

(47) Jean a parlé du sujet le plus difficile. (French)

Jean have talked about-the subject the most difficult

‘Jean talked about the most difficult subject.’

suppletive forms: du = de le, des = de les, duquel = de lequel, à les = aux, à le = au,

desquels = de lesquels ‘of the’, à lequel = auquel, à lesquels = auxquels ‘to the’

(Law 1998: 226)

(48) Gianni ha parlato del soggetto più difficile. (Italian)

Gianni have talked about-the subject most difficult

‘Gianni talked about the most difficult subject.’

suppletive forms: al = a il, alla = a la ‘to the’, sul = su il, sulla = su la ‘on the’, nel =

in il, nei = in i ‘in the’, del = di il, dello = di lo ‘of the’, col = con il ‘with the’

(Law 1998: 226)

In the French example given in (47), the preposition *de* coalesces with its following determiner *le* to yield a suppletive form *du*. Similar observations characterize the D-P coalescence in the Italian example in (48) ((*di + il = del*); see also German examples below in (52a, b). Law (1998, 2005) (see also Beermann 1990 and van Riemsdijk 1998: 639)

proposes that there is a syntactic constraint on suppletion, as defined in (49), to account for the impossibility of P-stranding under *wh*-movement in Romance languages.¹⁶

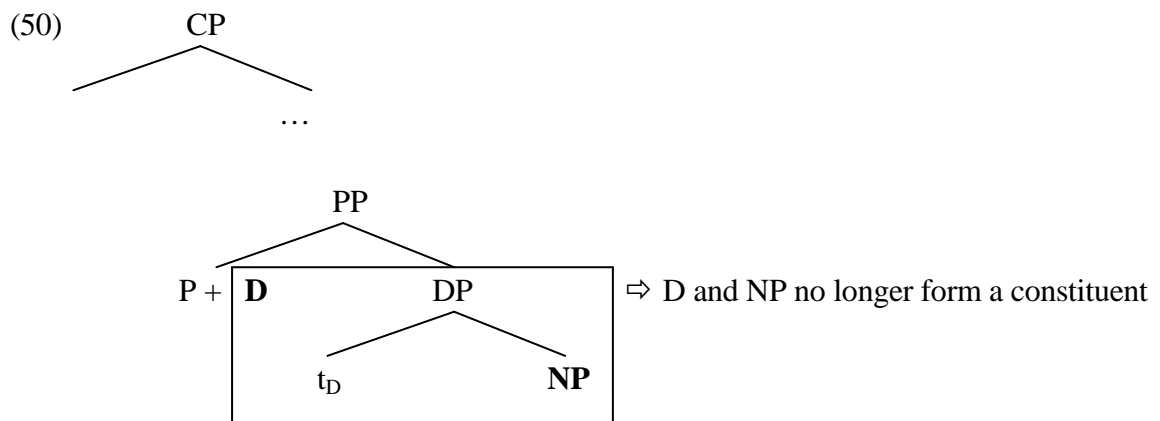
(49) Syntactic Constraint on Suppletion (Law 1998: 22)

Elements undergoing suppletive rules must form a syntactic unit X^0 .

This constraint essentially states that determiners must incorporate onto their governing prepositions to be reanalyzed in the post-syntactic component as a suppletive element. Of course, there are cases (e.g., *de la* in French) where the D-P coalescence does not occur, but it is not likely that general syntactic operations such as D-to-P incorporation should be constrained by this type of unpredictable morpholexical gap. Rather, a more plausible analysis would be one in which the D-to-P incorporation occurs across the board in French and Italian whether or not its effect is morphophonologically realized in the form of coalescence.

¹⁶ To account for the fact that V bears agreement and tense inflection despite the lack of V-to-T movement in English, Bobaljik (1995) proposes that V undergo Morphological Merger with T post-syntactically under adjacency. The principle in (49), thus, might not be consistent with Bobaljik's proposal. However, as Heidi Harley (personal communication) notes, this depends on where the representation is checked. If the constraint in (49) is checked before Morphological Merger, then it would be violated, but if it is checked at PF, after morphological operations had applied, the Morphological Merger operation would have indeed created a single V + T head and the constraint would be fulfilled. I will not pursue this matter further in this chapter.

The constraint in (49) provides a straightforward answer for why Romance prohibits P-stranding; once the D incorporates into the P, the N and D no longer form a constituent. As a result, the movement of the D + N becomes impossible, as illustrated in a schematic derivation in (50).



Law's analysis predicts that the D-P coalescence should be impossible when independent syntactic conditions block incorporation of D onto P in the syntax. This prediction is confirmed by the observation that the coalescence is impossible in examples shown in (51a, b) from French.

- (51)a. Je lui ai demandé [_{CP} *de le/*du* lire] 'I asked him to read it.'
- I him have asked to it read

- b. Nous sommes prêts [_{CP} *à le/*au* faire] ‘We are ready to do it.’
 we are ready to it do

(Law 1998: 227, 228)

(51a) illustrates that coalescence does not occur between the head of the embedded complementizer and its following clitic subject which is attached to the verb. Similarly, (51b) illustrates that it also does not happen between the complementizer and the embedded clitic object which is dependent on the verb. The impossibility of D-P coalescence in (51b) is what is expected under the constraint in (49) above because the incorporation of the complement of the verb into the C/T head *à* would be an instance of ex-corporation, which is generally assumed to be impossible.¹⁷

It is important to clarify the status of the constraint in (49). Law (1998, 2006) maintains that this constraint is a necessary condition that must be satisfied in the syntactic component for the D-P sequence to be reanalyzed as a suppletive form at the post-syntactic component. van Riemsdijk (1998) shows, based on examples structurally similar but more

¹⁷ Alternatively, the impossibility of D-P coalescence in (51a, b) could be blocked by the intervention of the CP projection between the D and P if CPs serve to demarcate possible domains for prosodic phonology, as independently proposed in Bošković and Lasnik (2003), An (2006), Richards (2006), and Sato (in press c). See van Riemsdijk (1998: 655) for a similar observation. For the purposes of the present discussion, however, I simply follow Law’s exposition here.

complicated than examples such as (47-48), that there are cases in where D-P coalescence are blocked even though D and P are phonologically adjacent when certain syntactic configurations are not met. I repeat one construction in German to illustrate this point. Consider the following examples.

- (52)a. von [DP [D e] [AP dem König treu ergeben] [N Dienern]]
 of the_{Dat} king faithfully devoted servant
 ‘of the servant that is faithfully devoted to the king’
- b. * vom König true ergebenen Dienern.
 of-the_{Dat} king faithfully devoted servant
 ‘of the servant that is faithfully devoted to the king’ (van Riemsdijk 1998: 655)

In (52a), the preposition *von* ‘of’ selects the DP complement. Within this complement, the adjective *ergeben* ‘devoted’ governs the dative DP complement *dem König* ‘the king’ to its left. If phonological adjacency were the only relevant condition that governs the D-P coalescence in German, we would predict that the contraction of *von* and *dem* would yield the suppletive form *vom* (see van Riemsdijk 1998: 653 for the basic morphophonemics of coalescence in German), since the determiner is linearly adjacent to the preposition without any prosodic boundary such

as CP in between. This prediction is not borne out by the ungrammaticality of the example shown in (52b). On the other hand, the failure of D-P coalescence in this example naturally falls into place if we assume that independent syntactic constraints, either the Empty Category Principle, the Head Movement Constraint or the phasehood of DPs, that blocks the head movement from within the embedded AP to the P over the intervening null D. For example, the Head Movement Constraint correctly blocks the required head movement on a par with examples such as **Have_i you could t_i left?* This observation, therefore, shows that D-to-P coalescence has its source in the syntax, even though its morphophonological effect is realized post-syntactically in the form of suppletion. This point becomes important in section 4.2 when I propose a new analysis of P-stranding.

4.1.3. Repair of “Imperfections” at the Syntax-Phonology Interface

The final and most important idea I pursue in this chapter is the idea that syntax makes “mistakes”, some of which can be remedied later by syntax-external operations such as deletion. This idea has been circulated since the late 1960s, when Ross (1969) observed that the sluicing transformation ameliorates island-violations that would otherwise yield ungrammatical sentences. Some original examples noted by Ross are given in (53-54) with his own judgments indicated.

(53) The Complex NP Constraint

- a. * She kissed a man who bit one of my friends, but Tom doesn't realize which one of my friends she kissed a man who bit.
- b. ? She kissed a man who bit one of his friends, but Tom doesn't realize which one of my friends.

(Ross 1969: 276)

(54) The Sentential Subject Constraint

- a. * That he'll hire someone is possible, but I won't divulge who that he'll fire is possible.
- b. ?? That he'll hire someone is possible, but I won't divulge who.

(Ross 1969: 277)

The contrast between (53a) and (53b) shows that the sentence that would be ungrammatical due to the Complex NP Constraint is improved when sluicing deletes the part (namely the embedded TP) that involves the violation of this constraint. Similarly, the grammaticality of the example in (54b) shows that the Sentential Subject Constraint is also ameliorated under sluicing. Ross argues based on this observation for the necessity of global trans-derivational comparison, as stated in (55).

(55) If a node is moved out of its island, an ungrammatical sentence will result. If the island-forming node does not appear in surface structure, violations of lesser severity will (in general) result. (Ross 1969: 277)

In other words, “ungrammaticality is a property not of merely deep or surface structure, or of pairs of trees which are related by rules, but rather of derivations.” (p. 277). In contrast to Ross’s original judgments cited above, recent researchers (e.g., Lasnik 1999, 2001, 2007, Fox and Lasnik 2003; Boeckx and Lasnik 2006) are in agreement that sluiced versions as in (53b) and (54b) are perfect rather than marginal. Following Chomsky (1972), Merchant (2001) proposes a revision of Ross’s statement by arguing that sluicing ameliorates certain island violations as illustrated in (53a) and (54a) because they constitute islands only at PF; thus, (53b) and (54b) become grammatical because the island-violation is nullified at PF by deleting the structure that encodes such a violation.

Elaborating this point further, Boeckx and Lasnik (2006) claim (cf. Aoun et al. 1987 and Aoun and Li 2003) that both derivational and representational constraints must be admitted into the theory of grammar. They provide data concerning *wh*-island effects and superiority under sluicing in Serbo-Croatian and resumption in Egyptian Arabic. I just review their argument based on the examples from Serbo-Croatian, given in (56-59). (The angled bracket in (58) means that the words are deleted/unpronounced.)

(56) Ivan i Marko ne znaju....

Ivan and Marko Neg know.

a. ko je šta kupio.

who is what bought

‘Who bought what’

b. * šta je ko kupio.

what is who bought

‘What who bought’

(Boeckx and Lasnik 2006: 152)

(57)a. Somebody bought something, but...

b. i. Ivan i Marko ne znaju ko šta.

Ivan and Marko Neg know who what

‘Ivan and Marko know who what.’

ii.* Ivan i Marko ne znaju šta ko.

Ivan and Marko Neg know what who

‘Ivan and Marko don’t know what who.’ (Boeckx and Lasnik 2006: 152)

(58) * idio sam knjigu koju se pitam ko prodaje <knjigu koju>.

seen am book which SE wonder.1sg who sells

‘I saw a book which I wonder who sells.’ (Boeckx and Lasnik 2006: 151)

(59)a. Every journalist went out today to find out who was selling a certain book....

b. ali ne znam koju (knjigu).

but Neg know which book

‘but I don’t know which (book).’ (Boeckx and Lasnik 2006: 152)

The examples in (56a, b) illustrate that Serbo-Croatian exhibits the superiority effect in multiple *wh*-questions. The contrast in grammaticality between (57a) and (57b) shows that this effect persists after sluicing. Interestingly, however, the *wh*-island effect behaves different from superiority in this respect. Serbo-Croatian exhibits the *wh*-island effect, as shown in (58). However, when sluicing is applied to eliminate the TP that contains the island violation, the result becomes grammatical, as illustrated in (59b). Boeckx and Lasnik suggest one possible approach to the observed difference in “reparability” between superiority and *wh*-island effects. Specifically, they propose that superiority is a derivational constraint *within syntax* that thereby is immune to interface operations such as deletion whereas the *wh*-island effect is a

representational constraint imposed on chains that thereby is subject to repair at the syntax-external system. This bipartite approach to syntactic violations is also hinted at in Almeida and Yoshida (2007) and Lasnik (2007) in light of the P-stranding pattern in Brazilian Portuguese, who speculate that there is more than one source of P-stranding ban, one derivational and hence irreparable, and the other representational and hence reparable.

The survey of the literature concerning syntactic violations conducted here provides support for the assumption that syntax could make certain mistakes, some of which external interpretive components can remedy by domain-specific operations such as deletion. This assumption plays an important role in the following section. We return to the issue of derivational vs. representational theories of grammar in section 5 when we discuss Fortin's (2007b) analysis of Indonesian sluicing.

4.2. Towards an Etiology of P-Stranding Violations across Languages

In this section, I propose a novel, parametric analysis of the typology of P-stranding under sluicing that draw on these independently motivated assumptions. The proposed analysis was in Table 1, repeated here.

Table 1: The Parametric-Theory of P-Stranding at the Syntax-Phonology Interface

Parameters Languages	[+wh] feature percolation from the D to PP	D-to-P incorporation in the syntax
English	OPTIONAL	NO
Indonesian	OBLIGATORY	NO
French	OBLIGATORY	YES

The proposed analysis claims that the three-way contrast in P-stranding possibilities under sluicing and *wh*-movement among English, Indonesian, and French is naturally derived by two simple parametric choices: a) whether or not the [+wh] feature MUST percolate onto PPs and b) whether there is D-to-P coalescence/syntactic incorporation. English can optionally percolate the *wh*-feature onto the PP and lacks syntactic incorporation of D to P. Indonesian must percolate the feature onto the PP and lacks syntactic incorporation of D to P. French must percolate the feature onto the PP and has syntactic incorporation of D to P.

This cross-classification of the two parameters might predict languages that optionally percolate the feature and have syntactic incorporation of D to P, the fourth logically possible combination of the above features. In such a language, the [+wh] feature would optionally percolate to P, while D would mandatorily incorporate into D. Note the present parametric analysis correctly predicts that the existence of this type of languages could not be detected,

because even though the non-percolation option would potentially allow the *wh*-phrase to be directly accessible to the interrogative C as in English, P-stranding would be independently blocked by the D-to-P incorporation, as in French and Italian. The only convergent derivations would have [+wh] percolation with the D-to-P incorporation, and so in practice this grammar would produce a P-stranding pattern indistinguishable from the French pattern.

4.2.1. English

Consider first why English allows P-stranding both under *wh*-movement and sluicing in conformity with Merchant's (2001) P-Stranding Generalization. The answer is quite straightforward under the proposed parametric analysis of P-stranding summarized in Table 2. English allows P-stranding under *wh*-movement because this language has the option of not percolating the [+wh] feature of the nominal complement of P onto the PP. When this option is chosen, the interrogative C attracts the closest element, namely, the *wh*-phrase, onto its specifier, deriving the P-stranding configuration. This yields the example in (3a), repeated here as (60a). When the relevant feature is percolated, then the pied-pied counterpart of (3a) results in (3b), repeated here as (60b). English also allows P-stranding under sluicing because the preposition left behind within the PP is elided by the deletion of the TP that contains this constituent. This yields (3c), repeated here as (60c).

(60)a. Who was he talking with?

b. With whom was he talking?

c. Peter was talking with someone, but I don't know (with) who. (Merchant 2001: 92)

4.2.2. Indonesian

Consider now the Indonesian P-stranding paradigm. I have shown in the present chapter that Indonesian allows P-stranding under sluicing but nonetheless disallows this option under regular *wh*-movement, in violation of Merchant's P-Stranding Generalization.

Consider the examples in (5a-c), repeated here as (61a-c).

(61)a.* Siapa yang kamu ber-dansa dengan?

who that you Vz-dance with

'Whom did you dance with?'

b. Dengan siapa kamu ber-dansa?

with who you Vz-dance

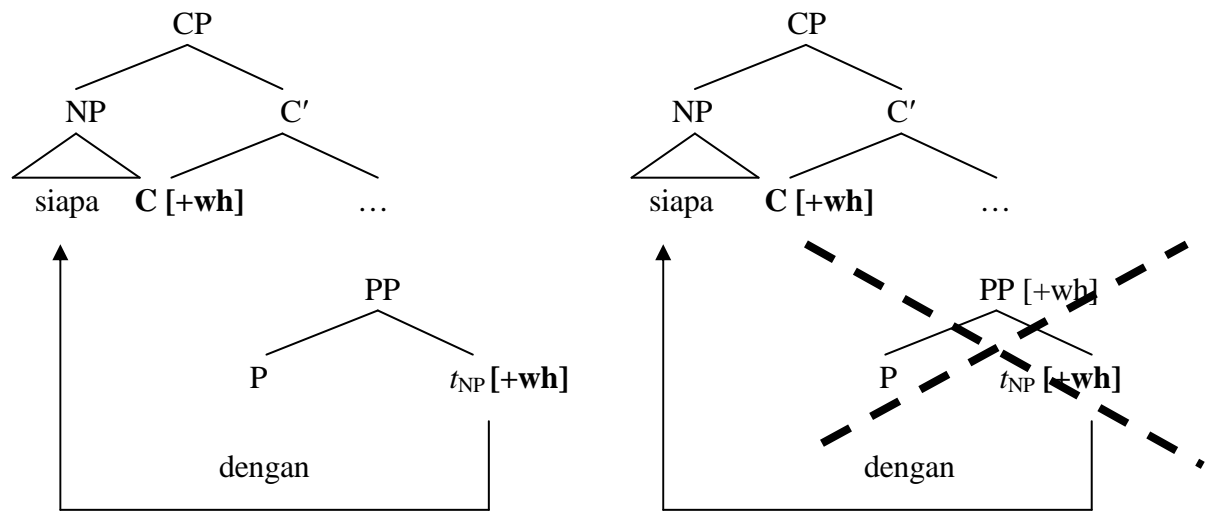
'With whom did you dance?'

- c. Saya ingat Hasan ber-dansa dengan seseorang, tapi saya tidak
 I remember Hasan Vz-dance with someone but I Neg
 tahu (dengan) siapa.
 know (with) who
 ‘I remember Hasan danced with someone, but I don’t know (with) whom.’

Indonesian does not allow P-stranding since the [+wh] feature of the nominal complement of P must percolate onto the PP. As a result, the PP, which is closest to the interrogative C, is attracted to its specifier. The P-stranding example in (61a) is thus deemed ungrammatical. The question is, then, why P-stranding does not yield ungrammaticality under sluicing, as illustrated in (61c).

It is at this point that the role of the syntax-external phonological system plays an important role in remedying imperfections created by syntactic computation. Consider the schematic derivation in (62) for the P-less sluice in (61c).

(62)a. Syntax-Phonology Interface (no repair) b. Syntax-Phonology Interface (repair)

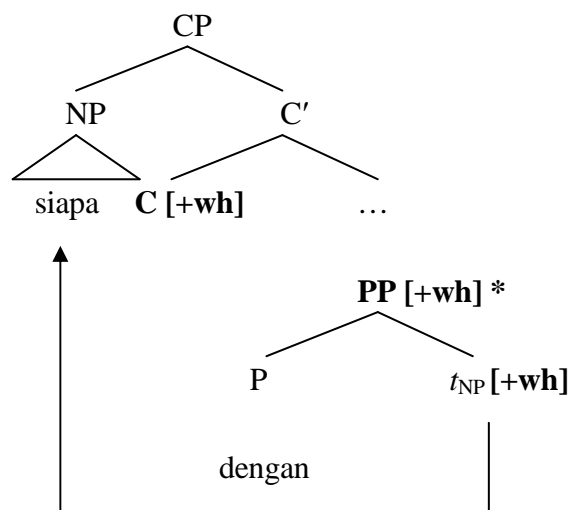


Let us propose that *a failure of the [+wh] feature to percolate* is repaired at the syntax-phonology interface and that a representational constraint to verify percolation rules out the offending PP at the interface.¹⁸ If the offending PP remains at PF, the representational constraint is violated, as shown in (62a). If the offending PP is deleted at the interface as shown in (62b), however, the representational constraint has nothing to apply to. Thus, the failure of percolation is repaired. One could, of course, think of an alternative formulation of the Indonesian data where what is repaired is minimality violations; when the NP moves to [Spec, CP], this movement violates the minimality constraint because the PP is a closer

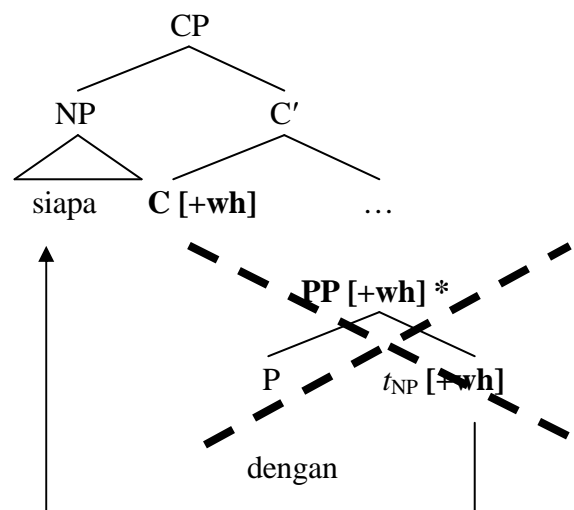
¹⁸ I thank Heidi Harley (personal communication) for suggesting this idea and useful discussion.

element to C than the NP. Under this alternative, the derivation of the Indonesian facts would be as in (63a, b).

(63)a. Syntax-Phonology Interface (no repair)



b. Syntax-Phonology Interface (repair)

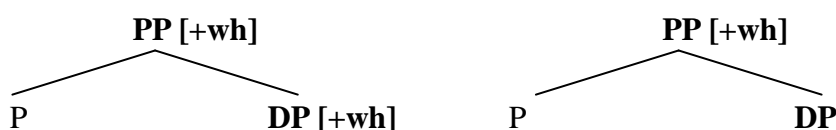


Within the first analysis, the representational constraint to verify percolation is violated but is repaired at the interface. Within the second analysis, minimality violations occur but are repaired at the interface. There is a number of ways to choose between the two analyses. First, if minimality is a derivational constraint and if derivational constraints are syntax-internal and inviolable, then the failure-of-percolation approach to interface repair is empirically superior to the minimality-based approach because the minimality-violating movement should not be able to occur in the first place. It is true that a growing body of work in recent minimalist research, as

mentioned in section 4.1, accumulates evidence that “true” minimality violations such as subjacency violations are *in principle tolerable* within syntax and only representationally illegitimate at the interface. But, crucially, the violation caused in (63a, b) seems to be of a different sort than subjacency; rather, it is a superiority-type/A-over-A-type violation (Chomsky 1964). As we saw earlier, Boeckx and Lasnik (2006) claim that superiority violations are inviolable derivational constraints within syntax. Thus, the alternative analysis of P-stranding in Indonesian would lead to the conclusion that superiority violations behave as representational constraints in Indonesian but behave as derivational constraints in Serbo-Croatian, which is a highly undesirable conclusion in the view of learnability. Second, for the minimality-based analysis of Indonesian to go through, it must be the case that feature percolation is feature copying; even after the percolation takes place, the [+wh] feature should still remain on the *wh*-phrase, for otherwise minimality violations would not occur in the first place. This original formulation of percolation (percolation-as-copying approach) was illustrated earlier in (46b), repeated here as (64a). True, Lieber (1980) and Williams (1981) provide cases based on the alternation of strong verbs such as *stand-stood* and *withstand-withstood* in which feature percolation could be construed as feature copying in morphology, it is not clear whether comparable phenomena exists in syntax to support this copying mechanism. This problem, however, does not arise if it is a failure to percolate that is repaired,; feature copying is not

necessary component of percolation. This alternative idea (percolation-as-percolation approach), which I adopt, is illustrated in (64b).

(64)a. Percolation as Copying Approach b. Percolation as Percolation Approach



These two considerations, therefore, suggest that the failure-of-percolation analysis is superior to the minimality-based analysis on both empirical and conceptual grounds.

The proposed repair-based analysis of the Indonesian P-stranding pattern is quite close in spirit to Lasnik's (1995, 1999, 2001) analysis of pseudogapping that also draws on the idea of "repair-by-deletion"; see section 4.3 for further discussion of his analysis. Lasnik argues that the overt V raising to *v* that is obligatory under non-elliptical contexts in English does not happen in pseudogapping constructions in English precisely because the constituent (namely, VP) that contains the violation is eliminated by PF deletion.¹⁹ Therefore, the

¹⁹ As Heidi Harley (personal communication) pointed out, the analogy here is not completely parallel because what is repaired at the PF interface is different between the two analyses; it is failure of feature percolation that is repaired within the present analysis whereas it is the defectiveness of verbs (i.e. verbs with scattered features) that is repaired within Lasnik's analysis. See Chomsky (1995) and Lasnik (1995, 1999, 2001) for details on why verbs with scattered feature cause crash at the interface.

present analysis provides a natural explanation for the apparently atypical P-stranding pattern under sluicing in Indonesian from the interaction of independently motivated assumptions concerning feature percolation and interface repair strategies.

4.2.3. French

Let us finally consider why French does not allow P-stranding under *wh*-movement or sluicing. The examples that illustrate this pattern are repeated here as (65a-c) from (4a-c).

(65)a.* Qui est-ce qu' elle l'a offert à?

who Q she it-has offered to

'Whom has she offered it to?'

b. À qui l'a-te-ell offert ?

to whom it-has-she offered

'To whom has she offered it?'

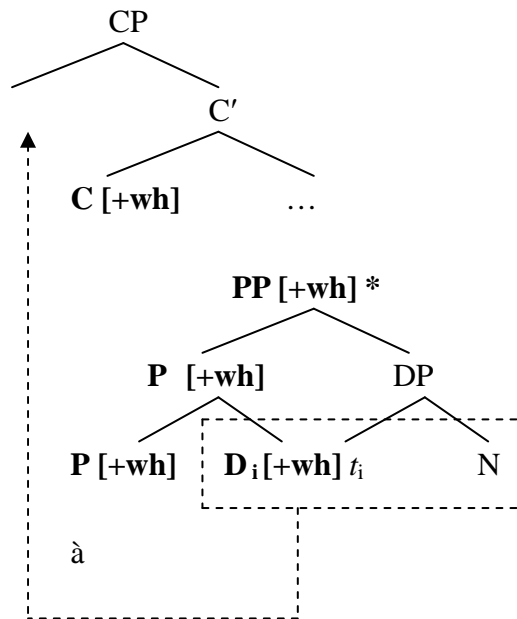
c. Anne l'a offert à quelqu'un, mais je ne sais pas *(à) qui.

Anne it-has offered to someone but I Neg know not to whom

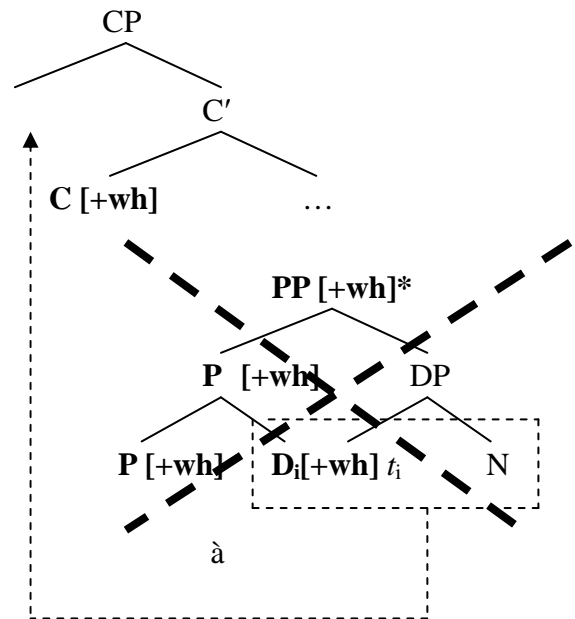
'Anne has offered it to someone, but I don't know (to) whom.'

French does not permit P-stranding under regular *wh*-movement as in Indonesian because the [+wh] feature obligatorily percolates onto the PP that dominates the *wh*-phrase. As shown in the example in (65c), French also does not allow P-less sluices. What is crucial here is that languages such as French have D-to-P syntactic incorporation, as we saw earlier in section 4.1. Consider the derivation in (66a, b) for the unacceptable P-less sluice in (65c).

(66)a. Syntax



b. PF



In the derivation in (66a), the D head undergoes syntactic incorporation into the P in accordance with Law's (1998, 2006) constraint given in (49). The [+wh] feature of the moved D head percolates into the dominating PP, which is thereby now marked with the

same feature. This derivation crashes, because, when the interrogative C with the [+wh] feature attracts the element with the matching feature, the DP is no longer a constituent that includes the D, hence cannot be attracted by C. Notice that the P-less sluice could be potentially derived if the D head underwent excorporation to be attracted by the C head. However, this possibility is blocked because the excorporation would cause both the minimality violation on attraction (because the PP is closer to C than the D head) and the ECP-like violation (because the trace of the excorporating element cannot be properly licensed). The point here is that whatever derivation would possibly yield the P-stranding sluice in French crashes because of the interaction of independently motivated *syntactic constraints* on D-to-P incorporation. Thus, when the derivation in (66a) reaches the phonological component, as shown in (66b), it is too late to repair violations associated with D-to-P incorporation within this derivation by deletion because the violations are within syntax. This pattern is different from that in Indonesian because this language does not have D-to-P incorporation, as the lack of D-to-P coalescence shows.

It is clear, then, that the notion of interface repair by way of deletion plays a crucial role in the proposed account of the three-way contrast between English, Indonesian, and French with respect to P-stranding. The most important point of the proposed analysis, therefore, is that the syntax-external phonological component can repair certain illicit configurations

created in syntax by deleting them but not all configurations: it cannot undo mistakes concerning the D-to-P coalescence that are syntactically derivationally conditioned. The present analysis, therefore, provides powerful support for the idea behind the thesis of minimalist interface, namely, that interface components can conduct domain-specific operations they avail of to repair certain syntactic imperfections but only within the parametrically defined curve set by syntax.

4.3. *Pseudogapping in Indonesian and French*

The proposed analysis makes certain interesting predictions regarding pseudogapping in Indonesian and French. The proposed analysis argues that the P-stranding is tolerated in Indonesian only under sluicing because the offending part of syntactic structure (namely, the PP that records the violation of the minimality constraint on movement) is removed by deleting it at the syntax-external phonological component. The proposed analysis, thus, makes a prediction that not only deletion of TP but also deletion of smaller constituents than TP would also have the ameliorating effect.²⁰ One construction in point in Indonesian is pseudogapping. Following the analysis of pseudogapping in English proposed by Lasnik (1995, 1999, 2001) (see also Jayaseelan 1990 and Merchant 2008), let us assume that the

²⁰ Thanks to Heidi Harley (personal communication) for pointing this out.

focused element undergoes movement into the specifier of a higher projection FP (such as AgroP in Lasnik's analysis or the ν P-internal focus position) that dominates the VP: a pseudogapping construction arises when the VP remnant is deleted at PF. The derivation of Indonesian pseudogapping examples with the focused object DP as in (67a) is shown in (68), ignoring the structure above FP that would contain the TP whose specifier the external argument *Hasan* moves and the split ν P configuration. (67b) shows that the PP remnant in pseudogapping is also grammatical.²¹

(67)a. Esti ber-danca dengan Fernando dan Hasan [_{DP} Rifi]_i [_{VP} ~~ber-danca dengan~~ *t_i*].

Esti Vz-dance with Fernando and Hasan Rifi

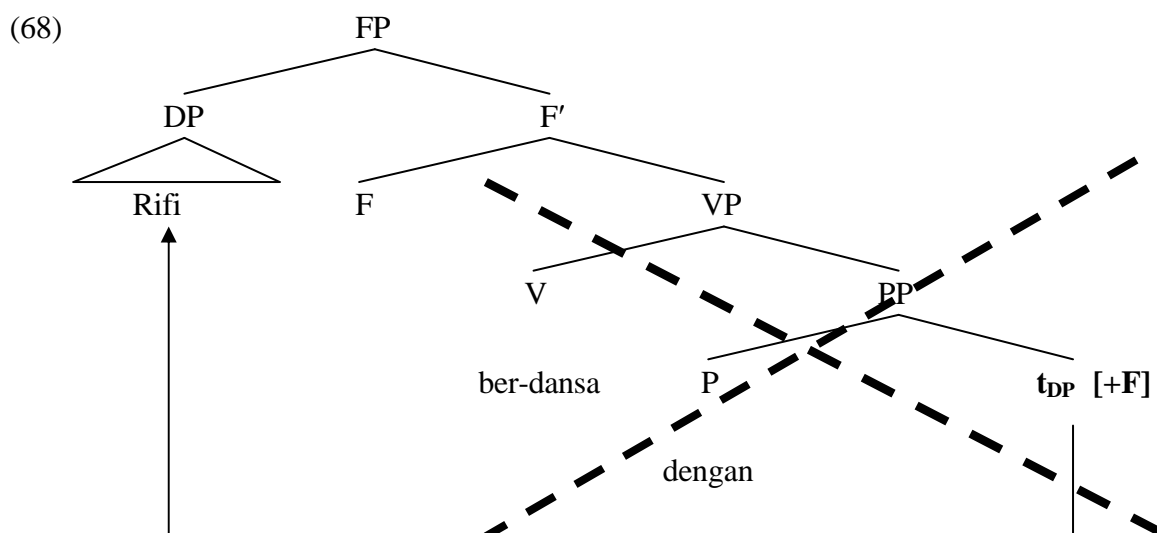
'Esti danced with Fernando, and Hasan (danced with) Rifi.'

b. Esti ber-danca dengan Fernando dan Hasan [_{PP} dengan Rifi]_i [_{VP} ~~ber-danca~~ *t_i*].

Esti Vz-dance with Fernando and Hasan with Rifi

'Esti danced with Fernando, and Hasan (danced) with Rifi.'

²¹ The nature of the functional projection into whose specifier the remnant moves does not matter for the purposes of the present discussion. It might be AgrOP, AspP, or ν P-internal FocP. Thus, I notate it here as FP.



In this derivation, the object of the preposition undergoes movement/object shift into the specifier of the FP. The VP portion of this structure undergoes deletion at PF. This sequence of operations yields the P-less pseudogapping construction in (67a). Note that this derivation yields the P-stranding violation that is active in Indonesian. To achieve this result, let us suppose that there is a feature such as [+focus] that must obligatorily percolate from the focused NP onto the dominating PP and violation of this condition is recoded in the PP that dominates the launching site of the movement, as argued independently by Lasnik (1999, 2001). Then, the movement illustrated in (68) violates the representational interface constraint that detects the failure of percolation if no deletion of the offending PP applies. The grammaticality of the example in (67a), then,

clearly indicates that the relevant violation can be repaired if constituent that contain the PP (namely, VP) is eliminated by deletion at the phonological component. The acceptability of the pseudogapping without the preposition, thus, provides further confirmation for the proposed repair-based approach to Indonesian P-stranding.²²

Note that the current analysis predicts that P-less pseudogapping should be ungrammatical in French because it has the D-to-P syntactic incorporation that cannot be ameliorated by interface repair. This prediction is indeed borne out by the contrast between (69a) and (69b).²³

(69) a. * Jean a danca avec Marie et Robert [DP Suzanne]_i [VP ~~danse avec~~ *t_i*].

Jean has danced with Marie and Robert Suzanne

‘Jean has danced with Marie, and Robert (danced with) Suzanne.’

b. Jean a danca avec Marie et Robert [PP avec Suzanne]_i [VP ~~danse~~ *t_i*].

Jean has danced with Marie and Robert with Suzanne

‘Jean has danced with Marie, and Robert (danced) with Suzanne.’

²² This analysis predicts that languages which allow P-stranding under pseudogapping should also behave like Indonesian with respect to sluicing. Javanese is certainly one language of this type. Whether this prediction holds also true for other Austronesian languages closely related to Indonesian (Madurese, Balinese, etc.) is an important task to be undertaken in my future research.

²³ Thanks to Summaya Racy (personal communication) for her help in constructing the data in (69a, b).

I take this difference in P-stranding under pseudogapping between Indonesian and French to provide further confirmation for the proposed parametric theory of P-stranding.

5. Comparison of the Proposed Analysis with Fortin's (2007b) Analysis

In this section, I compare the proposed analysis of the P-stranding pattern in Indonesian with the most recent alternative presented by Fortin (2007b). As mentioned in the first section of this chapter, Fortin shows with me that the P-stranding pattern in Indonesian presents a genuine counterexample to Merchant's P-Stranding Generalization. I show, however, that there are several serious empirical shortcomings within her analysis that are successfully resolved under the present interface-based approach to P-stranding.

5.1. Fortin's (2007b) LF Copy + Long Distance Agree Analysis

Fortin's (2007b) analysis follows the tradition of the LF Copy theory of sluicing as originally proposed by Chung et al. (1995) but presents a refinement of their original theory within the Minimalist Program. She argues that the sluice is a deficient syntactic structure with no TP. The *wh*-remnant is base-generated in the specifier of CP, with the sluiced clause being supplied semantic content by copying the antecedent TP into the TP

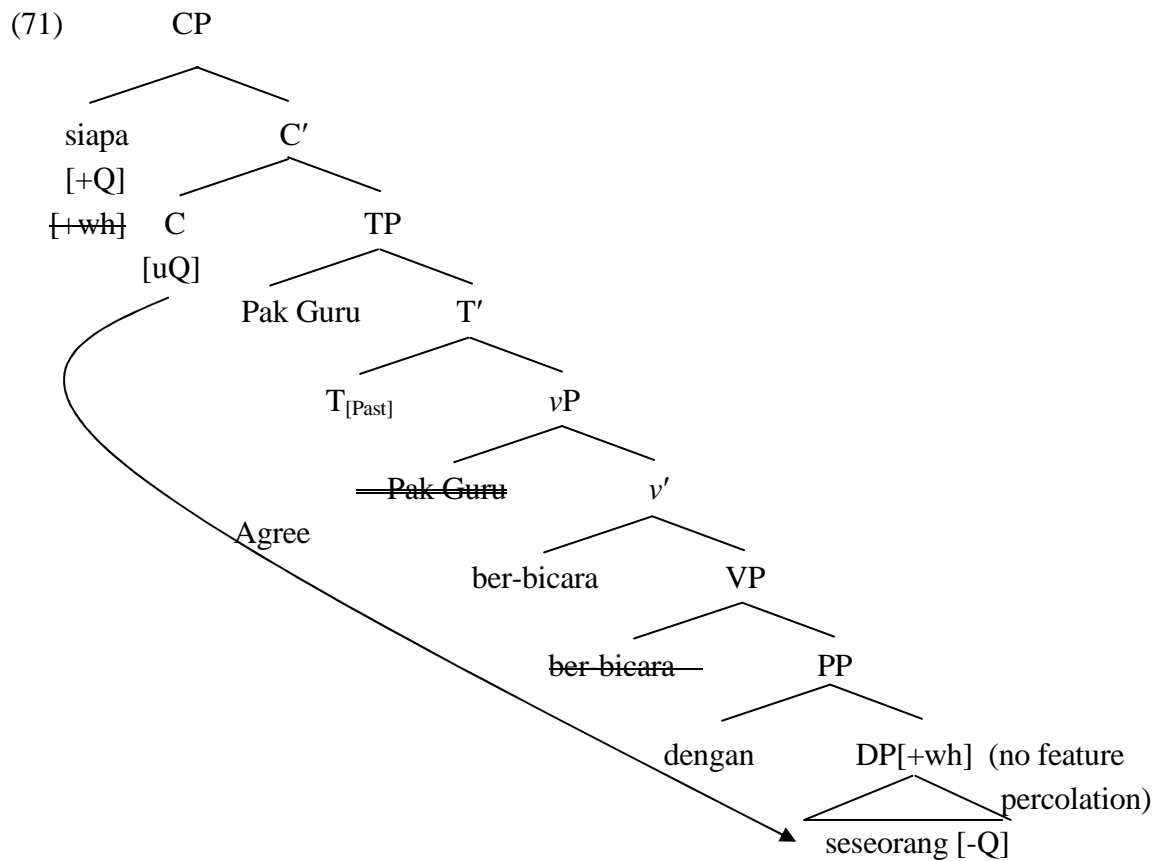
part of the sluice via sideward movement in the sense of Nunes (2004).²⁴ To illustrate the specifics of her LF-Copy analysis, consider the syntactic derivation of the example in (70) that involves a DP-remnant, as shown in (71).

(70) Pak Guru ber-bicara dengan seseorang, tapi saya tidak tahu **siapa**.

Mr. teacher Vz-speak with someone but I Neg know who

‘Pak Guru spoke to someone, but I don’t know who.’ (Fortin 2007b: 326)

²⁴ There are several assumptions made by Fortin such as her treatment of sprouted arguments. Since those assumptions are not germane to the discussion, I just mention only those assumptions relevant to P-stranding.



(Fortin 2007b: 327)

In this derivation, the *wh*-phrase *siapa* ‘who’ is base-generated in the specifier of the CP. The TP is reconstructed by copying the antecedent TP into the derivation via sideward movement. The probe C, then, agrees with the NP *seseorang* ‘someone’. Fortin assumes that there is a three-way matching relation between the *wh*-phrase in [Spec, CP], the interrogative Q, and the indefinite NP, such that the syntactic category of the *wh*-phrase delimits the syntactic category of the goal NP that C agrees with. In the derivation above,

the syntactic category of the indefinite matches with the *wh*-phrase in [Spec, CP] because the [+wh] feature does not percolate onto PP. The P-less sluice in examples as in (70), thus, can be naturally accommodated under her analysis.

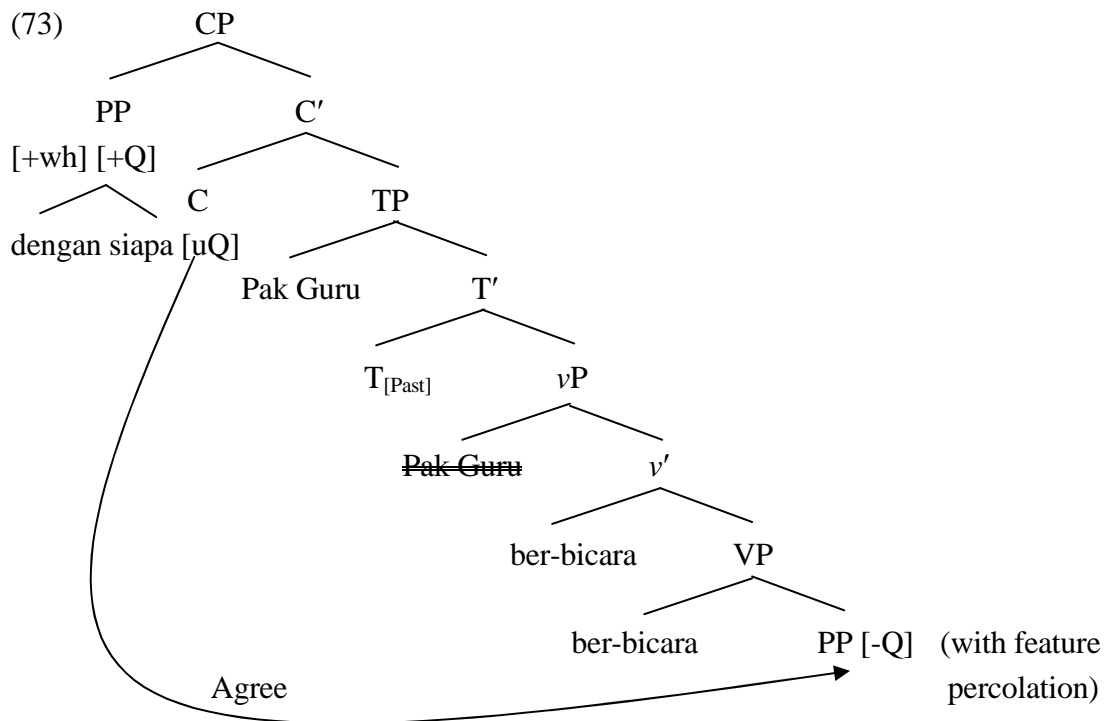
Consider now the derivation of the sluice with the PP remnant in (72), as shown in (73).

(72) Pak Guru ber-bicara [PP dengan seseorang], tapi saya tidak tahu **dengan siapa**.

Mr. teacher Vz-talk with someone but I Neg know with who

‘Pak Guru spoke to someone, but I don’t know with who.’

(slightly modified from Fortin 2007b: 327)



In this derivation, the prepositional *wh*-phrase *dengan siapa* ‘with whom’ is base-generated in the specifier of the CP, with the empty TP being supplied with semantic content by copying the antecedent TP into the empty TP of the sluice. Again, the three-way matching requirement between the interrogative PP, the goal C, and the indefinite PP in the recycled TP requires that the correlate match the interrogative PP in syntactic category. This result is achieved here by the percolation of the [+wh] feature onto the PP, as also assumed in the interface-based approach proposed in the previous section, so that the whole PP becomes the closest goal from the perspective of the probe C.

5.2. *Four Problems with Fortin’s (2007b) Analysis*

There are three serious empirical difficulties, some of them noted by Fortin herself, which show that her LF Copy + Agree-based analysis is not the right way to understand P-stranding under sluicing. The first problem concerns the Case checking of the *wh*-phrase base-generated in [Spec, CP], which has been shown to be a perennial difficulty for the general LF Copy Theory of sluicing as in Chung et al.(1995) by Merchant (2001: 151). Under the standard minimalist assumption that the Case feature of a *wh*-phrase is checked via Agree with an appropriate functional head such as *v* and T, Fortin’s analysis would not be able to provide a natural explanation for how the Case of the base-generated *wh*-phrase can be checked. One

could get around this problem by introducing new mechanisms such as Case Transmission but that is clearly an ad hoc assumption unless it is given independent motivation.

Second, Ross (1969) and Merchant (2001) observe that, in many languages with rich case morphology, the *wh*-remnant in a sluice is marked for the case that it would bear in a corresponding, non-elliptical *wh*-question. As Ross originally notes, this case-matching effect receives a straightforward account if the sluicing is derived from the regular *wh*-question, as in the movement + deletion approach adopted in this chapter. It is not clear, however, whether Fortin's analysis derives this result as naturally as the deletion analysis.

The third problem concerns the percolation of the [+wh] feature of the complement of a preposition onto the PP. As in the proposed analysis, Fortin assumes, following Chomsky (1972) and Stepanov (2001), that the [+wh] feature may or may not percolate depending on languages. As is clear from the derivations in (71) and (73) above, she assumes that the feature percolation is optional in Indonesian, a view that the proposed analysis does not support, as shown in Table 3. However, Fortin's assumption overgenerates. For example, in the derivation in (71), the *wh*-feature does not percolate whereas in the derivation in (73), it does. The question here is what prevents the percolation in (71) but not in (73). To put the same question in a different way, if the feature percolation were truly optional in Indonesian, we would predict that the P-stranding should be acceptable under non-

elliptical *wh*-questions in Indonesian because nothing in Fortin's analysis seems to necessitate the percolation. The present analysis, by contrast, is sufficiently constrained in this regard: the [+wh] feature *always* percolates onto PP in Indonesian whether the construction in question is a non-elliptical *wh*-question, sluicing, or pseudogapping.

Based on these four considerations, I conclude that the present interface-based approach is superior to Fortin's LF Copy + Agree-based approach on empirical grounds.

5.3. *Derivationalism vs. Representationalism*

It is interesting in this connection to think about Fortin's primary motivation for pursuing the LF Copy analysis explicated above. It comes from the idea, as expressed by Chomsky (1995, 2000, 2001, 2004, 2005), that the shortest movement /minimal search requirement is included in the definition of Move/Attract: movement of an element to some position over some other closer position, where closeness is defined in terms of asymmetric c-command, is simply impossible in syntactic computation. Note that this idea is in striking contrast with the idea behind Merchant's (2001) repair theory of PF island violations (see also Lasnik 1999, 2001, 2005, 2007). This is because, for a PF island to be ameliorated at all, the syntax should in principle allow violations of the minimality constraint. Indeed, this kind of representational view of subjacency violation has been adopted since Ross (1969)

and Chomsky (1972) and still around in recent work as in Lasnik and Saito (1992), Chomsky and Lasnik (1993), Merchant (2001), and Lasnik (1999, 2001, 2005, 2007). The amelioration-based account of the island effects, therefore, seems to show that subjacency violations are *tolerable* in syntax, in contrast to Chomsky's own idea that syntactic computation embedded shortest move/minimal search into the definition of Move. Nonetheless, the proposed account could still be formulated in derivational terms. For example, Richards (2001) proposes the *Principle of Minimal Compliance*; once syntactic constraints such as subjacency are satisfied by movement in a certain stage of the derivation, violations of the same constraints in later stages of the same derivation can be tolerated.²⁵ Thus, whether subjacency should be best formulated in (fully) representational or (strictly/partially) derivational terms is far from settled. See chapter 7 for further discussion on this debate.

The debate between derivational and representational theories of syntax has been active since the advent of generative enterprise and teasing apart two theories on empirical grounds has proven quite subtle, depending on complex considerations of extremely intricate data. To the extent that the proposed analysis of the typology of P-stranding is tenable, however, our

²⁵ Thanks to Andrew Carnie (personal communication) for directing my attention to Richards' (2001) *Principle of Minimal Compliance* in this context.

current investigation indicates that certain violations such as subjacency, minimality, etc. are tolerated and repaired at the interface whereas other violations such as the failure of the syntactic D-to-P incorporation, superiority violations, etc. are not. Thus, it seems that both derivational and representational constraints are needed to describe and explain the full range of data in natural language. See also Aoun and Li (2003) for much relevant discussion.²⁶

6. Conclusions

Let us summarize the content of this chapter and consider the implications of the idea proposed here for the syntax-phonology interface and the thesis of minimalist interfaces.

6.1. *Chapter Summary*

In the current chapter, I have observed that the P-stranding pattern in Indonesian presents the first genuine challenge to the P-Stranding Generalization established by Merchant (2001) in support of his movement + TP deletion analysis of sluicing. I have presented arguments based on the distribution of the question particle *-kah* and the obligatory lack of the complementizer *yang* in non-nominal *wh*-questions that the source of PP-sluicing in this language is via regular *wh*-movement, as in English, rejecting alternative accounts of

²⁶ Thanks to Simin Karimi (personal communication) for useful discussion on the debate and bringing my attention to the work by Aoun and Li (2003).

wh-questions as reduced clefts as in Cheng (1991) or headless relative clauses as in Cole et al. (to appear). This observation is important because, to the extent that it is correct, the P-stranding pattern in Indonesian provides the first genuine counterexample to the P-Stranding Generalization.

Based on this result, I develop a new theory of the typology of P-Stranding that crucially relies on the recent idea of interface repair, coupled with two independently motivated assumptions concerning the percolation of the [+wh] feature of the nominal complement of the preposition onto the PP and the D-to-P coalescence as the result of syntactic D-to-P incorporation. I have demonstrated that the three-way contrast in P-stranding between English, Indonesian, and French receives a straightforward explanation in a way that the apparently atypical pattern of Indonesian sluicing is naturally accommodated. I have also shown that the current analysis makes a correct prediction that deletion of smaller categories than TP should also ameliorate P-stranding violations in Indonesian but not in French. I have shown that this prediction is verified by the grammaticality of the pseudogapping construction with the DP remnant in Indonesian and the ungrammaticality of the corresponding construction in French.

I have compared the proposed analysis with the most recent alternative LF copy analysis of P-stranding pattern in Indonesian by Fortin (2007b) and shown that there are

several empirical shortcomings concerning the case-matching effects documented in Ross (1969) and Merchant (2001), the Case checking of the *wh*-remnant base-generated in the specifier of CP, the optionality of the [+wh] feature percolation, and the effects of the morphophonology of certain prepositions on the reparability of the P-less sluice, that are successfully resolved under the current analysis.

6.2. *Implications of Indonesian Sluicing for the Theory of the Syntax-Phonology Interface*

The proposed analysis of sluicing and P-stranding in this chapter provides strong evidence that syntax *could make* certain derivational mistakes and some of them can be repaired by operations such as deletion at the syntax-external phonological component. It is important that only certain mistakes such as failure of percolation and subjacency-type violations can be repaired; some others such as the failure of the syntactic D-to-P incorporation cannot.

One important question is why the PF component is not omnipotent in that it can repair both the failure of D-to-P incorporation and the minimality violation. I believe that one answer to this question is available once we take seriously the interaction between syntactic computation, the PF computation and the A-P system. The representation that the A-P system receives from the PF interface would not contain information about syntactic violations of

minimality, since the A-P cannot read such language-specific information; it would be simply a string of words as in #I#wonder#what#who#bought, with accompanied information on intonation and stress. Then, violations such as failure of feature percolation/subjacency *must be* repairable by deleting relevant information at the PF interface before they reach the A-P system. On the other hand, the failure of the D-to-P incorporation often has repercussions in the form of coalescence (as in Romance). Thus, the effects of such a syntactic operation are easily detectable in pronunciation and orthography of the D-P combination, and hence relevant to the working of the A-P system; failure to apply this operation in syntax necessarily turns out to cause failure in the proper communication between the phonological component and the A-P system. Then, the failure of D-to-P incorporation should not be repairable at the PF interface since coalescence and cliticization, which must be reflected at the A-P system, are syntactically conditioned.

The analysis proposed in this chapter is one clear demonstration of how the thesis of minimalist interface works; interface components do whatever domain-specific operations they avail of to make them legible/usable for the language-independent sound system (A-P system in Chomsky's terms) but only within the range of options that is parametrically set by a particular language. I discuss further ramifications of the repair-based theory of P-stranding in chapter 7.

CHAPTER 4 THE MORPHOSYNTAX OF BARE NOMINALS IN INDONESIAN AND JAVANESE: A RELATIVIZED PARAMETRIC THEORY OF NOMINAL DENOTATION¹

1. Introduction

This chapter discusses the issue of syntax-semantics interface with a case study in the relation between the denotation and morphosyntax of bare nominals in Indonesian and Javanese. Though the first part of this chapter is devoted to the descriptive analysis of bare nouns in these two languages, it ultimately becomes clear in the rest of the chapter that the primary goal of this chapter is to construct a universal theory of the denotation of bare nouns within the Principles & Parameters approach to language variation (Chomsky 1986a, 1995).

In a series of his recent work (Chierchia 1998a, 1998b), Gennaro Chierchia proposes a semantic parameter, known as the *Nominal Mapping Parameter*, which states that languages differ in terms of what they allow their bare nouns to denote in the syntax-semantics mapping, namely kinds ([+arg, -pred] languages), properties ([-arg, +pred] languages), or both ([+arg, +pred] languages) under certain conditions. Chierchia argues that setting of this parameter uniquely determines the morphosyntactic profile of bare

¹ This chapter is a substantially modified version of the paper read at the Inaugural Meeting of the Arizona Linguistics Circle (ALC1) held at the University of Arizona, Tucson and the Mid-America Linguistic Conference (MALC) held at the University of Kansas, Lawrence. The papers read at those two meetings are to appear in Sato (in press d, forthcoming b).

nominals in a particular language with respect to the availability of bare arguments, the generalized classifier system, and plural morphology. This claim, therefore, includes a strong statement that all natural languages can be classified as one of three and only three languages with respect to the denotation of bare nominals.

In this chapter, I provide arguments, modeled after Chung (2000), that Indonesian and Javanese do not fit into any one of the three language types identified under Chierchia's semantic typology. I further argue that the very notion of "semantic parameter" as varying the nature of the mapping between the syntactic and semantic representation of a bare noun across languages is hardly groundable within the standard generative conception of the locus of parameter as restricted to properties of the lexicon (Chomsky 1995) or, more precisely, properties of functional categories (Borer 1984; Fukui 1986, 1995).

Based on these results, I propose an alternative, relativized parametric theory of the effects of the Nominal Mapping Parameter, whereby languages differ in terms of the complexity of nominal functional structures each language allows (i.e. DP>QP>ClP>NumP>NP) and in terms of the possible set of values the Num Head in each language can take (i.e. {singular, plural} or {neutral, plural}). Neither of these ideas is anything new. The idea that languages differ in terms of the complexity of nominal structure has been already proposed on various grounds in recent work as in Grimshaw

(1991, 2005), Massam (2001), Guilfoyle and Noonan (1992), and Vainikka (1993/1994).

The observation that languages differ in their possible set of Num values has been recently made in extensive studies of the number system in languages such as Malay and Indonesian as in Carson (2000) and Chung (2000). The results achieved in this chapter, therefore, provide further evidence for these proposals.

Given the Subset Principle of Wexler and Manzini (1987), the proposed parametric theory of bare nominals makes an interesting prediction that early acquisitional stages of all languages should look like Indonesian and Javanese since these languages instantiate the simplest nominal structure (NumP) under the proposed theory. I show that this prediction is borne out by the so-called *telegraphic speech* produced by English-learning children, at which English nouns share several fundamentally morphosyntactic properties with bare nouns in Indonesian and Javanese. The proposed analysis also sheds light on the recent debate on the nominal structure of Slavic languages such as Russian. Our discussion of the morphosyntax of bare nouns in Russian shows that bare nouns in this language only project up to NumP as do Indonesian and Javanese bare nouns and leads us to conclude that the universal DP hypothesis for article-less languages is incorrect.²

² The proposed parametric theory of nominal denotation contrasts sharply with the proposal made by Gil (2005). Gil argues that Riau Indonesian, the variety of Indonesian spoken by local people in east-central Sumatra, has no distinction between lexical and functional categories and capture its many correlates of a verb-initial language by the interaction of the single principle (heads precede modifiers) with iconicity and

I also discuss the broader theoretical consequences of the proposed analysis of bare nominals across languages by situating it within the context of the syntax-semantics interface, a central subject matter of this dissertation. In particular, empirical findings in this chapter provide strong support for the notion of “minimalist interfaces”, whereby the syntactic computation provides a parametrically defined curve that the universal semantics interface blindly follows, without any extrinsically determined mapping between the syntax and semantics of a particular expression as in Chierchia’s (1998a, b) Nominal Mapping Parameter.

The present chapter is organized in the following manner. In the next section, I review in detail the idea behind the Nominal Mapping Parameter proposed by Chierchia (1998a, b). In section 3, I turn to the denotation and morphosyntax of bare nominals in Indonesian and Javanese. The purpose of this section is empirical. I provide evidence that neither of these two languages fits into any one of the three language types under Chierchia’s semantic classification. I discuss Chung’s (2000) analysis of Indonesian bare nouns in

information flow. It is not clear whether Gil’s radical analysis applies to the dialect of Indonesian discussed in this dissertation. Although this chapter argues that this dialect represents the simple nominal functional structure (NumP) among languages of the world, it is an SVO language, unlike many dialects in Sumatra such as Riau Indonesian and Toba Batak, as argued for convincingly in Chung (to appear). Thus, whether Gil’s analysis could even apply to the SVO dialect under investigation is far from clear and needs extensive discussion that goes beyond the scope of this chapter. Thanks to Andrew Carnie (personal communication) for directing my attention to Gil (2005).

section 3.1. Though Indonesian allows bare arguments and has plural morphology, I argue, against Chung, that Indonesian is not a classifier language. I turn to parallel facts in Javanese in section 3.2, where I show that Javanese behaves as Indonesian in all relevant respects; it allows bare arguments, has plural morphology, and lacks a generalized classifier system. These empirical findings strongly indicate that the Nominal Mapping Parameter imposes too tight a mapping between the syntax and semantics of bare nouns in natural language and falsely exclude the morphosyntactic patterns of bare nominals in Indonesian and Javanese that are actually attested. These results lead us to pursue a different account of what Chierchia's Nominal Mapping Parameter is supposed to capture.

In section 4, starting with the observation that the notion of "semantic parameter" as constraining the mapping between the syntax and semantics of nominals is at odds with the standard view of the locus of parameter, I propose a novel, relativized parametric theory of the denotation of bare nominals across languages. According to this theory, languages differ in two dimensions: how high a language allows its bare noun to grow and which set of feature values each language allows its Num head to choose. I show that the morphosyntactic profile of bare nouns in languages such as Javanese, Indonesian, Japanese, English, and Italian is naturally derived from the combination of these two parameters. In section 5, I consider a number of predictions made by the proposed theory of bare

nominals. The proposed analysis combines with the Subset Principle proposed by Wexler and Manzini (1987) to predict that initial stages of all languages should go through stages where bare nominals behave similarly to those in Indonesian and Javanese. Drawing on the data and analysis presented by Guilfoyle and Noonan (1992) and Vainika (1993/1994), I show that this prediction is indeed borne out by the so-called *telegraphic speech* produced by English-learning children. I further show that the proposed analysis sheds a new light on the debate concerning the nominal syntax of Slavic languages such as Russian. Specifically, the proposed analysis, if correct, indicates that Russian nominals project only up to NumP as Indonesian/Javanese nominals and that the universal DP hypothesis for article-less languages needs to be seriously reconsidered. In section 6, I summarize major theoretical implications of the parametric theory presented here as they relate to the theory of syntax-semantic interfaces, a central subject matter of the dissertation.

2. Chierchia's (1998a, b) Nominal Mapping Parameter

The Nominal Mapping Parameter, recently proposed by Chierchia (1998a, b), claims that languages differ in terms of what they let their bare nouns denote in the syntax-semantics mapping: kinds, properties, or both under certain conditions. Chierchia claims that the setting of this semantic parameter serves to uniquely identify syntactic and morphological properties

of a bare noun in a given language. Chierchia's Nominal Mapping Parameter, which identifies three and only three language types, includes a strong statement that all languages should belong to one of them. In the first type of languages such as Chinese and Japanese, which he calls [+arg, -pred] languages, bare nouns are mapped onto kinds (type <e>). A *kind* is defined by Chierchia (1998a: 349) as "function[s] from worlds (or situations) into pluralities, the sum of all instances of the kind." This has three morphosyntactic consequences. First, since a kind is saturated in the Fregean sense, this type of language allows bare nominal arguments. Second, languages of this type also lack plural morphology for the following reason. As Chierchia (1998a: 351) puts it, "Fido is as good an instance of the dog-kind as Fido and Barky are. This means that the property corresponding to a kind comes out as being mass." The notion of *mass*, in turn, is defined in Chierchia (1998a: 347) to "come out of the lexicon already pluralized... a mass noun, such as, say, *furniture*, will be true in an undifferentiated manner of singular pieces of furniture as well as of pluralities thereof... quite literally the neutralization of the singular/plural distinction." Since a kind is essentially mass in that it cannot differentiate between singular and plural instances of a kind, [+arg, -pred] languages should not have plural morphology; mass terms, being "pre-pluralized", cannot be further pluralized. In other words, the extension of all bare nouns ends up being mass in this type of language. Third, [+arg, -pred] languages also develop a

generalized classifier system. This is because kinds cannot be individuated and hence need an appropriate counting level for each bare noun. Note that all the three morphosyntactic characteristics of bare nouns observed here are the automatic consequence of the denotation of a bare noun as a name of a kind.

The second type of language is what Chierchia (1998a, b) calls [-arg, +pred] language such as Italian and French, where bare nouns are mapped onto properties (type $\langle e, t \rangle$). This type of language does not allow bare nominal arguments; they need to be combined with determiners (either covert or overt depending on the language) to be able to serve as a saturated argument of type $\langle e \rangle$.

The third type of language is termed [+arg, +pred] language, and Chierchia mentions English and Russian as examples that belong to this language type. As the setting [+arg, +pred] indicates, this type of language expectedly shows a mixed morphosyntactic profile. It behaves as Chinese and Japanese in that mass and bare plurals are mapped onto kinds (hence [+arg]) whereas it behaves as French and Italian in that count nouns are mapped onto properties (hence [+pred]).

Bare nouns must be mapped onto some semantic type, either kinds, properties, or both, under Chierchia's theory of a semantic parameter. The fourth possibility that bare nouns

are not mapped onto any type (i.e. [-arg, -pred]) is excluded because such a type does not denote anything.

One important aspect of Chierchia's Nominal Mapping Parameter is that, as noted above, a particular setting of the denotation of a bare noun in a language uniquely predicts its morphosyntactic profile. In other words, Chierchia's theory imposes a rigid one-to-one mapping between the morphosyntax and the denotation of a bare noun in a given language. Thus, for example, if a language *L* develops a generalized classifier system, it must be the case that that language also allows bare nominal arguments and lacks plural morphology. This observation is emphasized by Chierchia (1998a: 354), who explicitly states that "for example, a language with the plural-singular contrast and a generalized classifier system is certainly logically conceivable; it could, in principle, exist. The point of view we are adopting offers a seemingly principled way for ruling it out."

3. The Denotation and Morphosyntax of Bare Nominals in Indonesian and Javanese

In this section, I show, based on evidence from Indonesian and Javanese, that these two languages do not fit into any one of the three language types that should exhaust all natural languages under Chierchia's semantic theory. The arguments presented below owe a great deal to and are modeled on those developed by Chung (2000), who argues against the

Nominal Mapping Parameter from facts in Indonesian. Chung shows that Indonesian is a language with bare nominal arguments and a generalized classifier system, but nonetheless does have plural morphology marked via full reduplication of the root. I review her analysis of Indonesian bare nouns below in section 3.1. Though Indonesian does have generalized bare arguments and plural morphology, I argue against Chung (2000) that Indonesian is not a generalized classifier language. In section 3.2, I turn to bare nominals and their morphosyntax in Javanese. I show that Javanese is identical to Indonesian in all criteria pertinent to Chierchia's Nominal Mapping Parameter; It allows bare nominals in all argument positions and has plural morphology marked by full reduplication but does not develop a generalized classifier system of the kind found in Japanese or Chinese.

3.1. *Bare Nominals in Indonesian*

Let us review evidence from Chung (2000) that Indonesian runs counter to the predictions of the Nominal Mapping Parameter. Her basic argument is that, in Indonesian, bare nominal arguments occur freely and numeral classifiers are required under certain conditions like a [+arg, +pred] language, but bare nouns have a singular-plural contrast like a [-arg, +pred] language; this cluster of morphosyntactic characteristics, thus, would remain mysterious under Chierchia's semantic theory.

First of all, Chung notes that in Indonesian, bare nouns can occur freely as direct objects, object of prepositions and subjects, as shown in (1a-c), respectively. Notice that there are no elements in Indonesian that correspond to definite or indefinite articles, with the relevant distinction being only made by reference to contexts.

- (1) a. Dia membeli buku. (direct object)
 he buy book
 ‘He bought a book.’ (Wolff et al. 1992: 715; Chung 2000: 159)
- b. Tutup-lah pintu dengan kunci. (oblique object)
 lock-Emp door with key
 ‘Lock the door with a key.’ (McDonald 1976: 128; Chung 2000: 159)
- c. Rem depan, tanpa aku rem, mengerem sendiri. (subject)
 brake.front without I brake put.on.brake itself
 ‘The front brake, without being braked by me, braked itself.’
 (McDonald 1976: 102; Chung 2000: 160)

The fact that bare nominal arguments occur freely as shown in (1a-c), therefore, indicates that Indonesian is either a [+arg, -pred] such as Japanese or [+arg, +pred] language such as

English. Suppose now that Indonesian is a [+arg, -pred] language for the sake of argument. Chung argues that Indonesian is a classifier language, in which numerals precede bare nouns and are immediately followed by a classifier, consistent with the [+arg, -pred] language. Interestingly, though Dardjowidjojo (1978) notes that Indonesian has as many as sixty classifiers, only three of them are in frequent use in contemporary Indonesian, according to Chung (p. 162); *orang* ‘person’ for counting persons, *ekor* ‘tail’ for counting animals, birds and fish, and *buah* ‘fruit’ for counting other objects. Her evidence that Indonesian is a classifier language is two-fold. First, classifiers are obligatory with the numeral *se-* ‘one’ in Indonesian. She observes (p. 163) that *se-* ‘one’ must either be followed by a classifier or else occur in the fixed expression *s(u)atu*, in which it is combined with the obsolete classifier *watu* ‘stone’ (Hopper 1986: 311)”. In the absence of a classifier, *se-* cannot occur. This is illustrated in examples as in (2a, b).

- (2) a. Kemudian di-ambil-nya se-helai serbet kertas yang baru. (*se* + classifier)
 later Pv-take-by.her one-Class napkin paper which new
 ‘Then she got a new napkin.’ (Purwo 1989: 318; Chung 2000: 163)

b. Kemudian di-ambil-nya kertas baru.

later Pv-take-by.her paper new

‘Then she got a new napkin.’ (Purwo 1989: 312; Chung 2000: 163)

Second, Chung notes (p. 164) that, at an earlier stage of the development of Indonesian, overt classifiers were more frequent than they are today after *dua* ‘two’ and higher numerals. To support this, she points out that “statistics reported in Hopper’s 1986 careful study of classifier use in the 19th century Malay of the *Hikayat Abdullah*, an autobiography published in 1849, suggest that roughly 80% of the numerals that combine with NPs are accompanied by an overt classifier.” The following examples illustrate this point.

(3) a. Maka di-tembak-lah dua-bělas puchok mēriam di-bukit.

then Pv-fire-Emp twelve Class gun from-hill

‘[A salute of] twelve guns was fired from the hill.’

(Abdullah 1963 [1849]: 222; Chung 2000: 164)

b. Ada pun takala mēmbuat rumah itu tiga orang orang China kuli jatuh dari atas.

as for when make house the three Class person Chinese laborer fell from top

‘In the course of its construction three of the Chinese workmen fell from the top.’

(Abdullah 1963 [1849]: 222; Chung 2000: 164)

Granted that Indonesian is a classifier language with generalized bare arguments, the only setting that would account for these two properties under Chierchia’s Nominal Mapping Parameter is a [+arg, -pred] type language. This setting, thus, predicts that Indonesian should not have plural morphology. Chung (p. 164) shows that this prediction is false because Indonesian does have a way of expressing plurality via full reduplication of the root, as illustrated in (4) (see chapter 6 for further discussion of the semantics of reduplication).

(4) Buah-lah kalimat-kalimat berikut menjadi kalimat-kalimat negatif.

make-Emp sentence-Pl following become sentence-Pl negative

‘Please make the following sentences negative.’

(Dardjowidjojo 1978: 27; Chung 2000: 16)

It is widely known in the literature on Indonesian (Dardjowidjojo 1978: Dyen 1964) that, though an unreduplicated bare noun can be constructed either as singular or plural, its reduplicated counterpart is necessarily interpreted as plural. This traditional observation becomes important in section 4. The crucial point here is that, under Chierchia's system, a kind, by definition, cannot differentiate between singular and plural instance of that kind. The fact that reduplication has the function of denoting plurality, as in (4), thus, constitutes strong evidence that Indonesian cannot be a [+arg, -pred] language such as Japanese or Chinese.³

Chung further points out two problems that would arise with the alternative potential analysis of Indonesian as a [+arg, +pred] language such as English. The first problem concerns the scopelessness of bare nominal arguments in Indonesian. Following Carlson (1977), Chierchia (1998a: 368) observes that in English, bare plurals in object position behave as kinds in that they obligatorily take narrow scope with respect to negation and intensional operators

³ In Indonesian, reduplication does not co-occur with numerals more than 2, as shown by the ungrammaticality of (i):

- (i) * Esti mem-beli tiga buku-buku kemarin.
 Esti Av-buy three book-Red yesterday
 'Esti bought three books yesterday.'

One analysis is that the reduplicative process is blocked by numerals on the grounds of expressive economy. Since numerals more than 2 denotes a more specific function (e.g., 2, 3, 4, etc.) than nominal reduplication ($n, n \geq 2$, defined in contexts), the use of a numeral independently blocks reduplication. This particular analysis, however, does not affect the content of this chapter. Thanks to Heidi Harley (personal communication) for useful discussion.

whereas indefinite singulars can take wide scope over these scope-bearing elements. This contrast is shown in (5a, b) with respect to negation.

(5) a. I didn't see spots on the floor

⇒ Neg>Indefinite (narrow scope): I did not see any spot on the floor.

* Indefinite>Neg (wide scope): There is a spot/are certain spots that I failed to see on the floor.

b. I didn't see a spot on the floor.

⇒ * Neg>Indefinite (narrow scope): I did not see any spot on the floor.

Indefinite>Neg (wide scope): There is a spot that I failed to see on the floor.

(Chierchia 1998a: 368)

If Indonesian is a [+arg, +pred] language as in English, we predict that bare nominals should also be able to take scope over negation as do English indefinites, under their indefinite reading. This prediction is clearly false, as evidenced by (6a, b), where the bare nouns *buku* 'book' and *perempuan* 'woman' must take narrow scope with respect to the negative element *tidak* 'not'. To express the wide scope reading of the bare nominal, a relative clause existential sentence must be used as in (6c).

(6) a. Ali tidak jadi membeli buku.

Ali Neg finished buy book

‘Ali didn’t finish any book(s).’/*‘There was a book that Ali didn’t finish.’ (Chung 2000: 161)

b. Ia tidak melihat perempuan.

he Neg see woman

‘He saw no women.’ (Purwo 1989: 303; Chung 2000: 161)

c. Ada sebuah buku yang Ali tidak jadi beli.

exist one book that Ali Neg finish buy

‘There is a book that Ali didn’t finish.’

The second problem concerns the lack of reduplicated forms for generic statements.

Chierchia (1998a: 362-368) observes that bare nominals in English can be inflected for plural in generic statements as in (7).

(7) Dogs bark.

⇒ plural interpretation: There is more than one dog that barks/are barking.

generic interpretation: It is a general property of dogs that they bark.

(adopted from Chierchia 1998: 367)

We have seen above that Indonesian has a way of expressing plurality by full reduplication of the root, as shown in examples such as (4). Then, if Indonesian is a [+arg, +pred] language as in English, we predict that reduplicated nominals in this language should also be able to feed generic interpretation. Examples as in (8a), however, show that this prediction is false, because the reduplicated counterpart of the root *anjing* ‘dog’ can yield only a plural interpretation (Sneddon 1996: 17). To express the generic interpretation, the unreduplicated bare noun must be used instead, as shown in (8b), which gives rise to both plural and generic interpretations.

(8) a. *Anjing-anjing* *menggonggong*.

dog-Red bark-Red

‘Dogs bark/are barking.’

⇒ plural interpretation: There is more than one dog that barks/is barking.

* generic interpretation: It is a general property of dogs that they bark.

b. *Anjing* *menggonggong*.

dog bark

‘Dogs bark/are barking.’

⇒ plural interpretation: There is more than one dog that barks/is barking.

generic interpretation: It is a general property of dogs that they bark.

Chung (p. 168) briefly considers the final analytic possibility compatible with the Nominal Mapping Parameter whereby Indonesian is currently in the transition stage from a [+arg, -pred] language to a [+arg, +pred] language. Chung rejects this analysis on the ground that it would lead us to “expect the singular-plural contrast to be less in evidence at earlier stages of the language, when overt classifiers were more frequent.” Chung argues that prediction is not borne out by the 19th century Malay from *Hikayat Abdullah*, because examples as in (9) that contain nouns with overt plural inflection occur frequently.

- (9) Maka tukang-tukang kayu pun mēnarah-lah akan sēgala pērkakas rumah itu.
 then worker-Pl wood also smooth-Emp for all part house that
 ‘Carpenters started shaping planks of wood for various parts of the building.’

(Abdullah 1963 [1849]: 221; Chung 2000: 169)

To test Chung’s prediction, one would need a translation of the same autobiography in contemporary Indonesian so that we can compare the number of reduplicated nouns in the 17th century Malay with that of the corresponding nouns in contemporary Indonesian. This prediction, however, is impossible to test for the moment due to the absence of the

Indonesian translation of the text to date. Thus, I leave the fourth analysis of Indonesian aside in this chapter.

To sum up, Indonesian does not fit into any one of the three language types under Chierchia's (1998a, b) semantic typology. The free occurrence of determinerless, bare arguments shows that Indonesian is not a [-arg, +pred] language such as Italian or French. The presence of plural morphology marked by full reduplication shows that this language is also not a [+arg, -pred] language such as Japanese or Chinese. The obligatory narrow scope reading of bare nominals with respect to negation and the lack of the reduplicated form of a bare noun for generic statements means that Indonesian is also not a [+arg, +pred] language such as English or Russian.

Although Chung's argument against the Nominal Mapping Parameter from Indonesian is clear and I develop similar arguments against it based on parallel facts from Javanese in the next subsection, I point out here that it is problematic to analyze Indonesian as a (generalized) classifier language, as Chung does. In fact, Chung (p. 162-164) provides two arguments that suggest that Indonesian is not a classifier language. First, she observes (p. 162) that classifiers in Indonesian are more often than not omitted in colloquial Indonesian after *dua* 'two' and some number greater than two (Dardjowidjojo 1978: 64, 65; MacDonald 1976: 82, 83; Sneddon 1996: 134, 135). Second, she points out that "even in formal registers of the

contemporary language, an overt classifier need not occur after *dua* ‘two’ or higher numerals.”

Some examples to illustrate these points are given in (10), which is “excerpted from Indonesian translations of English-language articles on Indonesian syntax” (p. 163).

(10) Muda-mudahan makalah ini telah memenuhi dua tujuan pokok-nya.

hopefully then this already fulfill two goal principal-its

‘Hopefully, this paper has fulfilled its two major goals.’ (Purwo 1989: 333, Chung 2000: 163)

How about the two arguments made by Chung based on the obligatory presence of a classifier with the numeral *se-* ‘one’ and the classifier use in the 19th century Malay of the *Hikayat Abdullah*? The obligatory presence of a classifier with the numeral can be accounted for independently without necessarily assuming that Indonesian is a classifier language because the numeral *se-* ‘one’ is a clitic that needs a classifier as a host. Indeed, the non-clitic free morpheme meaning ‘one’, *satu*, can occur without any classifier, as the comparison between (2a) and (11) shows.

(11) Kemudian di-ambil-nya satu serbet kertas yang baru. (*satu* + no classifier)

later Pv-take-by.her one napkin paper which new

‘Then she got a new napkin.’

Chung’s second argument for her claim that Indonesian is a classifier language came from her observation that overt classifiers were more frequent than they are today after *dua* ‘two’ and higher numerals. We have seen above that this is supported by Hopper’s 1986 careful study of classifier use in the 19th century Malay of the *Hikayat Abdullah*, an autobiography published in 1849, according to which roughly 80% of the numerals that combine with NPs are accompanied by an overt classifier. This argument seems hard to evaluate at present for two reasons. First, as already noted above, there is no translation of the relevant autobiography in modern standard Indonesian to see whether this argument is still substantiated. Second, granted that we will have such a translation at hand, it is still not clear whether the comparison of 17th century Malay and the contemporary Indonesian versions of the same text can yield a meaningful result. Given a variety of differences between contemporary Malay and Indonesian that pose a formidable obstacle to field linguists working on dialects of Malay/Indonesian, it is possible that there would be even more dramatic differences between 17th century Malay and contemporary Indonesian.

At any rate, Chung (p. 164) concludes, based on the two arguments made above, that “when NP combines with a numeral, a classifier must be syntactically present even though it need not be phonetically overt.” This final conclusion, however, seems not to be supported by facts from other [+arg, -pred] languages such as Japanese or Chinese, which have a bona-fide classifier system. In Japanese, for example, when a noun combines with a numeral, a classifier must be overtly expressed. This is shown by the contrast between (12a) and (12b).

- (12) a. Taro-ga san-nin-no gakusei-o mita.
 Taro-Nom 3-Class-Link student-Acc saw
 ‘Taro saw three students.’
- b. * Taro-ga san-(no) gakusei-o mita.
 Taro-Nom 3-Class-Link student-Acc saw
 ‘Taro saw three students.’

As we will see in the next section, there is clear evidence that Javanese is not a classifier system. Given the large-scale linguistic and social interaction between Indonesian and Javanese in the Java island of Indonesia (Poedjosoedarmo 1982: 84; see also Chung 2000: 163), one likely scenario, compatible with all the facts noted above, is that contemporary

Indonesian has lost a classifier system due to the linguistic influence from Javanese. Based on the considerations above, I assume that Indonesian is a non-classifier language like Javanese, contra Chung (2000).

3.2. *Bare Nominals in Javanese*

In this subsection, I develop a similar argument as made by Chung from Indonesian against the Nominal Mapping Parameter from Javanese, a closely related Austronesian language spoken in Indonesia. I show that Javanese behaves like Indonesian in all relevant respects. Specifically, the free occurrence of bare nominal arguments shows that Javanese cannot be a [-arg, +pred] language such as Italian or French under Chierchia's system; it must be either a [+arg, -pred] language such as Japanese or Chinese or a [+arg, +pred] language such as English and Russian. I reject the first possibility on the ground that Javanese has plural morphology marked via reduplication and lacks a generalized classifier system. I reject the second possibility based on the obligatory narrow scope of bare nominals under their indefinite interpretation as well as the lack of reduplicated forms of bare nouns for generic statements. These results, therefore, show that Javanese presents itself as another counterexample to the predictions of Chierchia's Nominal Mapping Parameter.

Like Indonesian, Javanese permits bare, determinerless arguments to occur rather freely in any argument position, as illustrated in (13a-d).

(13) a. Buku larang.

book expensive

‘A book/the book/books {is/are} expensive.’

b. Esti tuku buku.

Esti buy book

‘Esti bought a book/the book/books.’

c. Esti nukokke uwong buku.

Esti buy man book

‘Esti bought a man/the man/men a book/the book/books.’

d. Esti entuk informasi seko buku.

Esti get information from book

‘Esti got information from a book/the book/books.’

In these examples, the bare noun *buku* ‘book’ occurs as the subject (13a), the direct object (13b), the indirect object (13c), and the object of preposition (13d). The free occurrence of bare

arguments in Javanese, thus, shows that this language cannot be a [-arg, +pred] language such as Italian or French. In other words, it should be either a [+arg, -pred] language or a [+arg, +pred] language. However, I demonstrate, following Chung's (2000) argument based on Indonesian, that Javanese does not fit into either one of the language types.

Let us now consider first the analysis whereby Javanese is a [+arg, -pred] language. Recall that, under Chierchia's semantic theory, [+arg, -pred] languages should have three morphosyntactic properties due to the specification of bare nouns as denoting a kind; a bare nominal argument, a generalized classifier system, and no plural morphology. Thus, if Javanese is a [+arg, -pred] language, the Nominal Mapping Parameter predicts that a) this language should have a generalized classifier system and that b) it should have no plural morphology, as in Japanese or Chinese.⁴ Both of these predictions are falsified by examples as in (14a, b) and (15a, b).

⁴ In fact, we will see in the next section that Japanese also does have plural morphology marked by *tachi*-suffixation to or reduplication of a nominal root, as shown in (ia, b).

- | | | | |
|--------|--|----|---|
| (i) a. | Shonen-tachi-ga kooen-de asonda.
boy-Pl-Nom park-Loc played
'Boys played in the park.' | b. | Hito-bito-ga kooen-ni kita.
people-Red-Nom park-Loc came
'People came to the park.' |
|--------|--|----|---|

This fact poses a problem for Chierchia's Nominal Mapping Parameter because Japanese, being a [+arg, —pred] language, should not be able to have plural morphology due to the kind-denoting requirement of bare nouns. See Nakanishi and Tomioka (2004) for a detailed analysis of the semantics of *-tachi*. See also note 30.

- (14) a. Esti tuku buku telu. b. Esti mangan pelem loro.
 Esti buy book three Esti eat mango two
 ‘Esti bought three books.’ ‘Esti ate two mangos.’

- (15) a. Esti tuku buku-buku. b. Esti nata meja-meja.
 Esti buy book-Red Esti arranged table-Red
 ‘Esti bought books.’ ‘Esti arranged tables.’

(14a, b) show that Javanese is not a classifier language (Poedjosoedarmo 1982; Robson 2002). This utter lack of a classifier system is in contrast with the optionality of classifiers in Indonesian as we have seen in the previous section. The examples in (15a, b) show that Javanese marks plurality via full reduplication of the root, as in Indonesian. These two facts clearly indicate that Javanese is not a [+arg, -pred] language such as Chinese and Japanese.

Let us next consider another analysis whereby Javanese is a [+arg, +pred] language such as English or Russian. Again applying the two arguments developed by Chung to bare nominals in Javanese, we can see that this analysis is incorrect. The first argument concerns the scopelessness of bare nouns in Javanese under their indefinite interpretation. Recall Carlson’s (1977)/Chierchia’s (1998a: 368) observation from the last subsection that bare plurals in English behave as kinds and

obligatorily take narrow scope with respect to negation. This is not the case with indefinite singular nouns which can take wide scope over negation. The relevant contrast was illustrated in (5a, b). Now, if Javanese is a [+arg, +pred] language, a bare nominal argument in this language should also allow a wide scope reading of the argument with respect to negation on its indefinite singular interpretation, as do English indefinites. This prediction is incorrect, as shown by (16a), in which the bare noun *kotoran* ‘spot’ must take narrow scope with respect to negation. Just as in Indonesian, a relative clause sentence is used as in (16b) to express the wide scope reading.

(16)a. Aku ora weruh kotoran ning jubin.

I Neg see spot on floor

‘I did not see spots on the floor.’

⇒ Neg>Indefinite (narrow scope): I did not see any spot on the floor.

*Indefinite>Neg (wide scope): There is a spot that I failed to see on the floor.

b. Ana kotoran sing aku ora weruh ning jubin.

exist spot that I Neg see in floor

‘There is a spot on the floor that I failed to see.’

⇒ * Neg>Indefinite (narrow scope): I did not see any spot on the floor.

Indefinite>Neg (wide scope): There is a spot that I failed to see on the floor.

The second argument against the classification of Javanese as a [+arg, +pred] language concerns the absence of reduplicated forms in Javanese for the generic use of bare nouns. We have seen above that bare nominals in English can be inflected for plural in generic statements as in (7a, b). If Javanese is a [+arg, +pred] language, then the prediction is that bare nominal arguments in Javanese should also be able to reduplicate when interpreted as generic. This prediction is incorrect, as shown in (17a).

(17) a. Asu-asu njegug. ‘Dogs are barking.’

dog-Red bark

⇒ plural interpretation: There is more than one dog that barks/are barking.

* generic interpretation: It is a general property of dogs that they bark.

b. Asu njegug. ‘A dog/the dog/dogs bark.’

dog bark

⇒ plural interpretation: There is more than one dog that barks/are barking.

generic interpretation: It is a general property of dogs that they bark.

In the example in (17a), the reduplicated noun *asu-asu* ‘dogs’ only allows plural interpretation. Instead, the non-reduplicated bare nominal is used for generic statements

in Javanese, as in (17b). True, this fact is naturally accounted for if Javanese is a [+arg, -pred] language as in Japanese because kinds are known to yield a universal reading (Chierchia 1998: 363a) but we have seen already above that this analytic possibility is incorrect in light of the lack of a generalized classifier system and the presence of plural morphology marked by reduplication. Thus, the two arguments concerning the mandatory narrow scope reading of bare nominals under their indefinite readings and the lack of reduplicated bare nouns for generic statements in Javanese provide evidence that this language cannot be a [+arg, +pred] language such as English. The last analytic possibility that is compatible with the Nominal Mapping Parameter is to resort to a finer-grained distinction within Chierchia's semantic typology. Under this possibility, Javanese may well be similar to Russian.⁵ In the next section, I show that this comparison is on the right track by demonstrating that Javanese is minimally different from Russian in the possible values for the Number head that each language is allowed to choose.

To summarize this section, I have shown, using the arguments developed by Chung (2000), that neither Indonesian nor Javanese can be classified as any one of the three language types that should serve to categorize all human languages under Chierchia's Nominal Mapping Parameter. This result is important because it raises important questions about the

⁵ Thanks to Heidi Harley (personal communication) and Simin Karimi (personal communication) for this observation.

extent to which the morphosyntactic properties of a bare noun in a language are predictable from its denotation and vice versa. In particular, the descriptive results achieved here clearly suggest that the Nominal Mapping Parameter imposes too tight a mapping between the denotation and morphosyntax of NPs.

4. A Relativized Parametric Theory of Nominal Denotation: From Indonesia to the World

In this section, I develop a purely syntactic account for the effects of the Nominal Mapping Parameter of Chierchia (1998a, b) within the framework of the Minimalist Program (Chomsky 1995). I start by pointing out a conceptual problem with Chierchia's notion of semantic parameter. I show that this view of parameters is hardly groundable within the standard conception of the locus of parameters within the Principles & Parameters approach to language variation. Resolution of this problem leads us to seek an alternative explanation in the realm of the morphosyntactic variation of the lexicon in each language, which is an unreducible source of linguistic variation (Borer 1984; Fukui 1986, 1995; Chomsky 1995). I propose a relativized parametric theory of nominal denotation, whereby languages differ in terms of a) the height/complexity of the functional super-structure above bare nominals and b) the possible set of Number values. Point (a) has been independently argued for by recent work such as Grimshaw (1991, 2005),

Massam (2001), Guilfoyle and Noonan (1992), and Vainikka (1993/1994). Point (b) is independently supported by extensive study of the Number system in languages such as Malay and Bahasa Indonesia as in Carson (2000) and Chung (2000). The proposed analysis, thus, provides further evidence for these claims.⁶

4.1. *A Relativized Parametric Theory of Nominal Denotation and Morphosyntax*

Let us start by pointing out a conceptual problem with Chierchia's notion of "Semantic Parameter". The Nominal Mapping Parameter essentially proposes that parameters can be postulated within the mapping between syntactic and semantic/LF representations. This proposal, however, is clearly at odds with the standard conception of the locus of the parameter as in the Principles & Parameters approach to language variation as outlined in Chomsky (1995). Instead, this approach takes it as a fundamental heuristic that parametric variation be restricted to the properties of the lexicon (Chomsky 1995, 2000) or, more specifically, functional categories alone, a hypothesis known as the *Functional Parameterization Hypothesis* (Borer 1984; Fukui 1986, 1995). A natural approach then should be one that

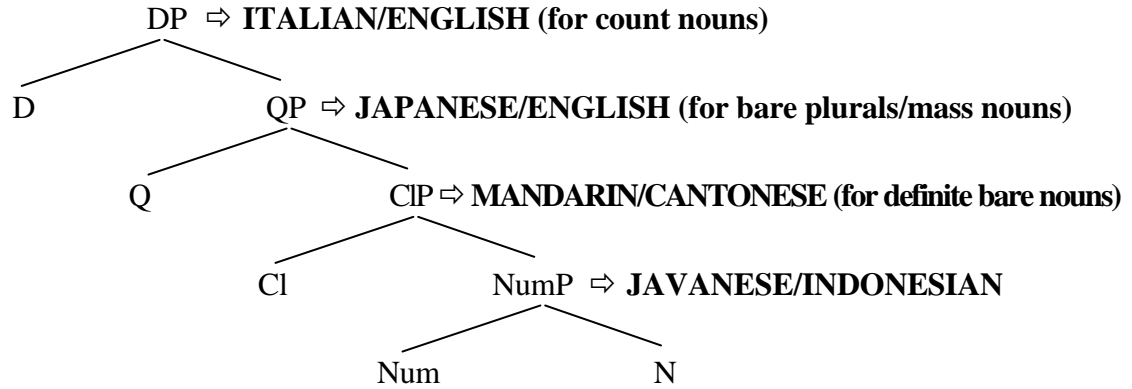
⁶ Thanks to Heidi Harley (personal communication) for suggesting the idea of "growth of bare nominals" and to Andrew Carnie (personal communication) for bringing my attention to several important works such as those cited here that independently proposed such an idea.

derives crosslinguistic variation in the denotation and morphosyntax of bare nominals from independently motivated variation in the makeup of the lexicon in each language.

With this agenda in mind, I propose that languages differ in two dimensions: a) how high a bare nominal can “grow” across languages and b) what set of the binary values the Num head in each language selects. For the first part of this claim, I propose the universal nominal morphosyntactic hierarchy $DP > QP > CIP > NumP > NP$, from which languages set the appropriate height of nominal projections for their bare nouns. Specifically, I argue that languages such as Javanese and Indonesian project up to NumP; languages such as Japanese project up to QP; languages such as Italian project up to DP. Finally, languages such as English project either up to QP like Japanese or DP like Italian, depending on the nature of a noun inserted into the N head in a manner to be explained below. A question arises as to whether there are languages which instantiate the CIP option.⁷ I address this question in section 5, where I claim, drawing on evidence from Cheng and Sybesma (1999), that definite bare nouns in Cantonese and Mandarin instantiate this option. This proposal is summarized in (18).

⁷ Thanks to Heidi Harley (personal communication) for this question.

(18) The Universal Nominal Morphosyntactic Hierarchy



As stated at the outset of this section, this idea of “growth of NPs” is not a new idea but rather has been argued for by recent work as in Grimshaw (1991, 2005), Massam (2001), Guilfoyle and Noonan (1992), and Vainikka (1993/1994) on various grounds. Grimshaw (1991, 2005) proposes an influential theory of extended projection whereby lexical heads such as V and N form a single projection with functional heads such as T/C and D/P on top of them conditioned by two constraints on projection. First, two categories in a projection chain must be categorically identical in terms of verbal and nominal features. Second, the F-value of X must not be higher than the F-value of YP” when X is the head of YP and YP is a projection of X (Grimshaw 2001: 4, 5).⁸ “F-value” is a functional status assigned as

⁸ Grimshaw (p. 4) defines the notions of a *head* and a *projection* as in (i):

follows: F_0 is assigned to the lexical categories, F_1 is assigned to the lowest level functional category; and F_n to the next successively higher functional categories.⁹ Thus, the lexical category V forms an extended projection with T and C because the F -value of V , being associated with the verbal feature, is not higher than that of a T or C . For the same reason, N (F_1) forms an extended projection with D (F_2) and P (F_3). One important respect in which Grimshaw's approach is similar to the proposed analysis based on the universal nominal functional structure in (18) is her claim that, under a universal tree structure/inventory (as in Cinque 1999), languages can choose whether or not to include each functional specification as a head in their possible inventory, as long as possible combinations of the functional head and their complement do not violate the two constraints noted above. Thus, if we have 20 functional categories in the universal inventory, a language might choose $\{F_{20}, F_9, F_5, F_0\}$ while another language might

-
- (i) X is a head of YP , and YP is a projection of X iff:
- a. YP dominates X
 - b. The categorial features of YP and X are consistent.
 - c. There is no inconsistency in the categorial features of all nodes intervening between X and YP (where a node N intervenes between X and YP if YP dominates X and N , N dominates X .)

For example, TP forms an extended projection with V because the F -value of V (F_0) is not higher than the F -value of TP (F_1) and no projection intervenes between TP and V . The same holds for the CP . See Grimshaw (2005) for details.

⁹ Grimshaw (2005: 64-71) also provides a useful summary of a variety of papers that appeared after the publication of Grimshaw (1991) that address important issues that arise in the theory of extended projections.

choose {F8, F4, F2, F0}. However, the two principles of extended projections above impose strict structure constraints on the relative hierarchy of these categories. Grimshaw's particular application of the theory of extended projection to nominal domains is different from the proposed analysis in that the highest (extended) projection of nominals is PP under her analysis whereas it is DP under the current analysis. Nonetheless, in a fundamental respect, her proposal is quite similar to the analysis I develop here, according to which languages choose a subset of the universal nominal structure given in (18) (or the universal morphosyntactic pool), and this choice determines the fate of the morphosyntactic behavior of elements embedded within the structural configuration parametrically chosen. Therefore, the proposed analysis can be understood as one particular way of implementing Grimshaw's theory of extended projection to nominal domains within a more elaborated array of functional projections such as Num, Cl, and Q.

Massam (2001) reaches similar conclusions based on her careful study of what she calls Pseudo Noun Incorporation/PNI in Niuean, an ergative-absolutive language of the Tongic subgroup from the Oceanic family. This language is a VSOX language, as illustrated in (19a). As Massam observes, this language also allows VOSX order as shown in (19b), which she analyzes as a case of PNI.

(19) a. Ne kai e Sione e tau talo aki e huki.

Past eat Erg Sione Abs Pl taro with Abs fork

‘Sione ate the taros with a fork.’ (Massam 2001: 155)

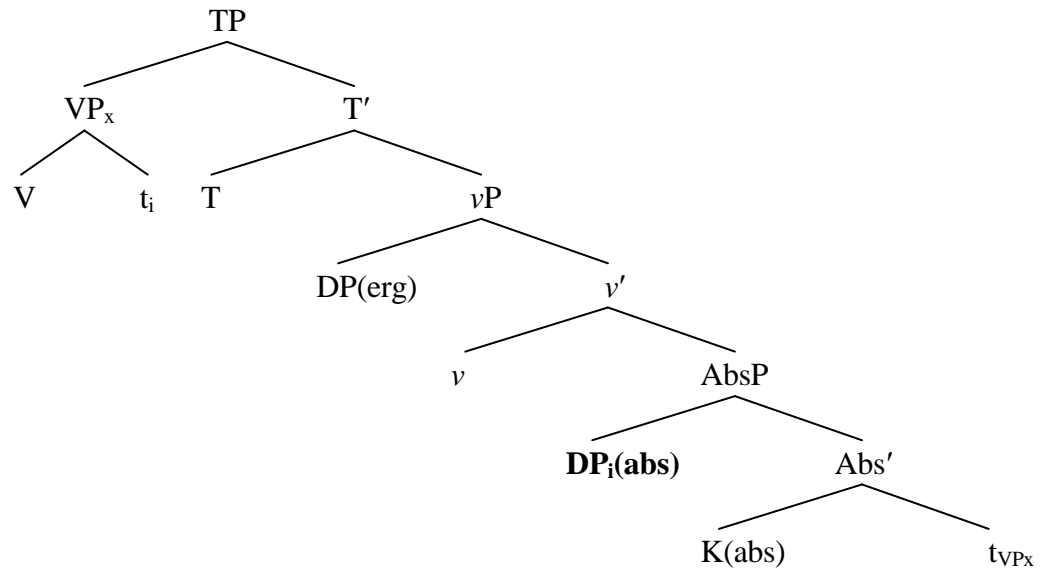
b. Neinu kofee kono a Mele.

Past drink coffee bitter Abs

‘Mary drank bitter coffee.’ (Massam 2001: 158)

Massam assumes that both examples are derived by the VP fronting into [Spec, TP]. For the VSOX order as in (19a), the DP argument undergoes movement into [Spec, AbsP] to check Absolutive Case. This movement is followed by the fronting of the VP that contains the verb and the trace of the direct object into [Spec, TP]. This derivation is illustrated in (20).

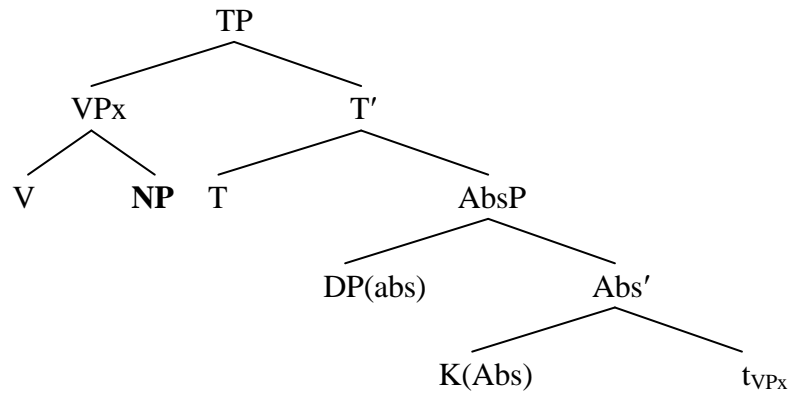
(20) The Derivation of the VSOX Order



(slightly modified from Massam 2001: 163)

For the VOSX/PNI order as in (19b), Massam crucially assumes that the direct object is an NP, not a DP. Under the assumption that Case features appear on an extended functional head such as KP, not on an NP, the direct object in (19b) cannot (and need not) check Absolutive Case, which is checked by the external argument of the sentence. The VOSX order is derived by fronting of the VP that contains both the verb and the NP object. This derivation is illustrated in (21).

(21) The Derivation of the VOSX Order



(slightly modified from Massam 2001: 165)

Massam shows that the analysis above that draws on the NP vs. DP distinction provides a natural account of several differences between the two types of Pseudo Noun Incorporation known as Generic PNI and Existential PNI first noted by Seiter (1980). The properties of the two types of PNI that are relevant for present purposes are summarized in (22) and (23), together with illustrative examples in each type of the PNI in (24a) and (24b).

(22) The characteristics of the Generic PNI Construction (Massam 2001: 172)

- a. NP is non-specific and non-referential
- b. no extended nominal categories or dependants of extended nominal categories
(i.e., tensed relative clauses) appear over N.
- c. durative/frequentative meaning

(23) The characteristics of the Existential PNI Construction (Massam 2001: 177)

- a. occurs with closed class of verbs *fai* 'have/be', *muhu* 'have plenty/be plentiful'
- b. relative clause can appear at the right edge of the sentence
- c. NP is referential, non-specific and indefinite

(24) a. Example of the Generic PNI (= (19b)) b. Example of the Existential PNI

Ne inu kofee kono a Mele. kua fai nakai e umu haau?

Past drink coffee bitter Abs Perf make Q Abs oven your

'Mary drank bitter coffee.'

'Have you made your oven yet?'

(Massam 2001: 173)

The three characteristics of the Generic PNI in (22a-c) directly follow from the NP status of the pseudo-incorporated nominal. First, property (22b) falls out from the fact that it is only a bare NP that undergoes movement into [Spec, TP] with the verb; if it were a DP, it would need to check Absolutive Case, contrary to facts. This property, Massam argues, also accounts for the fact that relative clauses cannot attach to pseudo-incorporated nominals under the assumption (Finer 1998; Ghomeshi 1996; Larson 1994; Kayne 1994) that relative clauses appear high within DPs. The property in (22a) is also naturally predicted under the standard assumption that a bare nominal has a non-referential, non-specific denotation. This property, in turn, yields the durative/frequentative meaning for the Generic PNI because the lack of referentiality ensures unbounded/non-delimited interpretation of the event denoted by the sentence, as argued for in Jackendoff (1990) and Tenny (1994), as cited in Massam (2001:171). Massam further shows that the three properties of Existential PNI in (23a-c) can also be derived if the nominal is topped with a quasi-determiner projection contributed by verbs of existence such as *fai* 'have/be' and *humu* 'have plenty/be plentiful' (23a). She proposes (p. 186) that "*fai* is essentially at one and the same time a verb and a determiner, in that it serves simultaneously as the main lexical predicate head for the sentence, and as a functional category which confers referentiality on its complement NP" in the sense of Higginbotham (1985) (see also Nichols 1997 and Johns 1999 for similar ideas as applied to

Noun Incorporation in Zuni and Inuktitut, respectively). Specifically, verbs like *fai* bind the referentiality index of the NP as a typical determiner does to its complement. This proposal thus not only derives the fact (23b) that relative clauses can appear in the Existential NPI (because verbs of existence serve the same function as D heads) but also yields the referential, non-specific, indefinite reading to the NP (23c). It is clear by now that Massam's analysis of the differences between the Generic PNI and the Existential PNI crucially depends on the fine-grained distinction between the NP and DP. Therefore, her analysis can be considered as another variation of the analysis proposed here whereby a parametrically different height of nominal functional projections across languages accounts for different morphosyntactic properties of their bare nominals.

Guilfoyle and Noonan (1992) argue for what they call the *Structure Building Hypothesis*, namely, that, in the early stages of language acquisition, children have a UG-constrained grammar solely consisting of lexical categories (N/V/A/P), *Lexical Grammar*, with functional categories such as I, C, D, and KASE gradually emerging after lexical categories based on the positive evidence from input data and the emergence of what they call *Functional Grammar*.¹⁰ They show that this hypothesis receives strong support from

¹⁰ Guilfoyle and Noonan (p. 241) acknowledge that Lebeaux (1988) and Radford (1990) developed similar ideas. For example, they state, citing Lebeaux, that "the early stages of acquisition are a pure representation of theta-theory." See also Clahsen (1991), Clahsen and Penke (1992), and Platzack (1991) for related discussion on the Structure Building Hypothesis.

the optionality of subjects in child language as well as the absence of NP-movement and subject-auxiliary inversion and several stages concerning the position of verbs in German because these properties are exactly what we expect if child English syntax lacks functional categories such as T and C. The proposed idea of “growth of bare nominals”, thus, can be naturally regarded as applying Guilfoyle and Noonan’s Structure Building Hypothesis to the domain of nominal syntax.

Vainikka (1993/1994) provides further evidence for the non-maturational gradual development of phrase structure as in Guilfoyle and Noonan’s Structure Building Hypothesis, based on the parallel acquisition of nominative case and tense-related materials. Her study carefully observes four of the developmental stages the child “Nina” went through. During the first stage of her acquisition (age 1; 11-2; 1), Nina used predominantly genitive subjects. During the second stage (age 2; 1-2; 2), she instead began to use nominative subjects predominantly. This shift in case marking for subjects makes sense if Nina’s grammar developed a phrase structure that includes functional material in the V projection up to TP; the subject is assigned genitive case in the specifier of VP but when the TP structure is learned, the subject is assigned nominative case in the specifier of TP as in Adult English. Vainikka shows that this analysis is supported by the fact that the TP-related material such as modals, auxiliaries, third person singular *-s*, and past tense *-ed*,

also gradually emerge in the second stage. This view of phrase structure thus can be considered another variation of the proposed analysis of nominal syntax.

Getting back now to the second part of the claim concerning the Num head, I claim that there are two possible sets of values for the Num P; {singular, plural} or {neutral, plural}. Languages such as English and Italian select {singular, plural} or {neutral, plural} values whereas languages like Japanese, Javanese, and Indonesian select {neutral, plural} values. There is independent evidence that the possible values for the number slot in Italian and English are significantly different from those for the same slot in Indonesian, Javanese, and Japanese. Thus, in her extensive study of the number system in Malay, Carson (2000) shows that bare nouns in Malay are *neutral* with respect to number unless reduplication tells us otherwise, and concludes that Malay chooses {neutral, plural} values for the Number head. For example, a (unreduplicated) bare nominal can denote either a singular or plural instance of the entity denoted by that nominal whereas its reduplicated form specifically denotes more than one instance of the same entity. As noted in section 3.1, the same argument was independently made by Chung (2000: 165, 167) for Indonesian. Consider examples as in (25a) (= (4)) and (25b).¹¹

¹¹ The translation for (25b) is slightly modified to reflect the fact that *kalimat* can be interpreted either as singular or plural, a judgment confirmed by my native language consultant.

(25)a. Buat-lah kalimat-kalimat berikut menjadi kalimat-kalimat negatif.

make-Emp sentence-Pl following become sentence-Pl negative

‘Please make the following sentences negative.’ (Dardjowidjojo 1978: 27; Chung 2000: 165)

b. Kalimat Dasar.

sentence basic

‘Basic sentence(s)’ (Wolff et al, 1992; Chung 2000: 165)

In the example in (25a), the reduplicated noun *kalimat-kalimat* ‘sentences’ must be construed as ‘more than one sentences’ whereas, in the example in (25b), the corresponding bare noun *kalimat* ‘sentence(s)’ can be construed either as a singular or plural. These examples, thus, illustrate that a bare nominal in Indonesian takes {neutral, plural} values for the Num head as in Malay. See also Dyen (1964: 7a.-10), cited by Chung (2000: 166-167), who makes a similar observation. .

There is evidence that Japanese and Javanese also select the {neutral, plural} values for the Num head as in Indonesian though dominant morphological processes to denote plurality seem to be different between Japanese (*tachi*-suffixation) and Javanese/Indonesian (reduplication) (see note 3 for relevant discussion). Consider examples in (26a, b) from Javanese and (27a, b) from Japanese.

- (28) a. Callie lan Tisa kuwi kucing. b. Callie kucing
 Callie and Tisa Cop cat Callie cat
 ‘Callie and Tisa are cats.’ ‘Callie is a cat.’
- (29) a. Callie to Tisa-wa neko-da. b. Callie-wa neko-da.
 Callieand Tisa-Top cat-Cop Callie-Top cat-Cop
 ‘Callie and Tisa are cats.’ ‘Callie is a cat.’

Notice that there is no morpheme in Indonesian/Javanese/Japanese that specifically denotes singularity. As stated above, these languages make the distinction between singularity and plurality of a noun only based on contexts in which it is found. I take this to suggest that there is no grammaticalized notion of singularity in these languages.

The number system in languages such as English and Italian exhibits a different picture from that in Japanese/Indonesian/Javanese. In English, there is a purely grammatical distinction between mass and count nouns that seems not predictable from their conceptual structures, though I could not make a similar argument for the number system in Italian due to the unavailability of the data. Consider the pair of words, *wheat* and *oat*, in English. (30a,

b) and (31a, b) show that *wheat* is a mass noun whereas *oat* is a count noun, even though they both denote “a grain of cereal”.¹²

(30)a. David did not eat many oats. (31)a. *David did not eat many wheat(s).

b. *David did not eat much oat. b. David did not eat much wheat.

(30a, b) show that the word *oat* can go only with the determiner (many) that requires a count noun. (31a, b) show that the word *wheat* can go only with the determiner (much) that requires a mass noun. These examples, thus, suggest that the count-mass distinction in English is grammaticalized in certain pairs of words such as *oat* and *wheat*. Based on this consideration, I assume that English and Italian can take one of the following three values for the Num head: singular, plural, and neuter.¹³

¹² Many thanks to Heidi Harley for (30a, b) and (31a, b) and very useful discussion. (31b) is acceptable if *many wheats* is interpreted as “many types of wheat”.

¹³ As Heidi Harley (personal communication) notes, an intriguing issue remains with the English number system. Borer (2005) claims that the alleged English plural morpheme *-s* is a kind of classifier. One argument for this claim comes from the fact that the proposition in (i) is true even if only one dog is removed; if *-s* denoted a plurality of the noun it attaches to, (i) should come out as a false statement.

(i) Any dogs will be removed.

I would like to address this issue in my future work.

With all the results discussed thus far in place, the proposed relativized parametric theory of nominal denotation can be summarized as in Table 3.

Table 3: A Relativized Parametric Theory of Nominal Denotation (preliminary version)

Languages	Height of Nominal Projections	Possible Sets of Num Values
Indonesian/Javanese	NumP	{neutral, plural}
Japanese	QP	{neutral, plural}
Italian	DP	{singular, plural, neuter}
English	DP or QP	{singular, plural, neuter}

Three assumptions are in order here before we move onto the actual structural analysis of bare nominals across languages. First, I assume that the N in the universal nominal morphosyntactic hierarchy in (18) is underspecified with respect to its denotation (cf. Distributed Morphology; see, for example, Marantz 1997 and Harley and Noyer 1999). Second, I assume the economy of projection as proposed on various conceptual and empirical grounds in recent work (Chomsky 1995, Fukui 1986, 1995, Fukui and Speas 1986, Law 1991; Speas 1994; Bošković 1997; Grimshaw 1993; Radford 1990; Safir 1993). Specifically, if a language like Italian and French selects the {singular, plural} set for the Num head, the

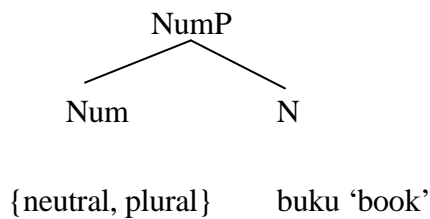
Cl(assifier) P does not project on the ground of expressive economy. The individuation function encoded by the singular value of the {singular, plural} set has the same function as that encoded by the classifier. Under the theory that semantic composition is computed in the bottom up fashion in a strictly local manner, the projection of the NumP with the relevant value makes the projection of the dominating CIP redundant. Finally, I assume that there is a feature checking/valuation relation of some sort between the Num head and its complement N. To be precise, the Num head with the {singular, plural} set values its complement as a count noun while the Num head with the {neutral, plural} set values it as a mass noun. I will return to this last assumption in the end of this section.

4.2. *Deriving the Denotation and Morphosyntax of Bare Nominals across Languages*

Consider first the nominal syntax of bare nouns in Javanese and Indonesian. We have seen in sections 2 and 3 that these two languages exhibit the following four morphosyntactic characteristics: i) they allow bare nominal arguments, ii) they lack a generalized classifier system, iii) all nouns are a type of mass nouns, and iv) they allow only narrow scope of bare nominals. All these properties straightforwardly follow if bare nominals in Javanese and Indonesian project up to NumP with the set {neutral, plural}, as shown in (32).¹⁴

¹⁴ See chapter 6 for a detailed explication of how reduplication is realized in Indonesian.

(32) The Nominal Structure of Bare Nouns in Indonesian and Javanese



First, the two languages in question allow bare nominal arguments because there is no DP, as shown in (32). Second, they don't have a generalized classifier system because they project only up to NumP for bare nominals. Third, the extension of all nouns is mass for the following reason. If the Num value is specified as neutral, the denotation of the NumP is a kind because it does not differentiate between singular and plural instances of the NumP. If the Num value is plural, the denotation of the Num P still comes out as a bare plural, which is also a kind under Carlson's (1977)/Chierchia's (1998a, b) theory. Thus, whichever value the Num head selects yields a kind, hence mass interpretation to the NumP. Finally, the obligatory narrow scope reading of bare nouns with respect to negation follows from the kind-denoting requirement (or whatever principle blocks the wide scope reading of such nouns). In this way, the clustering of the morphosyntactic properties observed in Javanese and Indonesian, which was shown to be unpredictable under

Chierchia's (1998a, b) Nominal Mapping Parameter, naturally follows from the interaction of the Num P structure and the {neutral, plural} set for the Num head.

One potential problem that remains with the proposed analysis is how to accommodate simple Javanese and Indonesian expressions as in (33a, b) and (34a, b), respectively.¹⁵

(33) Indonesian

- | | | | |
|---------------|-------|-------------|------|
| a. tiga | buku. | b. buku | ini |
| three | book | book | this |
| 'three books' | | 'this book' | |

(34) Javanese

- | | | | |
|---------------|-------|-------------|------|
| a. telung | buku. | b. buku | iki. |
| three | book | book | this |
| 'three books' | | 'this book' | |

The issue here is that expressions as in (33a, b) and (34a, b), which include numerals and demonstratives, might be incorrectly ruled out under present assumptions because of the lack of

¹⁵ This question was raised by Andrew Carnie (personal communication) and Heidi Harley (personal communication).

the QP or DP that would host the numeral and demonstrative words as their head in languages such as English and Italian. The claim that demonstratives are semantically functioning as determiners at least in the contemporary Indonesian has been made by MacDonald (1976: 85), who observes that the demonstrative *itu* ‘that’ is “coming to fulfill a function very much like that of the definite article.” A related question, of course, comes from our earlier observation that a bare noun in Indonesian/Javanese can have a singular indefinite reading, as we have seen in the examples in (6a, b) and (16a). Notice that these two questions, in fact, are two special cases of the more general problem of how it is possible that a bare noun, which otherwise would denote a mass extension, is also able to denote a singularity in Indonesian or Javanese.

I maintain, however, that this question receives a principled answer once we recall our earlier assumption that the singularity of a noun is not encoded in the grammar; rather, it is determined by contexts in which it appears. Consider first the issue raised by demonstratives. Fukui (1986, 1995) and Fukui and Speas (1986) argue that Japanese entirely lacks functional categories such as C, T, and D of the English-kind, or at best has a quite impoverished system of these function items, with elements such as subjects and modifiers all being attached to lexical projections as an X' adjunct (see also Guilfoyle and Noonan 1992 and Vainikka 1993/1994; see discussion in section 5). If this analysis can be extended to Indonesian and Javanese, then one can maintain that numerals and demonstratives as in (33a,

b) and (34a, b) are also modifiers of lexical projections such as NPs in the structure in (32) in these languages. There is independent evidence that demonstratives in Indonesian and Javanese are not D heads but instead modifiers of NPs. The evidence concerns the iterativity of such expressions. It is well known that lexical heads such as nouns do not impose any structural limit on iterating modifiers as long as they can be semantically interpreted and licensed. This is illustrated in examples in (35a-c) in English.

- (35) a. a big balloon.
 b. a red big balloon
 c. a red big expensive balloon

Demonstrative words in English such as *this* and *that*, by contrast, are instances of the D head, not modifiers of an NP, by the same criteria. This is shown by the contrast between (36a, b) and (36c), where *this/that* can never co-occur with other functional D elements such as *John's*. In other words, D elements are uniterable.

- (36) a. this/that book
- b. John's book
- c. * this/that John's book

Importantly, demonstratives in Indonesian and Javanese pattern with modifiers of the NP rather than D heads since they can co-occur with other possessor elements, as in (37c) and (38c).¹⁶

(37) Indonesian

- | | | | | | | |
|-------------|------|---------------|------|--------------------|------|------|
| a. buku | ini | b. buku | John | c. buku | John | ini |
| book | this | book | John | book | John | this |
| 'this book' | | 'John's book' | | 'this John's book' | | |

(38) Javanese

- | | | | | | | |
|-------------|------|---------------|------|--------------------|------|------|
| a. buku | iki | b. buku-ne | John | c. buku | John | iki |
| book | this | buku-Poss | John | book | John | this |
| 'this book' | | 'John's book' | | 'this John's book' | | |

¹⁶ It is important that the example in (37c) is not acceptable under the following reading: There are several people named John and the speaker talks about the book owned by one of them in contrast to the other. The only reading available to this sentence is roughly "this (deictic) book which belongs to John."

This contrast between (37c)/(38c) and (36c), therefore, provides independent support that there is no D in Javanese and Indonesian and that demonstratives in these languages are modifiers of the lexical NP projection. Further cross-linguistic evidence for this conclusion is provided by Bernstein (1997). Bernstein points out that, in languages such as Arabic and Greek, demonstratives and determiners can co-occur, suggesting that the former is not a D head. Bernstein further observes that, in other languages such as Spanish, Swedish, Norwegian and Scottish Gaelic, the demonstratives are the same words as ‘here’ and ‘there’.¹⁷ Crucially, this observation holds at least for Indonesian as well, as shown in examples in (39a, b), assuming that *sini* is an allomorph of *ini*.

¹⁷ Interestingly, Bernstein (p. 93) points out Javanese examples as in (i) as a case where a demonstrative and a determiner may co-occur.

- (i) ika n anak
 this the child
 ‘this child’ (Bernstein 1997: 93)

This example, however, is unacceptable according to my consultants, who speak modern Javanese. First, modern Javanese does not have any determiners like *n* ‘the’ in (i); recall our observation on Javanese made in section 3.2. Second, the demonstrative *iki* ‘this’ is used instead of *ika* in (i). Thus, Bernstein’s second point does not apply to Javanese.

(39) Indonesian

a. buku di *sini*.

book in here

‘A book/the book/books/the books are here.’

b. buku *ini*

book this

‘This book/these books’

The expression in (39a) is a TP due to the lack of copula in Indonesian where the word *sini* is used as a preposition. The expression in (39b) is a nominal phrase modified by the related word *ini*. This homophony between words meaning ‘here’ in Indonesian, thus, provide further support for the view that demonstratives are not necessarily D-elements but locative modifiers arguably attached to NPs.

More generally, it may well be that natural language uses whatever syntactic resources that are independently available to them to express the same meanings that other languages would express with a different (more articulated) syntactic structure.¹⁸ The fact that a bare nominal can, in principle, occur with numerals despite the fact that it denotes a mass extension also follows from the lack of the grammatical encoding of the singularity in Indonesian/Javanese. Syntax simply makes the denotation of a bare nominal underdetermined. It is the semantic interface that actually determines whether that particular

¹⁸ The idea in this paragraph owes a great deal to my discussion with Heidi Harley (personal communication).

instance of bare nominal denotes a count or mass extension. If a bare nominal occurs with numerals such as 2, 3, 4, etc, then the interface coerces its otherwise mass denotation into count denotation so it may be compatible with a specific numeral; if contexts in which a bare nominal X is found make it clear that it is intended to denote a singularity, then it will come out as denoting one instance of X at the semantic interface. This observation is supported from the fact that once coupled with a numeral, a bare noun triggers scope interaction with negation, as in English. This is illustrated by examples in (40) and (41), which minimally contrast with (6a) and (16a) in that the bare nominals are accompanied with a numeral.

(40) Ali tidak jadi membeli tiga buku.

Ali Neg finished buy three book

⇒ Neg>Indefinite (narrow scope): Ali did not buy three books.

Indefinite>Neg (wide scope): There are three books that Ali failed to buy.

(41) Aku ora weruh telung kotoran ning jubin.

I Neg see three spot on floor

⇒ Neg>Indefinite (narrow scope): I did not find three spots on the floor.

Indefinite>Neg (wide scope): There are three spots that I failed to find.

This observation, thus, provides important support for the idea behind the thesis of minimalist interface that the semantic component is endowed with domain-specific operations to yield an appropriate semantic interpretation to an otherwise underdetermined denotation of a particular noun. Notice that this underdetermined interface theory of nominal denotation predicts that there should not be any noun in Indonesian or Javanese whose count/mass denotation cannot be computed solely from its (prototypical) conceptual manifestation in the real world. This prediction seems to be borne out from the fact that Indonesian does not have pairs of words such as *oat* and *wheat* in English, though I must leave comprehensive examination of this prediction for my future research.¹⁹

¹⁹ Another important issue that the proposed analysis does not fully resolve is what restrictions are imposed on this recursion. For example, it is impossible to freely concatenate more than one numeral in Indonesian, as in (i), the intended meaning being that *dua tiga buku* means $2 \times 3 = 6$ books.

- (i) * Esti mem-beli dua tiga buku.
 Esti Av-buy two three book
 ‘Esti bought

I believe that the impossibility of this type of recursion violates certain general semantic or pragmatic constraints such as expressive economy and Gricean Maxims. In terms of expressive economy, *enam* ‘six’ is a more economical way of denoting 6 than multiplication of 2×3 . The same word also satisfies the Maxim of Quantity better than 2×3 because it is more concise than 2×3 . This question might lead us to a new comparative analysis of the syntax-semantics interface with respect to complex cardinals. See Ionin and Matushansky (2006) for a promising compositional approach to the semantics of numerals which do behave in this way, for example, *three hundred*. The comprehensive discussion of the cross-linguistic typology of number systems goes beyond the scope of this chapter. I thank Heidi Harley (personal communication) for useful discussion on the question raised here.

Let us now turn to the nominal syntax of bare nouns in Japanese. Japanese has the following morphosyntactic characteristics; i) bare arguments, ii) the generalized classifier system, iii) the extension of all nouns is mass, and iv) the obligatory narrow scope of bare nouns with respect to negation. The first three properties were noted by Chierchia (1998a, b); the last property is illustrated in the contrast between examples in (42a, b).

(42)a. John-ga yuka-de yogore-o mituke-naka-tta (koto)

John-Nom floor-Loc dirt-Acc find-Neg-Past (fact)

‘(The fact that) John did not find a spot/any spot/spots on the floor.’

⇒ Neg>indefinite (narrow scope): I did not see any spot on the floor.’

* indefinite>Neg (wide scope): There is a spot that I failed to see on the floor.’

b. John-ga yuka-de mituke-naka-tta yogore-ga aru (koto).

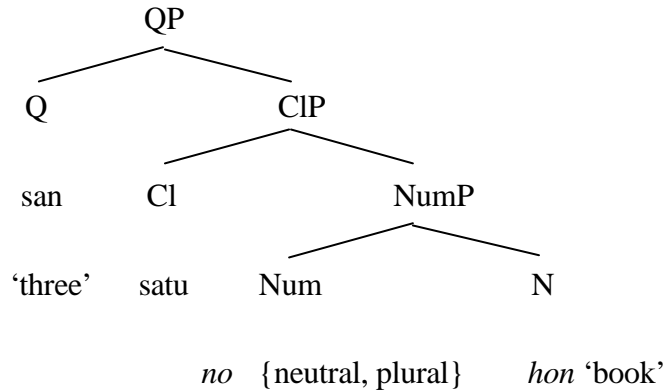
John floor-Loc find-Neg-Past dirt-Nom exist

‘(The fact that) There is a spot that John did not find on the floor.’

⇒ * Neg>indefinite (narrow scope): I did not see any spot on the floor.’

indefinite>Neg (wide scope): There is a spot that I failed to see on the floor.’

(43) The Nominal Structure of Bare Nouns in Japanese



The above-noted morphosyntactic profile of nominals in Japanese directly follows. First, Japanese allows bare nominals in any argument position because there is no DP projection in the structure in (43). Second, Japanese has a generalized classifier system due to the projection up to QP which dominates the CIP. Third, the extension of all nouns is mass for the same reason that the extension of all nouns is mass in Javanese and Indonesia: whichever value the Num head takes, the denotation of the NumP is a kind, which is mass. Finally, bare nominals in Japanese can only take narrow scope with respect to negation due to their kind-denoting requirement that is independently known to block wide scope readings. Thus, the morphosyntactic profile of [+arg, -pred] languages as in Japanese follows as an automatic consequence of the fact that Japanese nominals project only up to QP without the DP projection with the number values {neutral, plural}. Notice that the proposed analysis of

nominal structure in Japanese predicts that there are no grammaticalized mass-count distinctions in nominals, as exhibited by the *oat* vs. *wheat* pair, that is not solely computable from its conceptual realization. This prediction receives support from the observation that the translation equivalents of English mass nouns such as *furniture* are count nouns in the lexicon of Japanese speakers; it represents a conceptually identifiable discrete object such as couch, chair, and so on. The proposed analysis also predicts that bare nominals in Japanese should be able to exhibit scope interaction with negation if we introduce numerals. The comparison of (42a) and (44) shows that this prediction is confirmed.

(44) John-ga yuka-de mittu-no yogore-o mituke-naka-tta (koto)

John-Nom floor-Loc three-Link dirt-Acc find-Neg-Past (fact)

‘(The fact that) John did not find a spot/any spot/spots on the floor.’

⇒ Neg>indefinite (narrow scope): I did not see three spots on the floor.’

indefinite>Neg (wide scope): There are three spots that I failed to see on the floor.’

This fact thus provides strong support for the assumption that singularity is not a grammar-internal property of Japanese but rather is the byproduct of semantic disambiguation at the semantic interface.²¹

The idea that Japanese does not have anything like D heads in English has been a traditional one within generative research on this language since the seminal work by Fukui (1986, 1995) and Fukui and Speas (1986). Noguchi (1997) provides further arguments in favor of this position based on the contrast between English and Japanese with respect to the availability of variable binding for pronouns. One potential candidate for the D head in Japanese is a class of demonstratives such as *kono* ‘this’ and *ano* ‘that’. However, the iterativity test we introduced above in our discussion of Indonesian and Javanese shows that these elements are not D heads but instead modifiers of some other projection such as NP or the whole QP in the structure in (43) because, as Fukui (1986, 1995) notes, these elements can co-occur with other possessor or pronominal adjectives, as illustrated in (45c).

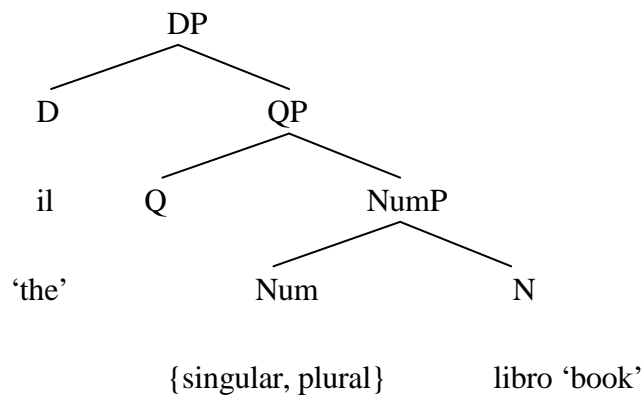
²¹ As Heidi Harley (personal communication) points out, the present analysis of scope variability in Indonesian/Javanese and Japanese allows us to raise a new question of why the phrase headed by the numeral phrase can undergo QR but bare nominals cannot. I speculate that numerals serve (part of) the same discourse function as Ds; D has the function of mapping the property denoted by a common noun (NP) onto whatever real-world entity the property holds true for. This discourse-oriented function is also shared by numerals.

- (45) a. ano kuruma b. John-no ano kuruma c. ookina John-no ano kuruma
 that car John-Gen that car big John-Gen that car
 ‘that car’ ‘*John’s that car’ ‘*big John’s that car’

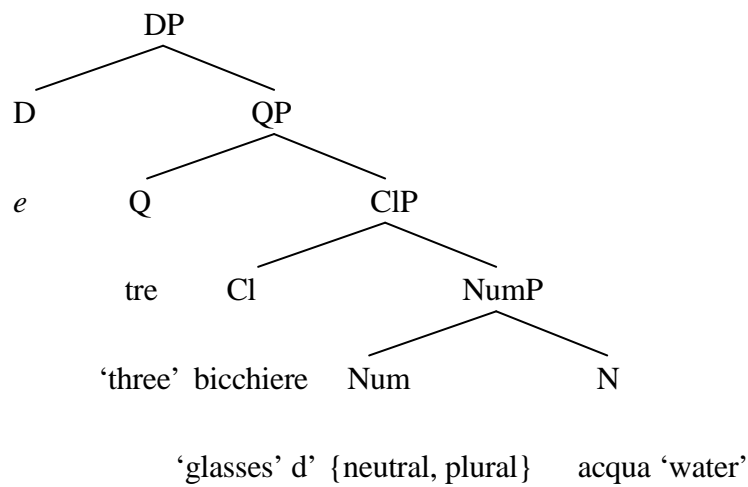
(Fukui 1995: 106, 107)

Let us now consider the nominal syntax of bare nouns in Italian. Italian, being one of the examples of the [+arg, -pred] language under Chierchia’s (1998a, b) Nominal Mapping Parameter, does not allow bare arguments. The unavailability of bare nominal arguments in this language directly follows if we assume that Italian nouns must always project up to DPs, hence instantiate the maximally complex nominal structure among languages of the world. We have also seen that languages such as Italian can take either {singular, plural} or {neutral, plural} for the Num value. The structure for Italian nominals, thus, will be as in (46) or (47), depending on whether the Num value is specified either as {singular, plural} or {neutral, plural}. Note that nominals in Italian project up to DP in both cases. I assume that *de* ‘of’ in the structure in (47) is inserted post-syntactically as a linker.

(46) Minimal Nominal Projection in Italian (for count nouns)



(47) Minima Nominal Projection in Italian (for mass nouns)



We saw in section 2 that Italian disallows bare nominal arguments. This property falls out because Italian nominals always project up to the DP.

Evidence that Italian nominals always project up to DPs comes from the subject-object asymmetry in Italian noted by Chierchia (1998a: 356), who observes that bare nominal arguments are allowed in direct object positions in certain cases but never permitted in subject positions, as illustrated by the contrast between (48a) and (48b). The same observation is also made by Longobardi (1994: 616), who points out the contrast between (49a) and (49b, c).

- (48) a. *Bambini sono venuti da noi. b. Ho preso bicotti con il mio latte.

kids be come by us I-have taken cookie with the my milk

‘Kids came by us.’

‘I ate cookies with my milk.’

- (49) a. * Acqua viene giù dalle colline.

water comes down from the-hills

‘Water comes down from the hills.’

- b. Viene giù acqua dalle colline.

comes down water from the-hills

‘Down from the hills comes water.’

c. Ho preso acqua dalla sorgente.

I took water from the-spring

‘I took water from the spring.’ (Longobardi 1994: 616)

The asymmetry observed here naturally follows if Italian nouns always project up to a DP with an empty head; see Contreras (1986) for a similar analysis in Spanish. One standard assumption in the generative framework has been that empty heads must be properly licensed by appropriate heads (Chomsky 1981, 1986a, b; Rizzi 1990). Under this assumption, the null D head in (48b) that dominates the bare noun *biscotti* ‘cookie’ is correctly licensed by the verbal head *preso* ‘take’. This licensing option is unavailable for the empty D head that dominates *bambini* ‘kids’ in (48a). A similar story holds for the contrast between (49a) and (49b, c). Thus, the subject object asymmetry provides support for the DP structure for Italian nominals. Alternatively, one could maintain that nominals in Italian themselves do not project to DP but the specifier position of TP requires a D head for EPP reasons. Then, nominals in subject position must be headed by a D head whereas nominals in object position need not (hence cannot) be headed by a D head.²²

²² I thank Andrew Carnie (personal communication) for suggesting this alternative analysis and Simin Karimi (personal communication) for directing my attention to Contreras (1986).

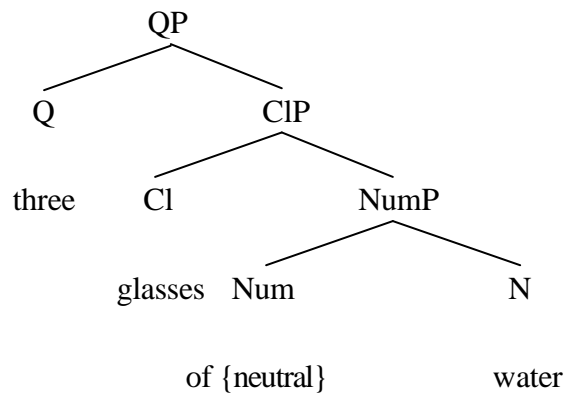
Consider finally the syntax of bare nominals in English. We have seen in the previous section that, under Chierchia's Nominal Mapping Parameter, English belongs to the [+arg, +pred] language type. This means that this language behaves as Japanese and Chinese in that the extension of its bare plural and mass nouns is a kind (hence [+arg]) whereas behaving as Italian and French in that it prohibits count nouns from occurring without determiners (hence [+pred]). I propose that this dual behavior of English nouns is exactly what we predict under the proposed analysis if English can choose the Japanese-type QP-structure or the Italian-type DP structure. Consider first the Japanese-type structure assigned to English when bare plurals and mass nouns are involved. In this case, English allows bare arguments, requires a classifier system, and does not have plural morphology, as in (50a-c). These properties mirror exactly those observed in Japanese.

- (50) a. I drank water.
 b. I drank three glasses of water.
 c. * I drank waters.

Thus, I propose the nominal structure for bare plurals and mass nouns as shown in (51), which is the Japanese-type nominal structure; it projects up to QP with the Num value

being chosen from the {neutral, plural} set. I assume that the preposition *of* is inserted in the post-syntactic morphological component between classifiers and their nominal heads as in *three glasses of water*.

(51) The Nominal Structure in English (for bare plurals and mass nouns)



The three morphosyntactic properties of bare plurals and mass nouns in English noted above are derived automatically by virtue of the fact that English has the Japanese-type QP structure in this context. The bare nominal option is possible because there is no DP on top of the QP. The Num specification in (51) requires that the denotation of the NumP be a kind. Thus, a certain set of classifier-like expressions such as *glass*, *cup*, and *piece* is required for nouns in (51) to set up an appropriate counting level for each noun, just as in languages like Indonesian, Javanese, and Japanese wherein all bare nouns denote a kind. There is no plural morphology observed in bare

plurals or mass nouns because they are true in an undifferentiated manner of a singular or plural instance of the entity denoted by this type of noun.²³

If English is like Japanese, the proposed analysis also leads us to the prediction that bare plurals and mass nouns cannot take wide scope over negation due to their kind-denoting requirement. This prediction is indeed confirmed by (5a), repeated here as (52).

(52) I didn't see spots on the floor.

⇒ Neg>Indefinite (narrow scope): I did not see any spot on the floor.

*Indefinite>Neg (wide scope): There is a spot/are certain spots that I failed to see on the floor.

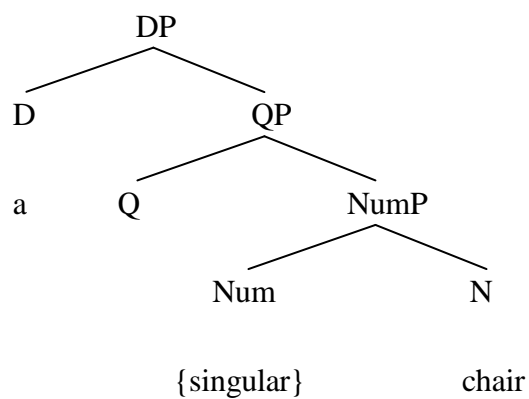
Consider now the structure for count nouns in English. When count nouns are involved, English does not allow bare nominal arguments (53a, b), lacks any classifier (53c), has plural morphology (53d), a cluster of properties that we have seen to characterize nominals in Italian.

²³ I have no good answer at this point of research for what would be wrong with using the null varieties of Q, Cl, and the Num heads in the derivation in (51) if D could be phonologically unrealized in Italian. This question is one special case of the broader question of why there are certain tendencies in natural language for a particular value (such as third person, not first or second person) to be null rather than overt. One might find clues to this question from learnability, but this is a speculation.

- (53) a. I bought a pen.
 b.* I bought pen.
 c.* I bought a piece of pen.
 d. I bought pens.

Thus, English takes the Italian-type nominal functional structure shown in (54). I assume that the indefinite article *a* is base-generated under the D head. Alternatively, *a* realizes the Q head (that is raised to the D head), since it denotes a singular instance of a discrete, countable entity; recall Chierchia's (1998b: 91, 92) observation noted above; see also Longobardi (1994).

- (54) The Nominal Structure in English (for count nouns)



The determiner-less bare option for count nouns is impossible for count nouns because the nominal structure projects up to the DP. When the singular value is selected for the Num head, the denotation of the NumP is a singular instance of the chair, which is compatible with the function of the indefinite article *a*. When the plural value is selected instead, the denotation of the NumP is bare plurals, which is a name of kind, as argued for in Carlson and Chierchia.

There are two potential problems with the proposed analysis for nominal syntax in English. The first problem concerns the selectional relation between the Num head and its N complement.²⁴ The question is, why it is that only bare plurals and mass nouns are inserted in the structure in (51) whereas only count nouns are inserted in the structure in (54)? What blocks count nouns and bare plurals/mass nouns from being inserted in the structures in (51) and (54), in that order? I maintain that there is actually nothing wrong with this choice *as far as syntax is concerned*; the syntax-external component interprets whatever syntactic object the narrow syntax creates and sends out. This position is related to what Hinzen (2006) calls *Semantic Blindness*; “as if syntax carved the path interpretation must blindly follow” (Uriagereka 2002: 275, as quoted in Hinzen 2006: 250). In other words, “the human language faculty provides forms that a possible human structured meaning may have, leaving a residue

²⁴ I thank Heidi Harley (personal communication) and Vicki Carstens (personal communication) for this question and useful discussion on how to address this question.

of non-structured meanings (concepts)." (Hinzen 2002: 235). It has been widely known that, when the meaning of an open class nominal element clashes with that of determiners, it is always determiners whose interpretation molds that of nouns (Harley and Noyer 2000: Harley 2006: 213-214). For example, mass nouns could occur with determiners that specifically select count nouns; thus, not only (55a) but also (55b) is possible in English.

- (55) a. I don't drink much coffee. (56) a. I had a cookie for breakfast.
 b. I bought two coffees this morning. b. That baby has cookie all over his face.

With an appropriate context, *two coffees* in (55b) can be interpreted as packaged coffees in cups or bags. Similarly, we can use count nouns together with determiners that specifically select mass nouns. In (56b), the bare nominal *cookie* is not interpreted as a discrete entity but instead as amorphous substance that cookies are generally made of. ((55a, b) and (56a) are from Harley 2006: 213.) These examples illustrate that the meaning of the noun is always bent to be compatible with the semantic contribution of the determiner which it co-occurs, not the other way around. In other words, there is no *ungrammatical* combination of nouns and functional elements within syntax *per se*, with the semantics trying its best to get a felicitous interpretation that is compatible with

world knowledge. Thus, when we have sentences like (57), to which no stretch of our encyclopedic knowledge can assign any reasonable interpretation, they are anomalous solely by our syntax-external criteria.

(57) # I had three oxygens in the kitchen. (Harley and Noyer 2000: 21)

The second potential problem with the proposed analysis concerns the modifiability of elements like *glasses*, *cups* and *pieces*, which I have thus far analyzed as a classifier on a par with Japanese classifiers.²⁵ This unified treatment appears to be incorrect, given that the former may be modified by adjectives while the latter may not, as the contrast between (58a) and (58b) shows.²⁶

(58) a. three big cups of coffee

b. * **ni** ookina **hai-no** kohii
 two big Class-Gen coffee
 ‘two big cups of coffee’

²⁵ As pointed out by Jaeshil Kim (personal communication) when I presented material in this chapter at the Mid-America Linguistic Conference held at the University of Kansas, Lawrence in October 2007.

²⁶ (58d) is acceptable in a context where the speaker talks about brands of coffee beans, for example.

- c. *ookina* **ni-hai**-no *kohii*
 big two-Class-Gen coffee
 ‘two big cups of coffee’
- d. #*ookina* *kohii*
 big coffee
 ‘big coffee’

The contrast between (58a) and (58b) receives a natural treatment in prosodic terms. We saw in section 3.1 that the numeral *se*- ‘one’ in Indonesian needs an overt classifier on its right due to its clitic/boundedness requirement. Following this line of thinking, (58b) is ungrammatical because the adjacency between the numeral *ni*- ‘two’ to the classifier *hon* is blocked by the intervening adjective *ookina* ‘big’ (cf. Bobaljik’s (1995) Morphological Merger under Adjacency). Thus, if we change the relative order of the adjective and numeral so that the two elements are phonologically adjacent, the result is grammatical, as in (58c). Notice that the adjective *ookina* ‘big’ modifies the classifier, not the head noun *kohii* ‘coffee’, in (58c), since the same adjective cannot co-occur with the same noun without the classifier *hai*, as in (58d).

To sum up this section, I have proposed a novel, relativized parametric theory of the morphosyntax and denotation of bare nominals across languages whereby different

morphosyntactic profiles exhibited by nominals in different languages are derivable from the interaction of two independently morphosyntactic/morphosemantic parameters: how high each language allows its bare nominal to grow and what set of Numb values each language can choose. The proposed analysis does not require any rigid mapping between the syntax and semantics of bare nominals, as in Chierchia's (1998a, b) Nominal Mapping Parameter but captures its effects through independently motivated observations about the available set of functional items and of the possible number values in each language, in a way that is compatible with the standard generative assumption about the locus of parameters.

5. New Typological Predictions of the Proposed Analysis

The relativized parametric theory of the denotation and morphosyntax of bare nominals makes certain predictions concerning several other languages and their morphosyntactic profile. In this section, I mention three languages, Child English, Russian and Chinese, whose morphosyntactic profile is naturally predicted by the proposed theory.

5.1. Child English Morphosyntax

The *Subset Principle*, proposed by Wexler and Manzini (1987), states that child language acquisition starts with the setting of the parameter values that rules out the most, so that

children can revise and modify their working hypotheses solely on the basis of positive evidence. Following the standard assumption that only positive evidence is available to children learning a language, Wexler and Manzini define the *Subset Principle* as in (59).

(59) The Subset Principle

Suppose one value of a parameter yields a language $L(i)$ and another value of the parameter yields a language $L(j)$. Suppose further that $L(i)$ is a smaller language than $L(j)$, that is, that $L(i)$ is contained in $L(j)$. $L(i)$ is a strict subset of $L(j)$. Then the learning strategy specified by the Subset Principle is that the learner select the value which yields $L(i)$ first. If this is the correct choice, there will never be evidence that it isn't, and the learner will stay with the value. If this is the wrong choice, then there will be positive evidence (sentences from $L(j)$) which are not in $L(i)$) which the learner will eventually hear; this evidence must exist, because $L(i)$ is a strict subset of $L(j)$. The Subset Principle specifies that when positive evidence which shows that $L(i)$ is the wrong language is encountered, the learner will switch to the parameter value that yields language. In short, the Subset Principle is a method for specifying a markedness hierarchy when alternative values yield languages that are in a subset relation. (Wexler and Manzini 1987: 44)

As mentioned in the previous section, the idea that children learn nominal syntax in a bottom-up way has been independently proposed by Guilfoyle and Noonan (1992) and supported by Vainikka (1993/1994). Guilfoyle and Noonan argue that Functional Grammar is not present in the early stages of language acquisition based on well-known acquisitional properties such as the optionality of subjects, the absence of NP-movement, subject-auxiliary inversion, and verb placement facts. To repeat one argument based on the absence of NP-movement, Borer and Wexler (1987) observe that children treat all passives as adjectival passives and claim that this observation naturally follows under the syntactic NP-movement account of verbal passives (see Baker et al. 1989 for a detailed analysis) if children cannot form A-chains. This maturation analysis is untenable according to Guilfoyle and Noonan because principles of UG such as the θ -Criterion, the EPP, and the Projection Principle (Chomsky 1981, 1986a) built into their grammar would not be able to block their association of two structural positions with a single θ -role. Instead, Guilfoyle and Noonan argue that the above observation is straightforwardly derived if there is simply no IP in the child language syntax, so that the movement of an NP can target: as a result, children treat all cases of passive as adjectival because formation of this type of passive is a process that does not require IPs (Wasow 1977; Levin and Rappaport 1986). This ‘bottom-up’ incremental view of phrase structure is also supported by Vainikka’s

(1993/1994) developmental evidence that the emergence of nominative subjects occurs around the same time as that of TP-related material like modals, auxiliaries, and past tense.

The relativized parametric theory, in combination with the Subset Principle, makes an interesting prediction; early stages of acquisition of nominal syntax in all languages should mirror the morphosyntactic profile of bare nominals in Javanese and Indonesian because these two languages instantiate the simplest nominal functional structure (NumP). Specifically, we predict that initial stages of all child languages should allow bare nominal arguments, lack a classifier system for nouns, and iii) have plural morphology. This prediction is indeed borne out by what the literature on acquisition calls the *Telegraphic Speech*, as illustrated in English examples as in (60); see also Chierchia (1998a: 400) for a similar remark made on the basis of his Nominal Mapping Parameter.

(60) Telegraphic Speech ²⁷

25	[dan? ɪ? t ^s ɪ?]	“don’t eat (the) chip”
	[b ^w d? tat]	“block (is on) top”

²⁷ The number on the left-hand side indicates the age in months at which the utterances on the right were uttered.

26	[mamis tu hæʃ]	“Mommy’s two hands”
	[mo bʌʃ go]	“Where bus go?”
	[dædi go]	“Where Daddy go?”
27	[ʔaj gat tu dʒʊʃ]	“I got two (glasses of) juice”
	[do bajʔ mi]	“don’t bite (kiss) me”
	[kʌdər sʌni ber]	“Sonny color(ed a) bear”
28	[ʔaj gat pwedɪʃ]	“I (‘m) play(ing with) this.”
	[mamis tak mɛns]	“Mommy talk(ed to the) men”

(Fromkin et al. 2003: 365)

The examples illustrate use of generalized bare arguments (i.e. *chip*, *block*, *bus*, *juice*, *bear*), lack of classifiers for mass nouns (i.e. *juice*), and presence of plural morphology (*hands*, *men*), a clustering of properties that we have seen to hold for Javanese and Indonesian. This behavior of nominals in Child English therefore provides evidence for the proposed cross-linguistic analysis of bare nominals and for the Structure Building Hypothesis of Guilfoyle and Noonan (1992) from the domain of nominal syntax.

The current analysis also makes two other predictions concerning child English syntax. First, nominals in child English syntax should show the interpretive property of taking narrow

scope with respect to negation and lack plural forms for generic statements, as in Indonesian and Javanese. I have not found any literature that focus on this specific property in child language acquisition.²⁸ However, the experiment conducted by Pérez-Leroux and Roeper (1999) provides strong experimental evidence that English-learning children are sensitive to the presence or absence of D heads in calculation of differences in scope between the following pair of examples: *Everybody went home* vs. *Everybody went to his home*. Second, recall that I have argued above that expressions such as demonstratives should be analyzed as modifiers of the NP projection in D-less languages such as Indonesian, Javanese, and Japanese, on the grounds that this type of element can co-occur with other modifying elements such as possessives and adjectives. If child English starts as Indonesian and Javanese, we predict that there should be instances where what are analyzed as D elements in Adult English should be able to co-occur with similar types of elements. This prediction is borne out by examples as in (61a-c), documented by Miller and Ervin-Tripp (1973), where “determiners” co-occur with possessives and adjectives.

²⁸ As Cecile McKee (personal communication) points out, there is a linkage problem as well because children in the scope experiments conducted in current acquisitional study often tend to be well past the telegraphic stage.

- (61) a. Is that the blue mine?
 b. this a Bonnie pants
 c. I know a that.
 d . These a Lidz pants
 e. mine, all a mine. (Miller and Ervin-Tripp 1973: 363)

The iterability of “functional” elements in child syntax, therefore, provides further evidence that nominal syntax in Child English does not project up to the DP for count nouns. This property also supports my argument, made in section 4, that languages do whatever they can do with their available syntactic and lexical resources available to their grammars to arrive at the same interpretation that other languages would yield with more elaborated functional structures.

5.2. *Slavic MorphoSyntax*

The relativized parametric theory proposed in this chapter also sheds new light on the nominal syntax in Slavic languages such as Russian, which Chierchia (1998a, b) briefly mentions as a language of the [+arg, +pred] type on a par with English. The morphosyntactic profile of Russian is as follows: i) bare nominals are permitted in argument positions, ii) a generalized

classifier system is missing, iii) bare nominals may take either wide or narrow scope with respect to negation. This last scope-related property is illustrated in (62a, b).²⁹ Note that this contrasts sharply with the behavior of bare nominals under negation in Indonesian and Javanese illustrated above.

(62) a. Ya ne vizhu ni odnogo pyanta na polu.

I Neg see no one spot on floor

⇒ Neg>indefinite (narrow scope): I did not see a single spot on the floor.

*Indefinite>Neg (wide scope): There is a spot on the floor that I failed to see.’

b. Ya ne zametila ni odnogo pyatna na polu.

I Neg notice no single spot on floor

‘There might have been one spot on the floor that I failed to notice.’

⇒ Neg>indefinite (narrow scope): I did not see a single spot on the floor.’

indefinite>Neg (wide scope): There is a spot on the floor that I failed to see.’

Under the proposed analysis, the morphosyntactic profile of Russian falls into place if bare nouns in this language project up to NumP, as in Javanese and Indonesian, but with its

²⁹ Thanks to Tatyana Slobodchikoff (personal communication) for useful discussion on Russian syntax.

possible set of the Num values being either {singular, plural} or {neutral, plural}, as in English and Italian. The availability of bare nominal arguments and the lack of a generalized classifier system results from the NumP nominal structure. The scope variability of bare nominals as shown in (62a, b) results when a N is selected by the Num head with the singular specification, as in English and Italian count nouns. In this way, the proposed analysis can serve to classify languages like Russian in terms of the complexity of nominal projections and the possible number values and derive their morphosyntactic profile from the interaction of the two parameters.

5.3. *Chinese Morphosyntax*

The proposed theory also has something to say about the nominal syntax in Chinese. Recall the universal nominal morphosyntactic hierarchy shown in (17). I have provided evidence that the DP, QP, and NumP options are instantiated by Italian/English (for count nouns), Japanese/English (for bare plurals and mass nouns) and Javanese/Indonesian, respectively. The proposed analysis predicts that there are languages that instantiate the ClP option. Of course, there may be conceptual reasons why languages of this type are not easy to find. As Iljic (1994: 104) notes (see also Croft 1994), classifiers have the function of individuation, which makes it possible to extract “discrete occurrences.” This individuation function may well naturally tie with

that of numerals to produce expressions as in *three cups of coffee*, thereby blocking the CIP option from being utilized across the board in natural language syntax. However, as we have seen above, English can choose among structures depending on the nature of the bare nouns involved. Therefore, the proposed analysis still leads us to expect that some languages might instantiate the CIP option in a restricted range of circumstances. Importantly, Cheng and Sybesma (1999) show that definite bare nominals in Mandarin and Cantonese project up to CIPs whereas indefinite bare nominals in these languages project up to Numeral P (QP under the proposed analysis). According to their view, the definite interpretation is derived from the CIP structure under their assumption that N-to-Cl movement feeds the generation of the ι -operator, which, according to Chierchia (1998a: 359), “defines selects the greatest element from the extension of a predicate and constitutes typically the meaning of a definite article.” (see Cheng and Sybesma 1999; 524 for detailed discussion). The indefinite interpretation, on the other hand, is derived from the NumeralP/QP option since “the numeral apparently has the effect of undoing the definiteness.” (Cheng and Sybesma 1999: 524). The argument above therefore shows that Cantonese and Mandarin are optional CIP languages.

This analysis further allows us to correctly predict several other morphosyntactic properties of bare nominals in Mandarin and Cantonese that are indeed borne out by Cheng and Sybesma’s findings. First, these two languages allow bare nominal arguments

because these languages do not project up to DPs. Second, these languages have a generalized classifier system due to the projection up to CLPs or QPs that dominate them. Third, Cheng and Sybesma (1999: 519) observes that “Chinese bare nouns can be interpreted as both singular and plural.” This property follows from the choice of the feature set [neutral, plural] in these languages.³⁰ Finally, bare nominals should not be able to take scope over negation, a prediction that is borne out by the obligatory narrow scope of bare nominals as in (63a) in contrast to (63b).³¹

(63)a. di ban shang yī gè bàn-diǎn wǒ měi kàn-dào.

floor board on one Class spot I Neg see

‘I did not see a spot on the floor.’

⇒ Neg>Indefinite (narrow scope): I did not see any spot on the floor.

* Indefinite>Neg (wide scope): There is a spot that I failed to see on the floor.’

³⁰ As stressed by Heidi Harley (personal communication), though, this feature set itself does not exclude the possibility of plural morphology in Mandarin and Cantonese because that is exactly the feature set in Indonesian, Javanese, and Japanese, which all have such morphology. One might regard (yī) *xīē* and *–men* as two possible markers of plurality in Mandarin but this view is quite controversial because Iljic (1994) argues that these morphemes are not plural but rather collectives that refer to wholes. For this reason, I leave the issue of whether Mandarin/Cantonese has genuine plural morphology open here.

³¹ I thank Sunjing Ji (personal communication) for her help in constructing the examples in (63a, b).

b. di ban shang you ge ban-dian wo mei kan-dao.

floor board on exist Class spot I not see

‘There is a spot that I failed to see on the floor.’

⇒ * Neg>Indefinite (narrow scope): I did not see any spot on the floor.

Indefinite>Neg (wide scope): There is a spot that I failed to see on the floor.

In this way, the proposed analysis correctly characterizes the morphosyntactic profile of bare nominals in Mandarin/Cantonese by assuming that they project up to CIPs/QPs depending on the definiteness with the Num value being set as {neutral, plural}.

5.4. *Other Languages*

The relativized parametric theory of nominal morphosyntax may well predict some other language types once the research goes beyond those languages discussed in this chapter. Thus, in a series of recent work, Schmitt and Munn (1991, 2002) present evidence that Brazilian Portuguese provides a counterexample to Chierchia’s Nominal Mapping Parameter and argue for a purely syntactic account of the morphosyntactic profile of bare nominals in this language by extending Bobaljik’s (1995) Free Agr Parameter to nominal domains. Importantly, Schmitt and Munn observe that bare singulars in this language must take narrow scope with

respect to negation and are underspecified for number. Under present assumptions, this observation means that the Num value must be selected from {neutral, plural} set as in Javanese, Indonesian, Japanese, and Mandarin/Cantonese. Schmitt and Munn conclude that bare singulars in Brazilian Portuguese are DPs. Within the present framework, combining this conclusion with the above-noted observation on Number, we can conclude that bare nominals always project up to DP, as in Italian and English, with the Num head being specified for {neutral, plural}, as in Javanese, Indonesian, Japanese and Mandarin/Cantonese. Schmitt and Munn note many detailed interpretive subtleties involved in the use of bare singulars caused by the nature of predicates, episodic contexts, and so on, so the situation is more complex than stated here. However, the fact that the core morphosyntactic property of bare nominals in Brazilian Portuguese can be characterized as the interaction of the morphosyntactic complexity with the possible Num values indicates that the proposed analysis is on the right track.

6. Chapter Summary: Theoretical Implications for the Syntax-Semantics Interface

Let us summarize what we have done in this chapter. I have discussed the issue of syntax-semantics interface with reference to the denotation and morphosyntax of bare nominals in several languages, in particular, Indonesian and Javanese. I have shown that these two languages do not fit into any one of the three language types predicted by Chierchia's (1998a,

b) Nominal Mapping Parameter. First, the free occurrence of bare arguments in Indonesian and Javanese shows that these languages are not [-arg, +pred] languages such as Italian. Second, the absence of a generalized classifier system and the presence of plural morphology marked by reduplication of a root argue against categorizing the two languages as [+arg, -pred] languages such as Japanese and Chinese. Finally, the obligatory narrow scope of bare nominals with respect to negation and the lack of pluralized/reduplicated forms for generic statements suggests that Indonesian and Javanese are also not [+arg, +pred] languages such as English. This result casts serious doubts on the rigid mapping between the syntax and semantics of bare nominals of the kind assumed in Chierchia's Nominal Mapping Parameter.

Following the standard conception of the locus of parameters in the Principles-&-Parameters approach to language variation (Borer 1984; Fukui 1986, 1995; Chomsky 1981, 1986b, 1995), I have proposed a related parametric theory of nominal denotation that derives different morphosyntactic profiles of bare nominals in different languages from the relative complexity of nominal functional structures and the possible set of Num values that are available in each language. The complete summary of this theory is summarized in Table 4 with a full range of languages discussed in this paper.

Table 4: A Relativized Parametric Theory of Nominal Denotation (final version)

Languages	Num Values	Nominal Syntax
Indonesian, Javanese	{neutral, plural}	NumP
Russian	{singular, neutral, plural}	NumP
Chinese	{neutral, plural}	CIP
Chinese, Japanese	{neutral, plural}	QP
Italian	{singular, plural, neutral}	DP
English	{singular, plural, neutral}	DP or QP

I have also shown that the proposed analysis makes new predictions concerning a few other languages that have not been discussed in detail in Chierchia (1998a, b). Coupled with the Subset Principle of Wexler and Manzini (1987), the proposed analysis predicts that nominals in the early stages of language acquisition should show the same set of morphosyntactic properties as those in Javanese and Indonesian because these languages instantiate the simplex nominal functional structure. I have shown that this prediction is borne out by telegraphic speeches produced by English-learning children. I have also argued that Russian is minimally different from Javanese and Indonesian in that the

former projects only up to NumP as in Indonesian and Javanese but with the Num values being either {singular, plural} or {neutral, plural}, as in English and Italian.

As is true with the other chapters in this dissertation, the major goal of this chapter has been to propose a new theory of language variation within the Principles & Parameters approach with special reference to the morphosyntax of bare nominals. The case study reported here highlights the importance of under-studied languages as in Indonesian and Javanese because generalizations concerning these languages often deviate from those that have been found among relatively well-studied languages such as English, Romance, and East Asian languages, hence reveal crucial gaps that otherwise would go unnoticed in modern syntactic theorizing. I have shown here that detailed examination of these languages as conducted here is instructive in informing a more “balanced” parametric theory of human language syntax, a desideratum in the current generative enterprise.

The proposed analysis has several important implications for the proper theory of syntax-semantics interface. First of all, the proposed analysis of bare nominals across languages provides support for a certain conception on the economy of derivation and projection at the syntax-semantics interface: the syntax-external interpretive component employs whatever syntactic resources are available to the language to express the same denotation that other languages would express with a more complex nominal functional structures. This was seen,

for example, in the treatment of demonstratives and numerals as adjuncts/modifiers of lexical projections such as NPs in Indonesian, Javanese and Japanese, which would constitute D and Q heads in languages such as Adult English and Italian. This point also applies to the morphosyntactic development of bare nominals in Child English, which suggests that child language acquisition also respects structural economy in the sense that it posits minimal structures necessary to analyze the data available to children until it ends with the QP/DP structure (Radford 1990). We also have noted that various denotations assigned to a particular nominal element (kind, predicate, indefinite, etc.) are the interpretive outcome of the semantics component that interprets a different height of the nominal functional projection and of the Num values. Similarly, when the functional requirement of a determiner clashes with the conceptual structure of its complement nouns (mass vs. count), it is always the former that prevails, forcing the semantic interpretive component to interpret the output of syntax in every way compatible with our knowledge of whether particular nouns can be conceptualized as discrete individualizable objects or amorphous discrete mass and how. These results provide strong support for the notion of “minimalist interfaces” introduced in chapter 1, whereby the syntactic computation provides a parametrically defined curve that the conceptual/semantic interfaces must blindly follow, without any extrinsically determined mapping between the syntax and semantics of a particular expression as in Chierchia’s (1998a, b) Nominal Mapping

Parameter. The manner in which the syntactic representation is mapped onto the semantic representation is language-invariant. The narrow syntactic computation will do whatever it can within a parametrically chosen set of morphosyntactic features and their projections in each language, and the universal semantics will just come up with an interpretation that is compatible with our conceptual knowledge of how things are represented in the external world.

Second, the proposed parametric theory of bare nominals argues against the common assumption in the generative literature (e.g., Higginbotham 1985; Stowell 1989; Szabolcis 1987, 1994; Longobardi 1994; Heim and Kratzer 1998) that it is only DPs that can serve as arguments; see also Baker (2003: 113) for the claim that D is not necessary for argumenthood.³² This assumption is understandable, given that NPs denote $\langle e, t \rangle$ whereas DPs denote a type $\langle e \rangle$; as a result, a NP must combine with a D to be saturated and computed as type $\langle e \rangle$. This assumption also makes sense if D can be considered to have the function of mapping the set of entities denoted by the N-set to a specific/definite entity thereof. To the extent that my analysis is correct, however, the proposed syntax of nouns in languages such as Javanese, Indonesian, Japanese, Russian, and English (for bare plurals and mass nouns) provides strong evidence against this commonly held view.

³² Recall that Chierchia's Nominal Mapping Parameter as well, DP is not necessary for argumenthood.

It is important to note in this context that there have been considerable debates concerning whether nouns in Slavic/article-less languages such as Serbo-Croatian, Bosnian and Polish are associated with DPs or NPs. See Progovac (1998), Bašić (2004), Leko (1999), Rappaport (2001), Rutkowski (2002), Rutkowski and Maliszewska (to appear), Franks and Pereltsvaig (2004), Trugman (2005), and Pereltsvaig (2007) for the DP analysis; see Corver (1992), Willim (1998, 2002), Trenkic (2004), Bošković (2003, 2004, 2005), Zlatić (1997), and Stepanović (1999) for the bare NP analysis. If the proposed analysis of Russian morphosyntax in section 5 is correct, it argues for the parametrized NP hypothesis for Slavic, hence necessitates a large-scale re-examination of the data in this language family that have been adduced in favor of the DP hypothesis.

I conclude this chapter with the following important question, which the presnet analysis would bring to light if it is on the right track. As stated above, the proposed theory claims that the denotation of a functional catgeory is language-invariant. For languages like Italian and English (in certain cases) whose nominals always project to DP, the syntax-semantics mapping is strictly compositional, for the reason stated above. Then, a question arises in those languages which I have argued to lack a D as to how a determiner-less nominal can denote an inidividual in the same way that a determiner-headed nominal does. To put in a more informal way, how can humans express information such as

definiteness and specificity in languages without D heads that other languages with D heads will express in fully syntactic ways given that those pieces of information are syntactically represented?³³ I believe that the present thesis of minimalist interfaces provides an interesting answer to this question. It is possible that some sort of repair takes place at LF in just as much as at PF (chapter 2). Under the present case, the linguistic semantic component may well develop general type-shifting operations to solve a type mismatch problem that necessarily arises in D-less languages. In fact, this is the tack that Chierchia (1998a) develops to account for the morphosyntactic behavior of Russian. More generally, given that all languages have some or other ways to express a particular thought/message, D-less languages will activate certain domain-specific operations that would remain unused in D-languages to express the same thought, whether via demonstratives that seem to have a similar function as D, general type-shifting operations, or even extra-linguistic contexts that impose the definite/specificity on a D-less noun. Of course, one could imagine that there should be certain restrictions on the kind of operations that the semantic component perform to remedy the crude syntactic representation; see Chierchia (1998a), for example, who claims that type shifting is a Last Resort operation at

³³ Thanks to Andrew Carnie (personal communication) and Simin Karimi (personal communication) for raising this question and useful discussion/suggestions.

the LF component. I leave detailed examination of the nature of such restrictions as an important task for future research.³⁴

The investigation conducted in this chapter, therefore, is construed as a case of the thesis of minimalist interface; the semantic component is subservient to the needs of the syntax, assigning whatever reasonable interpretation it can to the output of the syntax with several domain-specific operations in hand (type-shifting, contexts, pragmatics) but only within the parametrically defined carve set up by syntactic computation. I discuss further ramifications of this conjecture in chapter 7.

³⁴ It would be interesting to extend the parametric theory of nominal projections proposed here to clausal projections as well in order to see whether any parallelism holds between the nature of functional nominal structures and the nature of functional clause structures. Fukui (1986, 1995) argue that Japanese lacks functional categories such as T and C. The observation that Indonesian, Javanese, and Chinese all lack morphemes such as tense and agreement overtly manifested in Indo-European languages provides evidence that these languages may also lack verbal functional categories such as T and C. I wish to conduct a fuller examination of this parallelism in my future research. I am very grateful to Simin Karimi (personal communication) for suggesting this extension.

CHAPTER 5 WHETHER *WH*-IN-SITU MOVES OR NOT IN INDONESIAN:CHOICE FUNCTIONS AND INTERFACE ECONOMY ¹**1. Introduction**

In the last three chapters, I have explored the consequences of the thesis of the minimalist interface for the way syntax interfaces with the external phonological and semantic components. The results in chapters 2 and 3 based on the interaction of the active voice deletion with syntactic movement in Indonesian and Javanese as well as the amelioration of P-stranding violations by deletion show that the phonological component is fundamentally interpretive, subservient to the universal law of syntax, doing whatever it can within the parametrically defined range of options available to the linguistic interface. The results in chapter 4 based on the denotation and morphosyntax of bare nominals in Indonesian and Javanese as well as other languages show that the same characterization holds for the other syntax-external semantic component, which, I have argued, comes up with whatever semantic interpretation it can within the restricted set of morphosyntactic

¹ This is a substantially expanded version of the paper presented at the Mid-America Linguistic Conference held in the University of Kansas, Lawrence (October 2007). The paper presented there is to appear in Sato and Yuliani (forthcoming).

structures that each language parametrically selects from a universal set of morphosyntactic features and of possible number values.

The purpose of this chapter is to further substantiate the validity of the thesis of minimalist interfaces from the perspective of the syntax-semantics interface. For this purpose, I concentrate on the question of what the proper licensing mechanism of *wh*-in-situ is in Indonesian. The current minimalist interface thesis expects that this phenomenon is naturally explained as an optimal realization of the needs of the syntactic computation on the part of the semantic component that is responsible for instructions for the conceptual-intentional system. I argue that this is indeed what the facts on *wh*-in-situ in Indonesian actually attest by showing that it has syntactic and semantic properties that would remain mysterious under purely syntactic approaches to *wh*-questions and that their interpretation is straightforwardly derived by the asymmetric interaction of syntax with semantics via the notion of *Interface Economy* in the sense of Reinhart (1992, 1995, 1997, 1998, 2006), Fox (1998, 2000), and others.

The present chapter is organized as follows. In the next section, I provide a detailed overview of the syntactic and semantic properties of *wh*-in-situ in Indonesian. The discussion in this section owes a great deal to the seminal work on this construction by Saddy (1991), who observes that *wh*-in-situ in this language shows a spectacular range of

interesting properties such as the lack of island/ECP effects, the absence of pair-list readings and weak/strong crossover effects, and the quantificational uninformativity.² Drawing on Saddy's description and analysis, I show that commonly accepted analyses of *wh*-in-situ in other Asian languages such as Chinese and Japanese as in overt/covert *wh*-movement (Huang 1982; Nishigauchi 1986, 1990; Watanabe 1992, 2001; Richards 2001) and unselective binding (Baker 1970; Pesetsky 1987) cannot adequately account for these intriguing syntactic and interpretive characteristics uniquely associated with *wh*-in-situ in Indonesian. This discussion leaves us with the question we started with: what is the proper analysis of *wh*-in-situ in Indonesian. In section 3, I provide Saddy's own analysis of this construction as an interrogative definite description. Saddy argues, based on the interpretive and syntactic parallelisms between *wh*-in-situ in Indonesian and the formulation of "questions" on the game show "Jeopardy" in the United States, that the former receives the interpretation it has due to the fact that it is interpreted in the same way as a definite description is. I provide several empirical arguments against this treatment of Indonesian *wh*-in-situ. The relevant arguments concern the source of interrogative force in the two cases, the indefinite nature of *wh*-in-situ in Indonesian in terms of reduplication, and, most importantly, the observation that this type of *wh*-phrase behaves more like

² Although I present evidence against some of Saddy's reported results and analysis.

existential indefinites rather than definite descriptions, which do not contain a variable position. Based on this consideration, I show in section 4 that all the observed properties of *wh*-in-situ in Indonesian are naturally derived once we assume that the appropriate licensing mechanism of the *wh*-construal is via a choice function in the sense of Reinhart (1992, 1995, 1997, 1998, 2006). I further demonstrate that this analysis makes several correct empirical predictions, in particular, in the area of NP vs. non-NP asymmetries with respect to putative ECP effects. I conclude this section by comparing the choice function analysis with the fine-grained unselective binding approach to *wh*-in-situ in Indonesian recently developed by Cole and Hermon (1998, 2000). Though the two analyses initially appear to amount to the same predictions in the core cases, I argue that the choice function makes correct predictions concerning the so-called “Donald Duck” problem and the intermediate scope reading of *wh*-in-situ in multi-clausal contexts that would only be accounted for by ad hoc stipulations under Cole and Hermon’s unselective binding analysis. I conclude the present chapter in section 5 by discussing several theoretical implications of *wh*-in-situ in Indonesian and the proposed analysis for the proper view of the syntax-semantics interface and for the thesis of minimalist interfaces. I argue that the choice function analysis developed in this chapter is one optimal satisfaction on the part of the semantic component of the needs of the syntactic computation, further vindicating the

validity of the minimalist interface thesis, because a choice function interpretation can be considered as a syntax-external manifestation in the interface component of a failure that syntax itself cannot make up for.

2. *Wh-in-Situ* in Indonesian

This section provides an overview of the structural and interpretive properties of *wh*-in-situ in Indonesian. The discussion in this section draws heavily on the description and analysis of this construction presented by Saddy (1991). Accordingly, all the examples in this section are drawn from his work, unless otherwise indicated, though I have checked the same data with three native-speaker consultants for grammaticality judgments. I show that the two most widely held analyses of this construction in other Asian languages such as Japanese and Chinese – syntactic movement and unselective binding – are not the right approach to *wh*-construal for in-situ *wh*-phrases in Indonesian.

As we saw in chapters 2 and 3, Indonesian has three ways to form *wh*-questions: i) overt syntactic movement to the matrix, scopal [Spec, CP], ii) partial syntactic movement to the embedded, non-scopal [Spec, CP]s, and iii) *wh*-in-situ. The three strategies are illustrated in (1a-c), respectively.

(1) *Wh*-Questions in Indonesian

- a. [_{CP1} ***Apa***_i yang kamu pikir [_{CP2} Esti kira [_{CP3} Pak Yanto beli ***t_i*** kemarin]]]?

what that you think Esti expect Mr. Yanto buy yesterday?

‘What do you think Esti expects Mr. Yanto bought yesterday?’

- b. [_{CP1} Kamu pikir [_{CP2} ***apa***_i yang Esti kira [_{CP3} Pak Yanto beli ***t_i*** kemarin]]]?

you think what that Esti expect Mr. Yanto buy yesterday

‘What do you think Esti expects Mr. Yanto bought yesterday?’

- c. [_{CP1} Kamu pikir [_{CP2} Esti kira [_{CP3} Pak Yanto beli ***apa*** kemarin]]]?

you think Esti think Mr. Yanto buy what yesterday

‘What do you think Esti expects Mr. Yanto bought yesterday?’

In the example in (1a), the *wh*-phrase *apa* ‘what’ undergoes overt syntactic movement to the scopal, matrix [Spec, CP]. This option is always available for nominal *wh*-phrases such as *siapa* ‘who’ and *apa* ‘what’ and prepositional *wh*-phrases such as *cara apa* ‘in what way’, *dimana* ‘in what place, where’, and *untuk apa* ‘for what reason’, but obligatory for non-nominal *wh*-phrases such as *kenapa* ‘why’, *bagaimana* ‘how’, and *kapan* ‘when’; see section 4.3 for related discussion. The example in (1b) illustrates the partial syntactic movement option in Indonesian, where the same *wh*-phrase undergoes movement into the

intermediate, non-scopal [Spec, CP], though the example itself has a matrix *wh*-interpretation as the fully moved example in (1a). This option is available for nominal *wh*-phrases but not for non-nominal *wh*-phrases. Finally, the example in (ic) illustrates the in-situ option in Indonesian. This option is always possible for nominal *wh*-phrases but never possible for non-nominal *wh*-phrases, according to my consultants.

Saddy (1991) observes that the *wh*-in-situ construction in Indonesian exhibits a spectacular range of syntactic and semantic characteristics that would not be accounted for under standard analyses of the corresponding constructions in other languages such as English, Chinese, and Japanese. I review his main arguments in the rest of this section to show that the two most widely assumed analyses of *wh*-construal, syntactic movement and unselective binding, are not applicable for *wh*-in-situ in Indonesian.

2.1. “Overt” Syntactic Movement?

The first analysis of *wh*-in-situ in Indonesian, which is most easily dismissed, is the null operator/Q-feature movement in the syntactic component, as proposed by Watanabe (1992, 2001) and Richards (2001) for *wh*-in-situ constructions in Japanese (cf. Kayne 1998). Watanabe proposes that Japanese has an overt syntactic movement of a phonologically invisible *wh*-operator into the specifier of the scopal C, as in English.

Richards proposes, following the Single Cycle Model (Bobaljik 1995; Groat and O’Neil 1996; Pesetsky 1998), that the difference between “overt” and “covert” movement is purely phonological, determined by whether the head or the tail of a movement chain is pronounced; if the head is pronounced, we have “overt” movement whereas, if the head is pronounced, we have “covert” movement. This model, thus, eliminates the traditional distinction between overt and covert movement drawn in the classical minimalist period (Chomsky 1993, 1995). Working within this model, Richards argues that Japanese *wh*-in-situ occurs when the tail of the movement of the in-situ *wh*-phrase into the scopal [Spec, CP] is pronounced at the PF component.

The following consideration shows that the analysis of *wh*-in-situ presented in Watanabe (1992, 2001) and Richards (2001) for Japanese is not transportable to the corresponding construction in Indonesian. Syntactic island effects (Ross 1967) have been widely acknowledged as a textbook case for the occurrence of syntactic movement in the generative framework. Expectedly, overt syntactic movement in Indonesian shows various island effects, as in (2a-d).

(2) a.* *Apa_i* yang kamu katakan [dimana kita beli *t_i*]? (*wh*-Island)

what that you mention where we buy

‘What do you mention where we bought?’

b.* *Siapa_i* yang kamu suka [cerita yang mengkritik *t_i* itu]? (Complex NP Island)

who that you like stories that criticize the

‘Who do you like the stories that criticized?’

c.* *Siapa_i* yang kamu kira gambar *t_i* dijual? (Subject Island)

who Foc you think pictures be-sold

‘Who do you think pictures of were sold?’

d.* Dengan *siapa_i* yang kamu cemburui Bill [karena saya berbicara *t_i*]? (Adjunct Island)

with who Foc you get jealous of Bill because I spoke

‘With who did you get jealous of Bill because I spoke?’

(Saddy 1991: 190, 191)

If Watanabe’s/Richards’ analysis is correct for *wh*-in-situ in Indonesian, in the *wh*-in-situ counterparts to these examples, the relevant *wh*-phrase would undergo overt movement of the null operator into the matrix, scopal [Spec, CP] (for Watanabe) or the tail of the movement is pronounced at PF (for Richards). Thus, this analysis would predict that the

in-situ counterparts to (2a-d) should be as ungrammatical as (2a-d). This prediction is false, however. Saddy shows that *wh*-in-situ in Indonesian is a well-formed question in this environment, freely taking the matrix interrogative interpretation outside of syntactic islands. This point is illustrated by the grammaticality of the examples in (3a-d).³

- (3) a. Kamu katakana [kita mem-beli **apa** dimana]? (*wh*-Island)
 you mention we AV-buy what where
 ‘What did you mention where we bought?’
- b. Kamu suka [cerita yang mengkritik **siapa** itu]? (Complex NP Island)
 you like stories that criticize who the
 ‘Who do you like the stories that criticized?’
- c. Kamu men-gira gambar **siapa** dijual? (Subject Island)
 You AV-think pictures who be sold
 ‘Who do you think pictures of were sold?’

³ (3a) is modified from Saddy (p. 190) by changing the verb from *ingat* ‘remember’ to *katakan* ‘mention.’

d. Kamu men-cemburui Bill [karena saya berbicara dengan *siapa*]? (Adjunct Island)

you AV-jealous of Bill because I spoke with who

‘With who did you get jealous of Bill because I spoke?’

(Saddy 1991: 190, 191)

The fact that these examples are all well-formed *wh*-questions, thus, provides strong evidence that Watanabe’s/Richards’ analysis is not the correct licensing mechanism for *wh*-in-situ in Indonesian.⁴

In the next section, I consider whether the alternative LF movement analysis works for Indonesian and provide several arguments from Saddy that this analysis is also not tenable. I point out that those arguments also provide further evidence against extending the movement-based analysis proposed by Watanabe and Richards to *wh*-in-situ in Indonesian.

⁴ Of course, it is a separate issue whether Watanabe’s (1991, 2001) analysis still holds for Japanese, if not for Indonesian. The main argument for Watanabe’s analysis comes from the *wh*-island effect, illustrated in (i):

- (i) ?? John-wa [Mary-ga nani-o katta kadooka] Tom-ni tazuneta no?
 John-Top Mary-Nom what-Acc bought whether Tom-Dat asked Q
 ‘What did John ask Tom whether Mary bought?’ (Watanabe 2001: 208)

Provided that LF movement is immune to subadjacency (Baker 1970), the fact that (i) cannot be construed as the matrix *wh*-question indicates that movement of the embedded *wh*-phrase has taken place in overt syntax. Importantly, however, Toyoshima (2004) claims that the apparent *wh*-island effect is purely semantic in nature, hence does not necessarily support Watanabe’s analysis. Specifically, since the embedded question denotes a set of existentially closed propositions, the *wh*-interpretation of the cardinal predicate cannot be associated with the matrix set of propositions. See also note 11.

2.2. *Covert Syntactic Movement?*

Saddy provides four arguments that the covert syntactic movement analysis of *wh*-in-situ in languages as in Chinese and Japanese as proposed in Huang (1982) and Lasnik and Saito (1984, 1992) is also incorrect for *wh*-in-situ in Indonesian. The first argument comes from his observation, well-known among Indonesianists (see Cole and Hermon 1998, 2000 and Fortin 2007b, to mention a few), that the covert/LF movement in this language obeys the same set of island constraints as does the overt/syntactic movement, a position that has been also supported from work on other languages such as Japanese, Korean, and English as in Nishigauchi (1986, 1990), Choe (1987), Pesetsky (1987), and Reinhart (1991). This observation is based on the ungrammaticality of examples as in (4a-c) that involve partial *wh*-movement, that *wh*-phrases that remain within syntactic islands in overt syntax still give rise to ungrammaticality.

(4) Partial *Wh*-Movement in Indonesian

- a. *Kamu kira (bahwa)[cerita bahwa *siapa_i* yang *t_i* mengkritik Jonitu] di-jual?

you think that story that who that criticized Jon the PV-sold

‘Who do you think that the story that *t* criticized Jon was sold?’

b. * Kamu kira (bahwa) [cerita bahwa *siapa_i* yang Jon mengkritik *t_i* itu]di-jual.

you think that story that who that Jon criticized the PV-sold

‘Who do you think that the story that John criticized *t* was sold?’

c. * Kamu men-cemburui Bill [karena [_{PP} *dengan siapa_i*] yang saya berbicara *t_i*]?

you AV-get jealous of Bill because with who Foc I spoke

‘With whom did you get jealous of Bill because I spoke (to) *t*?’

(Saddy 1991: 195, 196)

In (4a), the *wh*-phrase *siapa* ‘who’ undergoes partial *wh*-movement into the intermediate, non-scopal specifier of CP. Since this short extraction itself does not cross any syntactic island, it cannot be the source of the ungrammaticality. Saddy argues that the ungrammaticality here directly follows if we assume that the LF/covert movement of the partially moved *wh*-phrase into the matrix specifier of CP obeys island constraints in Indonesian. According to this analysis, the *wh*-phrase *siapa* ‘who’ undergoes covert movement into the scopal specifier of CP for the purposes of scope taking. This movement, thus, renders the example in (4a) ungrammatical due to the Complex NP island. A similar story holds for (4b). In the example in (4c), the extraction of the prepositional *wh*-phrase *dengan siapa* ‘with who’ remains within the adjunct island but

the example is still ungrammatical. This is naturally predicted if the phrase undergoes further *wh*-movement into the matrix [Spec, CP], crossing the island at LF. Thus, examples as in (4a-c) provide us with independent reason to believe that covert movement (as well as overt syntactic movement, as we have seen in the previous subsection) obeys island constraints in Indonesian.

Now, if the covert movement analysis of *wh*-in-situ in languages such as Chinese, Japanese, and English (in multiple interrogative questions) as proposed by Huang and Lasnik and Saito is correct for Indonesian, we predict that the in-situ counterparts of the partially-moved examples in (4a-c) should also be ungrammatical because the LF representation of the in-situ variants would be identical to that of the examples in (4a-c) in that it involves violation of one of the island constraints. Crucially, however, Saddy shows that this prediction is falsified by the grammaticality of (5a-d), in-situ counterparts to (4a-d).

(5)a. Kamu kira (bahwa) [cerita bahwa *siapa* mengkritik Jon itu] di-jual?

you think that story that who criticized Jon the PV-sold

‘Who do you think that the story that *t* criticized Jon was sold?’

- b. Kamu kira (bahwa) [cerita bahwa Jon mengkritik siapa itu] di-jual?
 you think that story that Jon criticized who the PV-sold
 ‘Who do you think that the story that John criticized *t* was sold?’
- c. Kamu men-cemburui Bill [karena saya berbicara dengan siapa]?
 you AV-get jealous of Bill because I spoke with who
 ‘Who did you get jealous of Bill because I spoke with *t*?’

(Saddy 1991: 195, 196)

All these examples are impeccable as matrix *wh*-questions, contrary to what the LF movement analysis would predict. Thus, the grammaticality of (5a-d) provides strong evidence against the application of the LF movement hypothesis to *wh*-in-situ constructions in Indonesian. Note that the same examples are also problematic for the Watanabe/Richards-style overt movement analysis reviewed in the previous subsection because they would also incorrectly render (5a-d) ungrammatical. Thus, the contrast between (4a-c) and (5a-c) argues against a covert movement analysis of *wh*-in-situ in Indonesian, whether formulated as LF movement or as by Watanabe and Richards.

The second argument against the LF movement approach to *wh*-in-situ in Indonesian is based on the fact that this language does not allow complements that contain a *wh*-in-situ

for verbs such as *ingin tahu* ‘want to know, wonder’ that are obligatorily subcategorized for a +WH feature, as in English (e.g., *I wonder what you bought.* vs. **I wonder you bought what.*). This is illustrated by the contrast between (6a) and (6b).

(6) a. * Saya ingin tahu Jon men-cintai *siapa*.

I want know Jon AV-love who

‘I want to know who Jon loves.’

b. Saya ingin tahu *siapa* yang Jon Ø-cintai.

I want know who Foc Jon love

‘I want to know who Jon loves.’

(Saddy 1991: 207)

It is a fact of English grammar that the +WH complement property of interrogative verbs such as *wonder* in English can only be satisfied by the movement of a *wh*-word into the specifier of the CP complement of such verbs to check its inherent Q-feature against that of C. If the [+WH] subcategorization of the verb *ingin tahu* ‘wonder’ [+WH] must be satisfied by the [+WH] feature within its complement CP, then the fact that in-situ *wh*-elements do not satisfy the [+WH] requirement of this verb in (6a) suggests that they do not substitute into the specifier of CP, in contrast to overtly moved *wh*-phrases in (6b).

This contrast would remain mysterious under the LF covert movement analysis because the interrogative CP requirement would be satisfied by the covert movement of the in-situ phrase *siapa* ‘who’ into the specifier of the embedded CP. Note that this contrast is also problematic for Watanabe’s (1992, 2001)/Richards’ (2001) approach because the LF configurations for the moved and in-situ structures in (6a) and (6b) would be identical at LF. By contrast, the difference in grammaticality here naturally follows if we assume that the in-situ *wh*-phrase in (6a) literally remains in situ, meaning that no features of the phrase undergo any movement either in overt syntax or at LF; the subcategorization requirement is simply violated in (6a) because nothing moves to satisfy it, in contrast to (6b).⁵

⁵ Notice that the unselective binding approach predicts that (6a) should be grammatical, contrary to facts because this in-situ mechanism would allow the *wh*-phrase *siapa* ‘who’ to be interpreted in situ, rendering the embedded clause a well-formed question. Therefore, the contrast between (6a) and (6b) shows that unselective binding is inadequate to license an embedded question interpretation in Indonesian.

Carrying this conclusion over to Chinese, the embedded clause in (i) must involve covert/LF movement as proposed in Huang (1982), rather than unselective binding, because the latter operation does not work for creating embedded questions in Indonesian.

- (i) Zhangsan xiang-zhidao Lisi mai-le shenme.
 Zhangsan wonder Lisi bought what
 ‘Zhangsan wonders what Lisi bought.’

(Watanabe 2001: 204)

I thank Heidi Harley (personal communication) for pointing this out.

Thirdly, Saddy claims that *wh*-in-situ constructions in Indonesian do not show strong or weak crossover effects, as illustrated in (7a, b).⁶

(7)a. (*) Dia_i meng-harap Jon men-cintai *siapa_i?*

he AV-expect Jon AV-love who

‘Who_i does he_i expect Jon to love?’

b. (*) Prof dia_i meng-ira saya men-cintai *siapa_i?*

professor his AV-think I AV-love who

‘Who_i does his_i professor thinks I love?’ (Saddy 1991: 207, 208)

As we have seen briefly in chapter 2, the standard assumption on the crossover effect (Postal 1971; Wasow 1979) is that it arises when a variable-bound pronoun fails to be c-commanded both by a binder and by its variable at the surface/derived structure. Under this assumption, this effect can be formalized as the filter of the form * *wh_i pronoun_i... t_i* (see discussion below, though, for a clarificatory remark). The strong and weak crossover effects arise in examples in (8a) and (8b), respectively, because the pronoun coindexed with the binder is

⁶ The examples in (7a, b) are reported to be ungrammatical in Cheng (1991), Cole and Hermon (1998), Rogayah (1995). Thus, I have put the star in parenthesis here. See discussion below for more detailed discussion on these examples.

not c-commanded by the variable, namely, t . The ungrammatical examples in (9a, b) show that the relevant effect is also caused by quantifier raising (May 1985), a case of LF movement.

(8)a. * Who_i does he_i love t_i ? (strong crossover effect in overt syntax)

b. *? Who_i does his_i mother love t_i ? (weak crossover effect in overt syntax)

(9)a. * He_i loves everyone_i. (strong crossover effect at LF)

b. * His_i mother loves everyone_i. (weak crossover effect at LF)

Under the standard assumption, the (putative) lack of the weak/strong crossover effects in (7a, b) can be construed as evidence that the *wh*-phrase *siapa* ‘who’ remains in its thematic position both in overt syntax and at LF. If the syntactic movement in the sense of Watanabe (1992, 2001) and Richards (2001) occurred into the specifier of CP that c-commands the pronoun coindexed with the *wh*-operator, then the resulting configuration would cause the strong/weak crossover effect, as shown in (8a, b), contrary to facts. If the covert movement in the sense of Huang (1982) and Lasnik and Saito (1984, 1992) were correct, then the LF movement would cause the same violation as quantifier raising would, as shown in (9a, b). Thus, the (putative) absence of the

crossover effects in examples as in (7a, b) reported by Saddy cast further doubts on the validity of the syntactic movement as the mechanism of in-situ *wh*-construal in Indonesian.

Importantly, the preceding argument against the LF movement analysis based on the crossover effect crucially depends on the grammaticality of the examples in (7a, b) as reported by Saddy. It is quite debatable, however, whether this observation holds for Indonesian and Malay. For example, Cole and Hermon (1998) provide examples as in (10) to show that a crossover effect is observed in *wh*-in-situ in the dialect of Malay they document, contrary to what Saddy reports for Indonesian; see Cheng (1991) and Rogayah (1995) for the same observation.

(10) * Prof *dia*_i fikir saya meny-intai *siapa*_i?

Prof his think I AV-love who

‘Who_i does his professor think I love t_i?’ (Malay: Cole and Hermon 1998: 234)

My language consultants also concur with Cole and Hermon, reporting that the examples in (7a, b) are unacceptable when the pronominal *dia* is construed as a variable whose value co-varies with that of the *wh*-operator. It is not clear at this moment what causes this huge variation in the acceptability of the examples in (7a, b). However, provided the judgment cited by Saddy represents the minority one in the literature in Malay/Indonesian linguistics and in

the face of my own consultant work, I take it that the examples in (7a, b) are ungrammatical in (certain varieties of) Indonesian. This, in turn, might mean that we lose an argument against the covert (and null operator) movement approach to *wh*-in-situ in Indonesian.

Quite to the contrary, I maintain, following the suggestion made in Cole and Hermon (1998: 234), that the presence of the crossover effect does not automatically mean that the in-situ *wh*-phrase in (7a, b) undergoes syntactic movement because the crossover effect can be formulated in non-movement terms as a constraint on representations. Specifically, Cole and Hermon argue that the crossover effect can be analyzed as the byproduct of the Bijection Principle of Koopman and Sportiche (1983) that prohibits a single operator from binding more than one variable. This principle allows us to correctly block the examples in (7a, b) without also assuming syntactic movement because the base-generated *wh*-operator in [Spec, CP] binds both the pronoun and the variable (see detailed discussion of their unselective binding approach to *wh*-in-situ in Malay in section 4.4).⁷ For this reason, I conclude, contrary

⁷ As Heidi Harley (personal communication) points out, this representational analysis predicts that crossover effects should exist in *wh*-in-situ constructions in other in-situ languages such as Japanese. This prediction is correct, as shown in (i) in Japanese.

- (i) a. * Soitu_i-wa John-ga dare_i-o aisiteiru-to omotteiru no?
 that person-Top John-Nom who-Acc love-C think Q
 ‘Who_i does he_i think that John loves?’
- b. * Soitu_i-no-sensei-wa John-ga dare_i-o aisiteiru-to omotteiru no?
 that person-Gen-teacher-Top John-Nom who-Acc love-C think Q
 ‘Who_i does his_i professor think John loves?’

to Saddy, that the presence of the crossover effect is mute on the question of whether *wh*-in-situ in Indonesian undergoes syntactic movement or not.

Let us now come back to the last argument made by Saddy against the LF movement analysis of *wh*-in-situ in Indonesian. This argument is based on his observation that *wh*-in-situ constructions do not support a pair-list reading. Consider examples as in (11).

(11) Siapa mem-beli apa ?

Who AV-bought what

‘Who bought what?’ (Saddy 1991: 208)

Saddy reports that this multiple *wh*-question can only be interpreted as a request for a single pair as in *John bought a book*; thus, answers such as *John bought a book*, *Mary bought a magazine*, *Bob bought a shirt* are not a possible reply to this question. Let us suppose for the sake of argument that this observation holds for Indonesian multiple *wh*-questions. Since Higginbotham and May’s (1981) work on English multiple interrogatives, the availability of the pair-list reading for English sentences such as *who bought what* has been taken to be driven by the association of the two *wh*-phrases in the same Comp at LF (or the multiple specifiers of the same C, in the more modern terminology) via the semantic process known

as *absorption*. Higginbotham and May's central assumptions are summarized as in (12a-c), adopted from Barss' (2000) exposition (his (7-9), respectively) with a slight modification.

(12) Higginbotham and May's 1981 Absorption

- a. Absorption has a syntactic precondition of structural adjacency, which is defined as mutual *m*-command of the two operators.
- b. Absorption creates a complex *n*-ary quantifier, within which the original restrictions on the unary quantifiers are conjoined as a complex restriction and to which all the variables bound by the input quantifier are bound.
- c. Absorption is a sufficient, and necessary, precondition to the multiple-pair interpretation of multiple questions with singular *wh*-phrases.

(adopted from Barss 2000: 33, 34)

To the extent that this absorption analysis is correct, the lack of the pair-list reading in (11) indicates that *apa* 'what' does not undergo movement either in overt syntax or LF. Accordingly, this example provides evidence against the overt or covert movement approach to the *wh*-construal in Indonesian.

Again, however, my language consultants have all reported that the pair-list reading is readily available in sentences such as (11). This is also the judgment elicited from speakers of Malay by Cole and Hermon (1998: 225), who report that their Malay informants had no problem in giving a list interpretation for sentences as in (13).

(13) Siapa kamu fikir beli apa?

who you think buy what

‘Who did you think bought what?’ (Cole and Hermon 1998: 225)

This judgment indicates that the argument against the LF movement based on the pair-list reading is not strong as Saddy wanted it to be. We come back to this point again in section 4.

To sum up, I have reviewed four arguments presented in Saddy (1991) that the covert movement analysis as proposed in Huang (1982) and Lasnik and Saito (1984, 1992) for Chinese and Japanese is not an adequate mechanism of licensing *wh*-in-situ in Indonesian. Though the arguments based on the crossover effect and the availability of the pair-list reading in multiple questions do not necessarily argue for or against the LF movement analysis due differences in judgment between Saddy’s and my work, the two other arguments based on the

lack of island effects and the [+WH] subcategorization requirement provide relatively solid evidence that this analysis is not transportable to *wh*-in-situ in Indonesian.

2.3. *Unselective Binding?*

The third potential analysis of *wh*-in-situ, which is perhaps the most widely held analysis for *wh*-in-situ in languages such as Japanese and Chinese is that of unselective binding (Baker 1970; Pesetsky 1987); I defer discussion of the variant of this approach presented recently by Cole and Hermon 1998 until section 4.4). Drawing on the analysis of multiple *wh*-questions and their interpretations in English by Baker (1970), Pesetsky (1987) proposes that *wh*-interpretation may be achieved not only by syntactic movement but also a non-movement mechanism called *unselective binding*. Pesetsky claims that the choice between these two options is determined by the notion of *D(iscourse)-Linking*, which roughly corresponds to the morphological distinction in English of *wh*-words between “*which-X*” (*which man*, *which book*, etc) and everything else (*who*, *what*, etc). (recall our discussion of Cinque’s (1990) notion of referentiality in chapter 2). As Pesetsky (p. 107-108) remarks, *which*-phrases are *discoursed-linked* (*D-linked*), because “when a speaker asks a question like *which book did you read?*, the range of felicitous answers is limited by a set of books both speaker and hearer have in mind” whereas “no such requirement is

imposed on *wh*-phrases like *who*, *what*, or *how many books*.” Pesetsky argues that if a *wh*-phrase is D-linked, it contains a variable that is unselectively bound by a Q-morpheme located in the scopal C head position, and thereby is licensed without syntactic movement. On the other hand, if a *wh*-phrase is not D-linked, it must undergo syntactic movement, be it overt or covert, to be properly licensed by the scopal C. Pesetsky provides various types of evidence based on the presence/absence of superiority effects in English questions as well as the behavior of what he calls aggressively non-D-linked *wh*-phrases such as *what the hell* in English and Japanese to support this hybrid approach to *wh*-construal.

Saddy, however, points out a couple of potential empirical problems with Pesetsky’s version of unselective binding analysis when applied to *wh*-in-situ in Indonesian. The first problem concerns the morphological composition of *wh*-phrases in Indonesian. As we have seen above, Pesetsky’s analysis rests upon the correlation between the morphological composition of a *wh*-phrase and its interpretive mechanism: if a *wh*-phrase is D-linked, it is interpreted by unselective binding without movement whereas, if it is non-D-linked, it must be interpreted by syntactic movement. This correlation, however, does not hold in Indonesian because almost all *wh*-phrases in this language have “D-linked” expressions corresponding to English “*which-X*” form. For example, Saddy observes that *orang siapa* ‘which person’, which would be analyzed as a D-linked phrase in Pesetsky’s terms, is used interchangeably

with the non-D-linked form *siapa* ‘who’ but this difference in morphological composition does not change the interpretive and structural constraints observed so far in this section. Though this observation may not be a serious problem for Pesetsky’s theory, it at least indicates that Pesetsky-style D-linking is not directly applicable to Indonesian *wh*-questions.

The second potential problem with the extension of Pesetsky’s analysis to Indonesian in-situ questions is based on what Saddy reports as the quantificational un informativeness of *wh*-in-situ in Indonesian. Pesetsky employs D-linking to account for the triplet interpretation available for sentences as in (14), so that the D-linked phrase *which prize* may get matrix scope without movement by being bound by the matrix Q in the manner seen in (15).

(14) *Who* did every athlete expect to win *which prize*?

triplet answer: Gretsky expected Milli Vanilli to win an Oscar, Gefrion expected

George Burns to win Grammy, etc. (Saddy 1991: 204)

(15) $[_{S'} [_{Comp} Q_{i,j} who_i [_{S} e_i \text{ every athlete expect.... win which prize}_j]]$

Importantly, this analysis crucially assumes that D-linked in-situ phrases such as *which prize* must be able to interact in scope with other scope-bearing elements such as *every*

athlete; for otherwise, the triplet interpretation would be unavailable in examples such as (14). As Saddy (p. 205) put it, “it is a necessary property of Pesetsky’s Q-bound D-linked WH expressions that they interact quantificationally with other elements in the main clause” When applied to Indonesian *wh*-in-situ constructions akin to (14) above, Pesetsky’s analysis leads us to predict that this type of construction also should allow the triplet interpretation. Saddy observes that this prediction is false in Indonesian because this reading is precisely the kind of interpretation that BI *wh*-in-situ resists, as in (16a).

- (16) a. Setiap orang men-cintai *siapa*?
 every person AV-love who
 ‘Who did every person love?’ who>every, *every>who
- b. *Siapa*_i yang setiap orang Ø-cintai *t*_i?
 who Foc every person love
 ‘Who did every person love?’ who>every, every>who (Saddy 1991: 199)

According to Saddy, the example in (16a) with the *wh*-phrase in situ only allows the wide scope reading of the in-situ phrase with respect to the universal quantifier *setipa orang* ‘every person’ in subject position; the reading where the value of the person loved co-varies

with that of the lover is impossible. This latter reading becomes available only when the *wh*-phrase must undergo overt syntactic movement, as illustrated in (16b).

Saddy notes that the same contrast can be seen in multi-clausal environments, as the possible scope interpretation in the triplet of examples in (17a-c) illustrates.

(17) a. *Setiap orang tahu Tom mem-beli apa?*

every person knows Tom AV-buy what

‘What does every person know Tom bought?’ what>every, *every>what

b. *Setiap orang tahu apa yang Tom beli t_i?*

every person know what Foc Tom buy

‘What does every person know Tom bought?’ what>every, every>what

c. *Apa yang setiap orang tahu Tom beli t_i?*

what Foc every person know Tom buy

‘What does every person know Tom bought?’ what>every, every>what

(judgments as reported by Saddy 1991: 200)

In (17a), the in-situ *wh*-phrase *apa* ‘what’ necessarily takes wide scope over the universal quantifier *setiap orang* ‘every person’, even though the relative structural height of the

latter with the latter leads us to expect the opposite reading. Again, the wide scope reading of the universal quantifier over the *wh*-phrase is only possible when the latter undergoes syntactic movement, either partial movement, as shown in (17b), or full movement, as shown in (17c). The pattern of (un-)available scope readings seen in these examples, therefore, shows that *wh*-in-situ in Indonesian is quantificationally uninformative with respect to other scope-bearing expressions, unlike moved *wh*-phrases. Accordingly, this result would remain mysterious under Pesetsky's Q-binding analysis of triplet questions. Note that this contrast in scope interaction between moved *wh*-phrases and *wh*-in-situ also shows that the scope of the latter cannot be determined by overt or covert syntactic movement; if this were the case, there should be no scopal difference between these two expressions. Based on this conclusion, Saddy concludes that Pesetsky's analysis is incorrect.

An important remark I need to add here, however, is that I could not reproduce the same judgments as elicited by Saddy from his Indonesian consultants. According to my language consultants, both (16a) and (17a) allow the narrow scope reading of the in-situ *wh*-phrase with respect to the universal quantifier, the reading where the value of the thing bought and the person loved can vary with the value of the universal quantifier. To check this judgment, I asked my language consultants to provide scope judgments for the sentence in (15a) in the following manner: "Suppose that there are only five persons in the world in 2100 due to

critical food shortage: John, Bob, Susan, Mary, and Amy. Then, observing this world, I said the sentence in (15a). In this case, are the following scenarios compatible with the meaning of the sentence in (15a)?” Then, I gave two scenarios given below that diagnose narrow scope and wide scope reading of the in-situ *wh*-phrase *apa* ‘what’.

(18) Scenario A (narrow scope reading of in-situ *wh*-phrase)

: John loves Susan, Susan loves Bob, Bob loves Amy, Amy loves Bob, and Mary loves John.

Scenario B (wide scope reading of in-situ *wh*-phrase)

: John, Bob, Susan, Mary all love Amy (and, Amy loves herself).

All the three consultants reported to me that the sentence in (15a) can be used to describe both situations in (18a, b). A similar judgment task was conducted for the in-situ question in (16a), for which the same consultants reported the same ambiguity. This result, therefore, shows that the *wh*-in-situ in Indonesian *is* scopally informative, quite contrary to what Saddy reports. At this moment, I have no idea how scope judgments can vary in such a clear manner, as I do not know the linguistic backgrounds of Saddy’s language consultants. For the purposes of this chapter, I assume that my consultants represent the majority judgment,

keeping in mind, though, that Saddy's reported judgments might also hold for certain dialects of Indonesian. We return to this in section 4.

Now if this is so, we have lost one argument against Pesetsky's version of the unselective binding approach to *wh*-in-situ in Indonesian. However, in section 4.1, I review Reinhart's (1992, 1995, 1997, 1998, 2006) arguments based on the scope behavior of *wh*-in-situ in multiple questions in English and the "Donald Duck" Problem that cast serious doubts on the general applicability of unselective binding as a possible non-movement strategy to license in-situ expressions. Anticipating this forthcoming discussion, I assert here that Pesetsky's D-linking-based unselective binding analysis is not suitable for *wh*-in-situ in Indonesian.

2.4. *Section Summary*

Let me summarize what we have done in this section. I have reviewed here Saddy's (1991) arguments that none of the existing movement and non-movement-based accounts of *wh*-questions motivated in other in-situ languages can adequately account for a series of somewhat peculiar syntactic and semantic properties associated with *wh*-in-situ in Indonesian. The lack of island effects in Indonesian *wh*-in-situ in contrast to overtly moved *wh*-phrases and facts concerning verbal subcategorization show that neither "overt" movement analyses as in Watanabe (1992, 2001) and Richards (2001) nor covert movement

analyses as in Huang (1982) and Lasnik and Saito (1984, 1992) are viable ways of interpreting this type of interrogative phrase in Indonesian. The putative lack of strong/weak crossover effects, and the potential absence of the pair-list reading suggest that *wh*-in-situ and its associated features in Indonesian literally remain in their base-generated thematic position. The language-particular morphological composition of *wh*-phrases and the (alleged) scopally uninformative behavior of this class of expression in Indonesian indicates that Pesetsky's (1987) version of the non-movement approach to in-situ questions in terms of unselective binding is also problematic. I have also made it clear that some of the judgments reported by Saddy himself are contradictory to the judgments elicited in the present study and those reported in the literature on Indonesian/Malay *wh*-questions as in Cheng (1991), Rogayah (1995), and Cole and Hermon (1998).

Now that three widely held analyses of *wh*-construal have been eliminated as analytical options for *wh*-in-situ in Indonesian, we are left once again with the original question with which I started the present chapter: what is the correct licensing mechanism for *wh*-in-situ in Indonesian that captures all the apparently mysterious syntactic and semantic properties such as those discovered by Saddy (1991). The rest of this chapter is devoted to showing that the answer to this question can only be found in the syntax-external interpretive component and to arguing that the particular semantic analysis proposed below is an

optimal realization of the needs of syntax on the part of the component in the way that our minimalist interface thesis leads us to expect. Before presenting such an analysis, however, let us consider first in the next section what Saddy (1991) himself said about the now mysterious *wh*-in-situ construction in Indonesian.

3. *Wh*-in-Situ in Indonesian is Not an Interrogative Definite Description

In this section, I review Saddy's novel analysis of *wh*-in-situ in Indonesian as an interrogative definite description that draws on the Jeopardy game show question in English. Although I provide several arguments that such a treatment of the relevant class of *wh*-expressions is inadequate, I also stress that his particular analysis contains a number of important insights that the analysis to be developed in section 4 attempts to capture.

3.1. Saddy's (1991) Analysis of Wh-in-Situ as Interrogative Definite Description

Saddy proposes that *wh*-in-situ in Indonesian behaves as interrogative definite descriptions, drawing on an impressive range of syntactic and semantic parallelisms that hold between this class of expressions and words of the form "this-X/these-Xs" in game Jeopardy show questions in English. An example of English game show quizzes is given in (19).

(19) Question: For \$100, every armchair general watched this television station.

Answer: What is NBC?

(slightly modified from Saddy 1991: 208)

The “question” in (19) does not have interrogative force in the standard sense as a *wh*-question because it is syntactically a declarative statement; rather, it gains such force from the very context that this sentence is uttered in a game show; a host utters this sentence to competitors, expecting them to make a question such that it constitutes an appropriate answer to the definite DP *this television station*. In other words, the interrogative requirement here is that competitors come up with the member (s) of the definite description of the form ‘this-X/these-Xs.’ The reason Saddy brought this type of game show scenario into his work is because of his observation that statements as in (19) under the game show context exhibit exactly the same range of structural and interpretive properties that we have seen to characterize *wh*-in-situ in Indonesian. Those properties are illustrated in (20-23), with some examples.

(20) “This-X” Form Cause No Island Effects.

- a. For \$100, the viewing public wondered whether Reagan believed the claim that this man was a hero.
- b. For \$100, the viewing public accepted the circumstances tat this man cited as extenuating.
- c. For \$100, many people refused to take showers after seeing this movie.

(Saddy 1991: 210)

(21) “This-X” Form Cannot Satisfy Subcategorization Requirements.

- * Bill wonders (for \$100), every armchair general watched this television station.

(Saddy 1991: 210)

(22) “This-X” Form Cause No Weak or Strong Crossover Violations.

- a. For \$100, his_i mother likes this man_i.
- b. Oliver North has been hailed as a great American patriot. ? For \$100, he_i most admires this man_i.

(Saddy 1991: 211)

(23) “This-X” Form Blocks Pair-List Reading in Multiple *Wh*-Questions

Question: For \$100, this man married this woman

Response:* John Smith married Sally Jones, Tod Wilks married Martha Sachs...

(Saddy 1991: 211)

Based on these facts, Saddy concludes that the same range of properties of Indonesian in-situ *wh*-phrases naturally fall out if they are interrogative definite descriptions. This analysis of *wh*-in-situ in Indonesian also makes the correct prediction that it should be scopally uninformative because it is widely known that bona fide definite descriptions are not, as shown in (24), where the definite DP *the woman* must take wide scope over the quantified subject *every person*.

(24) The Quantificational Uninformativeness of Definite Descriptions in English

Every person likes the woman with red hair.

⇒ *every> the woman with red hair, the woman with red hair>every (Saddy 1991: 212)

Saddy argues that the definite description analysis of *wh*-in-situ in Indonesian, informed by the behavior of “this-X” form in game shows, provides a unified account of all the properties we have seen to hold for this class of *wh*-questions. As we have just seen, the scopally

uninteractive behavior of in-situ *wh*-phrases is a natural consequence of the fact that it is a definite description, namely, that they “pick out a specific individual or a set of individuals.” (p. 212). *Wh*-in-situ does not satisfy the WH-complement requirement of verbs such as *ingin tahu* ‘wonder’ because it does not move into the specifier of the complement CP but instead is licensed in situ by a non-quantificational mechanism by virtue of its definite nature. Similarly, the (putative) lack of pair-list readings and weak/strong crossover effects and the insensitivity to syntactic islands for the purposes of scope taking are derived because *wh*-in-situ is interpreted in situ as an interrogative definite description.

Saddy’s analysis is extremely ingenious in a number of important ways. First, it provides one unified, non-stipulatory account of all the otherwise mysterious syntactic and semantic characteristics associated with *wh*-in-situ in Indonesian from the single fact that this class of *wh*-words is an interrogative definite description. Second, as hinted at the end of his paper (p. 213), his analysis leads to the conclusion that, once a *wh*-element undergoes syntactic movement, that element becomes quantificationally, as evidenced by the sudden emergence of new scope interaction between it and other scope-bearing elements such as universal quantifiers. Though he does not have an answer to this question, he speculates (p. 212) that “part of the mechanism of quantificational WH construal is derivative from the simple fact of movement having taken place.” Finally, his analysis suggests that natural languages may

well develop a non-syntactic mechanism of licensing *wh*-in-situ without relying on syntactic movement. The last two points become crucial in the rest of this chapter, in which I develop a different implementation of Saddy's insight in a more sophisticated manner that is driven by the considerations of the thesis of minimalist interface.

Though Saddy's analysis has many things that inform us about the proper analysis of *wh*-in-situ in Indonesian as well as other in-situ languages, there are, nonetheless, several arguments that suggest that his most fundamental claim that *wh*-in-situ in Indonesian is a definite description is incorrect. I review those arguments in the next subsection.

3.2. *Problems with Saddy's (1991) Analysis*

There are three problems that cast doubts on the validity of Saddy's treatment of *wh*-in-situ in Indonesian as an interrogative definite description. The first potential problem analysis is that the source of interrogative force is quite different between English game shows and *wh*-in-situ questions in Indonesian.⁸ In game show questions, the moderator already knows the value(s) of the definite description provided by "this-X/these-Xs". Speakers of Indonesian, by contrast, utter *wh*-in-situ questions precisely because they would like to know what is the correct value for the *wh*-phrase. In normal discourse situations, they always utter these sentences to seek new information. It has been generally assumed (see Chomsky 1977 and Zubizarreta 1998 for

⁸ Thanks to Heidi Harley (personal communication) for discussion on this point.

extensive discussion) that *wh*-questions request new information. Thus, in English sentences such as *Who bought the book yesterday*, the *wh*-phrase *who* requests new information while the rest of the clause [*x bought the book yesterday*] is old information. This means that, semantically, the *wh*-phrase is not a definite description. This qualitative difference between “this-X/these-Xs” form in game shows and the *wh*-phrase in Indonesian, thus, poses a potential problem for Saddy’s attempt to unify the two phenomena.

The second argument against Saddy’s analysis is that there is evidence internal to Indonesian that *wh*-in-situ in this language contains a variable. Cole and Hermon (1998) observe that nominal *wh*-words in Malay can be used as a variable bound by non-*wh*-operators, as in (25a, b) and (26a, b). This observation also holds for Indonesian, according to my consultants.

(25) a. Dia tidak mem-beli ***apa-apa*** untuk saya.

he not AV-buy what-Red for me

‘He did not buy anything for me.’

b. Dia tidak mem-beli ***apa-pun*** untuk saya.

he not AV-buy what-also for me

‘He did not buy anything for me.’

(Malay: Cole and Hermon 1998: 239)

(26) a. Saya tidak kenal *siapa-siapa* di universitas itu.

I not recognize who-Red at university that
'I didn't recognize anyone at that university.'

b. Saya tidak kenal *siapa-pun* di universitas itu.

I not recognize who-who at university that
'I didn't recognize anyone at that university.' (Malay: Cole and Hermon 1998: 239)

In (25a) and (26a), the *wh*-word is bound by the existential quantifier that is overly represented by the reduplication of the question word itself (see Travis 1999, 2001, 2003 for the observation that reduplication also has the parallel function of creating existential quantification in Malagasy). Similarly, in the examples in (25b) and (26b), the *wh*-word is bound by the existential quantifier realized in the form of *-pun* 'also'. This use of the in-situ *wh*-words in Indonesian as an element bound by quantifiers, therefore, shows that this class of words contains a variable. This result is problematic for Saddy's treatment of the same words as interrogative definite descriptions because definite descriptions as rigid designators do not contain a variable under the most commonly held assumption. At the same time, this result indicates that *wh*-in-situ in Indonesian, like its equivalent in Chinese, behave more like normal *wh*-phrases whose movement leaves a variable in the launching site.

The final argument against Saddy's analysis is that it misses the important generalization that *wh*-in-situ in Indonesian behaves as existential indefinites rather than definite descriptions in terms of scope taking. It is widely acknowledged in the literature that certain weak/existential indefinites such as singular NPs (e.g., *someone*, *something*) and cardinal plurals (e.g., *two men*, *many women*) are insensitive to syntactic islands for scope-taking, as shown by the contrast between (27a-c) and (28a-c).

- (27) a. Someone reported that Max and *all the ladies* disappeared.

⇒ some>all, * all>some

- b. Someone will be offended if we don't invite *most philosophers*.

⇒ some>most, *most>some

- c. Many students believe anything that *every teacher* says.

⇒ many>every, * every>many

(Reinhart 1997: 338)

(28)a. Everyone reported that Max and *some lady* disappeared.

⇒ every>some, some>every

b. Most guests will be offended if we don't invite *some philosopher*.

⇒ most>some, some>most

c. All students believe anything that *many teachers* says.

⇒ all>many, many>all

(Reinhart 1997: 339)

(27a-c) show that the strong quantifiers such as *all*, *most*, and *every* cannot violate one or the other island constraints to take wide scope over another scope-bearing element in the matrix clause. This is not surprising if Quantifier Raising, an instance of LF movement, is constrained by the island constraints, as is overt syntactic movement. What is surprising, then, is the fact illustrated in (28a-c), that weak existential indefinites such as *some* and *many* take wide scope over the quantifier in the matrix subject position, in apparent violation of the island constraints that we have just seen to constraint the Quantifier Raising operation. This wide scope reading of certain existential indefinites has been a source of endless controversies in the semantic literature; see Reinhart (2006: ch2.) for a concise summary of these controversies. Whatever the ultimate analysis might turn out to be, this island-insensitive scope-taking behavior is similar to that of *wh*-in-situ in Indonesian. We have seen in section 2 that this class of phrases

can freely take wide(st) scope in a massive violation of the set of standard island constraints on movement. Given this parallelism, the null hypothesis seems to be that *wh*-in-situ in Indonesian should be treated as existential indefinites. This indeed has been a standard assumption on *wh*-phrases since the earliest days of generative research as argued for in Chomsky (1964); Kats and Postal (1964); Klima (1964), and Kuroda (1965); see also Karttunen (1977), and Engdahl (1986).

Based on the afore-mentioned three considerations, I reject Saddy's (1991) approach to *wh*-in-situ in Indonesian and seek an alternative account that nonetheless captures the three fundamental insights behind his work mentioned in the end of section 3.1.⁹

4. Choice Function as an Optimal Interface Strategy at the Syntax-Semantics Interface

We have seen thus far that *wh*-in-situ in Indonesian remains in situ throughout the syntactic derivation. Drawing on the evidence presented in Saddy (1991), I have shown

⁹ Of course, what is the proper analysis of the game show question with the properties as observed by Saddy is a separate question that I leave aside in this chapter. I note here, however, that the choice function analysis of the kind developed Kratzer (1998) and Matthewson (1999) (see section 4.2 for discussion of their analysis) has suitable theoretical properties to accommodate this type of question. Kratzer argues that indefinites in English are divided into specific and quantificational and that they must take widest scope when interpreted as specific in the form of choice function. Matthewson provides evidence for Kratzer's analysis from evidence in St'át'imcets. Since Saddy's core claim is that all the peculiar properties of the 'this-X/these-Xs' in the game show question are derivable from their denotation as a definite description, it may well be that Kratzer/Matthewson-style analysis provides a unified account of the observed properties. See Sato (in prep) for such an analysis.

that two analyses that have been most often adopted for *wh*-in-situ in other languages such as English (for multiple *wh*-questions), Japanese, and Chinese in terms of syntactic movement and unselective binding have several serious shortcomings in face of several structural and semantic properties associated with this construction. I have also provided evidence based on the reduplication and the island-insensitive behavior of *wh*-words for the purposes of scope-taking that Saddy's analysis of this class of expressions as an interrogative definite description cannot be upheld. Thus, the fundamental question that we still need to address is what mechanism is available to license *wh*-in-situ in Indonesian without moving it while, at the same time, capturing the observed properties associated with this construction.

Now, let us think what we can do about this state of affair under the thesis of minimalist interface. Given that the syntax-external semantic component is fundamentally interpretive, utilizing whatever domain-specific operations available within this component to assign a semantic interpretation to the output of syntactic derivation, the current thesis expects that the fact that *wh*-in-situ remains in situ in Indonesian is simply the result of the syntactic computation: under the Last Resort view of movement, for example, this class of expressions remains in situ since there is no morphological feature that legitimizes such movement. The minimalist interface thesis, therefore, entails that the syntactic and

semantic properties associated with *wh*-in-situ in Indonesian should be solely accounted for in terms of the syntax-external semantic interface.

Crucially, an important series of work by Reinhart (1992, 1997, 1998, 2006) appeared soon after the publication of Saddy (1991), which claims that there is an independent interpretive mechanism available for licensing indefinite expressions, including *wh*-phrases, in terms of allowing existential quantification over choice functions in the sense of Heim (1982).¹⁰ Reinhart argues that introducing this way of licensing allows for a unified explanation for the set of syntactic and semantic problems that would arise with analyses in terms of LF movement, unselective binding, and absorption. Let us, therefore, first review her work on choice function.

4.1. *Choice Functions*

Reinhart (1992, 1995, 1997, 1998, 2006) starts by showing that neither the LF movement (Huang 1982) nor unselective binding (Pesetsky 1987)/absorption (Higginbotham and May 1981) analyses of *wh*-in-situ in multiple questions in English are tenable on the grounds that they cannot derive several syntactic and interpretive properties associated with this type of expressions. Consider first examples as in (29a-d).

¹⁰ I am grateful to Heidi Harley (personal communication) for suggesting this idea to me.

- (29) a. Who fainted when you attacked whom?
- b. * Who fainted when you behaved how?
- c. * How did Max faint when you behaved?
- d. Who fainted when you behaved what way? (Reinhart 1998: 31, 44)

It has been standardly assumed since the seminal work by Huang (1982) that overt syntactic movement obeys both subadjacency and the ECP whereas covert LF movement is only constrained by the ECP. This line of analysis correctly predicts the asymmetry between (29a) and (29b) once we assume that in-situ *wh*-phrases undergo covert movement into the matrix C. The example in (29a) is grammatical because the trace of *whom* is head-governed by the verb *attacked*. The example in (29b) is ungrammatical because the trace of *how* does not satisfy the antecedent-government requirement of the ECP, on a par with the example in (29c). Reinhart notes, however, that this line of analysis inspired would, then, have no way of accounting for why there is a contrast between (29b) and (29d); it would incorrectly predict the latter to be ungrammatical because *what way* is just as much an adjunct as *how*. This contrast, therefore, shows that what matters is not LF movement but instead the difference in syntactic category (NP vs. non-NP) for the purposes of licensing *wh*-in-situ expressions in English (see also section 4.3 for relevant discussion on Indonesian examples akin to (29d)).

Another problem with this LF movement analysis is that this way of assigning matrix scope by LF movement is inconsistent within the Minimalist Program (Chomsky 1995). One of the central ideas of this framework is that movement is subject to the “shortest steps” requirement (recall also chapter 2). To illustrate, consider examples as in (30a), which is assigned the LF representation in (30b) under the matrix scope reading of the in-situ *wh*-phrase *what* (in which case, the question can be answered by sentences like *Max knows where to find bicycles*).

(30) a. Who knows where to find what.

b. for which $\langle x, y \rangle$, x knows where to find y (Reinhart 1998: 33)

The LF movement analysis would claim that the relevant in-situ phrase would take the matrix scope by undergoing LF movement into the specifier of the matrix, scopal C. Reinhart notes that this very movement is impossible within the minimalist framework because it is less economical in terms of the shortest movement requirement than its potential movement into the specifier of the embedded CP. Thus, the example here shows that the scope assignment of *wh*-in-situ in LF movement is untenable and that a non-movement licensing is needed.

Reinhart further shows that the non-movement approach to *wh*-in-situ in terms of unselective binding/absorption also fails in light of the interpretation of examples as in (31).

(31) Who will be offended if we invite which philosopher? (Reinhart 1998: 36)

Reinhart assumes the semantics of questions proposed by Karttunen (1977) and Engdahl (1986), namely, that the denotation of a question is the set of propositions which constitute true answers to it. The unselective binding/absorption mechanism would assign the interpretation in (32a), which would be more formally represented as in (32b) under Karttunen's model, for the sentence in (31).

Wrong:

(32)a. for which $\langle x, y \rangle$, if we invite y and y is a philosopher, then x will be offended.

b. $\{P(\exists \langle x, y \rangle)\} \ \& \ P = ((\text{we invite } y \text{ and } y \text{ is a philosopher}) \rightarrow (x \text{ will be offended}) \ \& \ \text{true}(P))$

c. Lucie will be offended if we invite Donald Duck.

(Reinhart 1998: 36)

It is important to note that, in the representation in (32a), the restriction is contained in the implication of an *if*-clause. Given the truth-theoretic conditions on such a clause, the sentence in (31) would come out true in cases where the value of *y* is a member of the non-philosopher set; for example, the sentence would be true if Donald Duck is inserted since he is not a philosopher, as in (32c). This is because the only case where a conditional is a false statement is when a true antecedent *if*-clause leads to a false consequence clause; if Donald Duck were the value of *y*, then the antecedent clause would be false, hence the truth value of the sentence in (31) as a whole would be true. Clearly, this is not what the English sentence means. What we need to ensure, then, is thus to pull out the restriction from the implication as in (33a) or its Karttunen-style equivalent in (33b).

Right:

(33) a. for which $\langle x, y \rangle$, *y* is a philosopher, and if we invite *y*, *x* will be offended.

b. $\{P(\exists \langle x, y \rangle)(y \text{ is a philosopher}) \ \& \ P = \wedge ((\text{we write } y) \rightarrow (x \text{ will be offended})) \ \& \ \text{true}(P)\}$

(Reinhart 1998: 36)

Thus, examples as in (31) show that the absorption/unselective binding is not adequate for assigning scope for *wh*-in-situ phrases in English *wh*-questions. At the same time, we can see

from these examples that the ultimate semantic mechanism should be able to ensure that the value of the *wh*-in-situ will be necessarily chosen from the set of members that satisfy its accompanying restriction.

Reinhart argues that all the problems noted above with LF movement or unselective binding/absorption are straightforwardly resolved once we allow existential quantification over choice functions in the form of existential closure in the sense of Heim (1982); see also Winter (1997) for a quite similar approach. Choice function is defined as in (34).

- (34) A function f is a choice function (CH (f)) if it applies to any non-empty set and yields a member of that set. (Reinhart 1997: 372)

Let me illustrate how this choice function mechanism works to solve the problems noted above. For example, the LF representation of examples as in (35a) under the wide scope reading of the indefinite is shown in (35b) under the choice function approach.

- (35) a. Every lady read some book.
b. $\exists f (\text{CH}(f) \ \& \ (\forall z) (\text{lady}(z) \rightarrow z \text{ read } f(\text{book})))$

In (35b), the indefinite *book* is replaced by a function variable to be bound by an existential operator that is base-generated in the highest level. The choice function here, thus, applies to the non-empty set of books and picks up one member out of this set. This representation says that there is a function f such that for every z , if z is a lady, z reads the book selected by this function. More informally, then, this representation states that there is a book that is read by every lady, i.e. the wide scope reading of the indefinite expression.

This analysis provides a straightforward solution to the problems we have noted above with LF movement or unselective binding/absorption approaches. First, the contrast between (29a, d) and (29b) naturally follows under the standard assumption since Szabolcsi and Zwarts (1993) that adverbial *wh*-phrases do not have an N-set and that they denote functions ranging over higher-order entities. The examples in (29a, d) are both grammatical because *whom* and *what way* contain an N-set (see also Higginbotham (1985) for his notion of argument index and discharge), which is a necessary condition for choice function to work. On the other hand, the example in (29b) is ungrammatical because *how* cannot be evaluated by choice function due to its lack of N-set; the movement that is required would be blocked by the shortest movement requirement.¹¹ The same conclusion is reached in Tsai (1999). He

¹¹ This is not what Reinhart actually argues for. Specifically, Reinhart (1997, 2006) argues, drawing on her earlier work in Reinhart (1981), that adverbial *wh*-phrase, base-generated in [Spec, QP], can only be licensed in [Spec, CP] and that a single sentence can host only one *wh*-phrase in such a position. Since the sole purpose here is to show that the contrast is a direct consequence of choice function, I have stated here that the

argues that the adjunct vs. argument asymmetry here is actually the nominal *wh* vs. adverbial *wh* asymmetry on the ground that adverbials do not contain a variable, hence must undergo covert syntactic movement that would result in island/ECP violations, acknowledging (p. 183) that this possibility has been actually entertained in Huang (1982).

The “Donald Duck” problem, which we have seen to arise with unselective binding/absorption above, is also directly solved as the consequence of the choice function because, as defined in (34), a choice function applies to a non-empty set of individuals and yields a member *out of this set*. Accordingly, the value for *y* in examples such as (31) is correctly chosen from the set of philosophers.¹²

We should add two other remarks on Reinhart’s analysis at this point. First, the choice function analysis closely mirrors the intuition expressed by Saddy (1991) that natural languages may well develop a non-syntactic mechanism of licensing *wh*-in-situ in its base position without relying on syntactic movement. Second, this licensing strategy seems to be a natural analysis developed by the syntax-external semantic component to reflect the needs of the conceptual-intentional system, given that the *wh*-in-situ literally remains in situ in

ungrammaticality of (29b) is due to the shortest movement as one possibility. This move is conceivable given her conception of the shortest movement requirement made earlier in the text. At any rate, this decision does not affect our discussion in this chapter.

¹² As mentioned in note 4, Toyoshima (2004) applies the choice function analysis to Japanese *wh*-questions. See also Lin (2004), who extends the choice function analysis to *wh*-questions in Mandarin Chinese.

Indonesian as far as syntax goes. Therefore, it stands to reason that the characteristic properties of *wh*-in-situ are derived as the result of the semantics developing its own interpretive strategy. This line of thinking is exactly what we expect under the thesis of minimalist interfaces, according to which the semantic interface does whatever it can to assign an interpretation to the output of syntax. Recall that the results of my earlier chapters have suggested that semantics assign interpretation to the output of syntax in a way that is required by the syntactic information contained within the output. Similarly, in the present case, semantics not only accepts the needs of syntax (in the sense that it must interpret *wh*-in-situ in its base position) but also must be rich enough to yield a number of new interpretations to satisfy the needs of the conceptual-intentional system that would be unavailable by the purely syntax-driven mechanisms of *wh*-scope assignment as in LF movement, unselective binding, and absorption. I believe that this state of affairs is quite natural in light of the fact that the semantics is the interface between the language faculty and the conceptual-intentional system in the sense of Chomsky (1995, 2000, 2001, 2004, 2005). It is undeniable fact that it serves as the information access point between language and the general concept system (belief, context, intention, etc.). I will return to this point in section 5.

Of course, it is a different question whether this conceptual expectation is satisfied or not in the particular case of *wh*-in-situ in Indonesian. In the rest of this chapter, I demonstrate

that this expectation is fulfilled by showing that a) Reinhart's version of the choice function approach straightforwardly derives all the properties associated with *wh*-in-situ and that b) it allows us to make further empirical predictions that are indeed correct with regards to the NP vs. non-NP asymmetry with regards to the putative ECP effects.

4.2. *Deriving the Properties of Wh-in-Situ in Indonesian*

Let us now see whether all the syntactic and semantic properties we discussed so far with respect to *wh*-in-situ in Indonesian follow under the choice function approach by Reinhart (1992, 1995, 1997, 1998, 2006). The relevant properties are summarized in (36a-e). Recall that the status of the properties given in (36c-e) is unclear, as shown in section 2.

(36) Syntactic and Semantic Properties of *Wh*-in-Situ in Indonesian

- a. *Wh*-in-situ is insensitive to syntactic island/ECP effects.
- b. *Wh*-in-situ is not able to satisfy verbs' +*wh* subcategorization requirement.
- c. *Wh*-in-situ is immune to weak/strong crossover effects.
- d. *Wh*-in-situ is scopally uninteractive, always taking non-overt wide scope.
- e. *Wh*-in-situ is not able to support a pair-list reading for multiple *wh*-questions.

The property in (36a) is directly derived from the simple fact that *wh*-in-situ in Indonesian remains in situ throughout the syntactic derivation (including on the mapping to LF). The property in (36b) is derived for the same reason; this class of expression cannot satisfy the [+WH] subcategorization requirement of verbs such as *ingin tahu* ‘wonder’ because it does not undergo movement nor form a formally clause-typed question. Similarly, to the extent that the property in (36c) holds in Indonesian (recall our earlier conclusion that this property itself is silent about whether the movement has occurred or not), the lack of crossover effects is a natural consequence of the fact that *wh*-in-situ does not undergo movement.

The properties in (36d, e) need more elaborate discussion. We have seen in section 3.3 that, according to Saddy (1991), unlike overtly moved *wh*-phrases, in-situ *wh*-phrases do not show scope interaction with quantifiers that c-command them but instead take only widest scope. To the extent that this judgment is real, his observation naturally follows from Reinhart’s assumption that existential closure of a function variable introduced by an NP can be introduced by the highest possible position. Thus, sentences as in (37a), repeated here, would receive an LF representation in (37b) under choice function.

(37) a. Setiap orang men-cintai siapa?

every person TR-love who

‘Who did every person love?’ who>every, *every>who

b. $\{P \mid \exists \langle f \rangle (CH(f)) \ \& \ (\forall x) \ \& \ P = \wedge (\text{person}(x) \rightarrow x \text{ loves } f(\text{person})) \ \& \ \text{true}(P))\}$

In the LF representation in (37b), the choice function applies to a set of persons in a given world and picks out one of them from this set. The representation states informally that the denotation of the sentence in (37a) is the set of true propositions, each stating that there is a single function f such that, for every x , if x is a person, x loves the person selected by f . This representation, therefore, corresponds to the wide scope reading of the in-situ *wh*-phrase over the universal quantifier in subject position, thereby yielding Saddy’s observation about the widest scope.

We have seen in section 3, however, that Saddy’s characterization of the scope of *wh*-in-situ in Indonesian not only could not be reproduced in the present study but also goes against the judgments reported by Cole and Hermon (1998) and others. This raises the question of whether the existential operator that binds the function variable can be introduced in the scope of another quantifier. If the answer is yes, the scope interaction between *wh*-in-situ and c-commanding universal quantifiers, as reported by

my language consultants, is naturally predictable. Indeed, Reinhart shows, citing examples as in (38a) from Ruys (1992), that the operator can be inserted within the scope of another operator, as in (38b). See also Winter (1997) for the claim that function variables can be bound by an operator at any level.

(38)a. Most linguists have looked at every analysis that solves some problem.

b. For most linguists x , $(\exists (f)) (CH(f) \ \& \ (\forall y) (\text{analysis}(y) \ \text{and} \ y \ \text{solves} \ f(\text{problem})) \rightarrow (x \ \text{looked at} \ y))$.

(Reinhart 1998: 40)

According to Reinhart (1997: 40), in the example in (38a), “the choice of a problem may vary with the choice of a linguist, in which case some problem is not “specific.” Nevertheless it can take scope over every analysis.” Thus, this intermediate reading of the indefinite *some problem* is naturally accounted for if the existential operator is below another quantifier *most*, as shown in (38b). This analysis, of course, predicts that the corresponding *wh*-in-situ in Indonesian should also be able to take this intermediate scope in Indonesian sentences structurally akin to the English sentence in (39a). This

prediction is indeed confirmed by the same consultant. The LF representation for (39a) then looks like (39b).¹³

(39)a. Tiga siswa mempertimbangkan setiap analisis yang memecahkan
 three student consider every analysis that solve
 masalah yang mana.
 problem that which

‘Three students considered every analysis that solved which problem?’

three> which>every (the nature of problem is different depending on each student; the nature of analysis is different depending on each problem)

b. For three students x , $(\exists (f)) (CH(f) \ \& \ (\forall y) (\text{analysis}(y) \ \text{and} \ y \text{ solves } f(\text{problem})) \rightarrow (x \text{ consider } y))$.

¹³ For this case, I have elicited the judgment reported in (39) as follows. I asked the three consultants whether the sentence can be used for the following scenario and wrote the following diagram: there are three students (A-C), each one of the students works on three different problems (1-3), and each one of the problems has two different analyses that have been proposed in the history of academia (U-Z).

(i) Student A	Student B	Student C
Problem 1	Problem 2	Problem 3
Analysis U	Analysis W	Analysis Y
Analysis V	Analysis X	Analysis Z

The consultants all reported that the sentence can be used to describe this scenario, which instantiates the intermediate reading we are interested in.

This fact would remain mysterious under Saddy's (1991) account because it crucially depends on the observation that *wh*-in-situ in Indonesian *always takes the widest possible scope*.

The same observation also allows us to tease apart several competing theories of choice function. Thus, Kratzer (1998) and Matthewson (1999) (see relevant discussion in note 5) both argue that choice function variables can be bound only by an operator *that is in the highest possible position*. Drawing on Fodor and Sag's (1982) observation that indefinites in English are ambiguous between a specific and a quantificational expression, Kratzer argues that an indefinite must have widest scope by being evaluated by choice function when it is specific. The consequence of this analysis is that the intermediate scope reading of the kind we observed in English and Indonesian should be impossible. Matthewson provides further arguments for Kratzer's ambiguity-based approach to the scope behavior of indefinites from her analysis of indefinite determiners in St'át'imcets (Lillooet Salish). Observing that indefinite determiners in this language come in two types, the polarity determiner *ku* and non-polarity determiners, Matthewson shows that a non-polarity determiner-headed element such as *ta twíw't-a* 'a child' must take the widest possible interpretation, as shown in (40).

(40) [tákem i wa7 tsunám'-cal]cuz' wa7 qwenúxw-ahhts'a7 ih-káw-lec-as

[all Det.Pl Prog teach-Intr going.to Prog sick-inside Hyp-far-Intr-3Conj

[ta twíw't-a]

Det child-Det

'Every teacher will be sad if a child quits.'

- i. Accepted in context: There is one child, who every teacher doesn't want to leave.
- ii. Rejected in context: For each teacher, there's one child who s/he doesn't want to leave.
- iii. Rejected in context: Every teacher will be sad if any child leaves.

√ widest *intermediate *narrowest

(Matthewson 1999: 119)

We have seen above, however, that this kind of intermediate scope reading *is* possible in English and Indonesian. Thus, examples as in (38a) and (39a) serve to tease apart the predictions of Reinhart's/Winter's analyses vs. Kratzer's/Matthewson's analyses regarding the possible position of the existential quantifier and support the idea in the former type of analysis that nothing in principle blocks the introduction of existential closure in the scope of another quantifier/operator.

Let us now turn to the property in (36d), namely, that *wh*-in-situ in Indonesian is not able to support a pair-list reading for multiple *wh*-questions, as reported by Saddy (1991). In section 3, we have seen that this property is also controversial, as my consultant work has indicated. Nonetheless, whichever judgment is correct in certain dialects of Indonesian follows under the choice function approach because nothing blocks base-generation of the existential operator in the highest position or in the scope of another quantifier; the first option yields Saddy's observation while the second option yields the scope interaction reported by my consultants.

Thus, the choice function approach proposed by Reinhart allows for a unified explanation of a set of structural and interpretive properties associated with *wh*-in-situ in Indonesian, some of which would remain mysterious under syntactic movement/unselective binding/absorption/definite description-based accounts.

4.3. *New Predictions: The NP vs. non-NP Asymmetry*

Given the definition of choice function given in (34), we can make the prediction that the availability of licensing via choice function crucially depends on whether a given in-situ *wh*-phrase in Indonesian can denote an N-set. In other words, we predict that the NP vs. non-NP constraint should be observed essentially in the same way as in the English

examples in (29a, b, d). Cole and Hermon (1998: 226) show that this prediction is indeed borne out. Examples in (41a-g) below are constructed in Indonesian based on the related but partial paradigm from Malay reported in Cole and Hermon (1998: 226).

- | | |
|--|---|
| <p>(41) a. Siapa mem-beli buku?</p> <p>who AV-buy book</p> <p>‘Who bought a book?’</p> | <p>e. Esti mem-beli buku dengan cara apa?</p> <p>Esti AV-buy book in way what</p> <p>‘In what way did Esti buy a book?’</p> |
| <p>b. Esti mem-beli apa?</p> <p>Esti AV-buy what</p> <p>‘What did Esti buy?’</p> | <p>f.* Esti mem-beli buku mengapa?</p> <p>Esti AV-buy book why</p> <p>‘Why did Esti buy?’</p> |
| <p>c. Esti mem-beli buku dimana?</p> <p>Esti AV-buy book where</p> <p>‘Where did Esti buy a book?’</p> | <p>g. Esti mem-beli buku untuk apa?</p> <p>Esti AV-buy book for what</p> <p>‘For what did Esti buy a book?’</p> |
| <p>d.* Esti mem-beli buku bagaimana?</p> <p>Esti AV-buy book how</p> <p>‘How did Esti buy a book?’</p> | |

The examples in (41a, b) are naturally predicted to be grammatical because the in-situ *wh*-phrases *siapa* ‘who’ and *apa* ‘what’ denote an N-set (the set of persons and the set of things, respectively). The example in (41c) also falls into place because *demana* ‘where’ is actually bimorphemic, consisting of *di* ‘in’ and *mana* ‘place’. What is crucial for us is the contrast between (41d, f), on the one hand, and (41e, g), on the other. The examples in (41d, f) are both ungrammatical because *bagaimana* ‘how’ and *mengapa* ‘why’ are *wh*-adverbials, hence need to undergo movement to be licensed but there is no morphological reason to legitimize this movement. The examples in (41e, g) are both grammatical because *cara apa* ‘what manner’ and *untuk apa* ‘for what’ clearly contain an NP that contributes an N-set required for choice function to apply. Note that this contrast provides another argument against the Huang-style ECP-based LF movement analysis of *wh*-in-situ in Indonesian, as the LF movement of the in-situ *wh*-phrases in (41e, g) would uniformly violate the ECP, rendering these examples ungrammatical.

4.4. *Cole and Hermon’s 1998 Unselective Binding vs. Choice Function*

A different implementation of the non-syntactic, in-situ approach to *wh*-in-situ has been independently proposed by Cole and Hermon (1998, 2000) on the basis of data from Malay. In this subsection, I consider whether the choice function approach and their fine-

grained unselective binding approach would make any different empirical predictions concerning *wh*-in-situ in Indonesian and show that the “Donald Duck” Problem and the intermediate scope reading of *wh*-in-situ in multi-clausal *wh*-questions pose non-trivial problems for Cole and Hermon’s analysis.

Cole and Hermon (1998: 240) propose that “in *wh*-in-situ in Malay the (*wh*-OP) question operator is merged at the root Spec CP, and, therefore, unselectively binds a *wh*-variable in its scope.” This analysis derives essentially the same set of facts concerning Indonesian *wh*-in-situ as the choice function analysis. There are no island effects in syntax or at LF because this class of expression does not undergo syntactic movement but instead gets interpreted in situ by being bound by an operator base-generated in the matrix specifier of CP. The fact that *wh*-adverbials cannot be interpreted in situ also follows from the inability of such elements to be bound by unselective binding along lines suggested by Reinhart and Tsai. It appears, then, that their analysis might amount to the same thing as our choice function approach. Indeed, they note (Cole and Hermon 1998: 240) that “since it does not affect the issues under consideration in this paper, we will maintain the pretense that the question operator binds the *wh*-variable directly rather than through the mediation of a choice function, and shall continue to employ the term ‘unselective binding’.” The same position is maintained in Cole and Hermon (2000: 106), who remark that “Reinhart 1995 argues that the correct mechanism for in situ interpretations of *wh* is

a choice function rather than unselective binding. We leave this issue open since the precise mechanism for in situ interpretation is irrelevant for our analysis.” Cole and Hermon (2000), however, is more explicit in their analysis of *wh*-in-situ, as shown in their schematic representation of their unselective binding analysis (p. 109).

(42) Unselective Binding of *Wh*-in situ:

$$[_{CP} OP_i [_{CP} \dots wh_i]]$$

where *wh* is a variable in a base-generated position and *OP* is base-generated in scopal position and binds *wh*. (Cole and Hermon 2000: 109)

The explication of Cole and Hermon’s (1998, 2000) analysis above, therefore, might give the impression that the choice function approach is a notational variant of their proposed version of unselective binding. Nonetheless, I show below that the predictions do diverge in a number of domains related to the “Donald Duck” Problem and the intermediate scope reading and that the choice function analysis is superior to Cole and Hermon’s analysis on these grounds.

The first divergence in prediction between the choice function approach and Cole and Hermon’s (1998, 2000) version of the unselective binding concerns the interpretation of

in-situ *wh*-phrases in Indonesian contained within an *if*-clause. Recall that Reinhart showed that the LF representation of examples such as (31), repeated here as (43), which would be derived under selective binding, fails to express the fact that the value of the denotation of the in-situ *wh*-phrase must be selected from the set of philosophers because this approach would leave the restriction (philosopher) in the implicational clause at LF, as shown in (32a, b) (= (44a, b)), and render the sentence in (32c) (= (44c)), true even though the value of *y* is Donald Duck, contrary to facts.

(43) Who will be offended if we invite which philosopher? (Reinhart 1998: 36)

Wrong:

(44)a. for which $\langle x, y \rangle$, if we invite *y* and *y* is a philosopher, then *x* will be offended.

b. $\{P(\exists \langle x, y \rangle)\} \ \& \ P \models ((\text{we invite } y \text{ and } y \text{ is a philosopher}) \rightarrow (x \text{ will be offended}) \ \& \ \text{true}(P))$

c. Lucie will be offended if we invite Donald Duck.

(Reinhart 1998: 36)

As we saw in section 4.1, this Donald Duck problem won't arise under the choice function approach because the value of *y* must be selected from the non-empty set of philosophers in a

given model/world by virtue of the very definition of choice function given in (34). Therefore, Cole and Hermon's analysis cannot capture the correct interpretation unless it is accompanied with special mechanisms of pulling out the restriction out of the antecedent of an implicational clause to circumvent the "Donald Duck problem. Several technical additions, of course, would not be inconceivable that would potentially solve this problem; for example, the representations in (44a, b) could be converted to something like (45a) below, in which the value of y can only be selected from the set of philosophers. However, as Reinhart (1992) shows, whatever addition that yields this presentation would rule in not only the desired reading in (45b) but also the undesired reading in (46c).

- (45) a. Lucie will be offended if Donald Duck is a philosopher and we invite him.
- b. Lucie will be offended if we invite Kripke.
- c. Lucie will be offended if Kripke is a philosopher and we invite him. (Reinhart 1992: 3)

The point here is that no such special addition is required under the proposed approach; instead, the effects of such stipulations are automatically derived from the definition of choice function as applying to a (non-empty) set of individuals and picking up one member from that set.

The second empirical domain in which the predictions of the two competing approaches would diverge concerns the intermediate scope reading illustrated by data as in (39a), repeated as (46a). The LF representation in (46b) is derived under the choice function, which correctly captures the reading where *masalah yang mana* ‘which problem’ takes wide scope over *setiap analisis* ‘some analysis’ but takes narrow scope with respect to *tiga siswa* ‘three students.’

- (46)a. Tiga siswa mempertimbangkan setiap analisis yang memecahkan
 three student consider every analysis that solve
 masalah yang mana.
 problem that which

‘Three students considered every analysis that solved which problem?’

three> which>every (the nature of problem is different depending on each student;

the nature of analysis is different depending on each problem)

- b. For three students x , $(\exists (f)) (CH(f) \& (\forall y) (\text{analysis}(y) \text{ and } y \text{ solves } f(\text{problem})) \rightarrow (x \text{ consider } y))$.

It is not clear, however, whether this intermediate scope reading would be derived under Cole and Hermon’s version of the unselective binding approach. In their 1998 paper, they

assume that the *wh*-in-situ in Malay is bound unselectively by the operator base-generated in the root [Spec, CP]. Accordingly, it would falsely predict that the intermediate scope reading observed in (46a) would be impossible. The same problem remains with their analysis updated in their 2000 paper explicated in (42) because it base-generates the operator in the scopal [Spec, CP]. Crucially, however, the intermediate reading in (46a) requires that the operator must be base-generated in a position in the matrix clause that is lower than the specifier of the matrix TP but higher than the complement of the matrix VP. Thus, the relevant reading would be impossible, contrary to facts.

Based on these two divergences between Cole and Hermon's (1998, 2000) analysis and the choice function analysis, I conclude that the two analyses are not entirely the same; the latter analysis makes better empirical predictions concerning the "Donald Duck" Problem and the intermediate scope reading in multi-clausal *wh*-questions. The two problems could be technically solvable by several special amendments on the mapping from syntax to LF under Cole and Hermon's analysis but the fact that these amendments are not necessary but instead derived from the way choice function independently works provide support in favor of the choice function analysis adopted in this chapter.

5. Conclusions

Let me summarize the whole chapter. I have started by reviewing several peculiar structural and interpretive characteristics associated with *wh*-in-situ in Indonesian originally discovered by Saddy (1991). The lack of island effects and the failure of the [+WH] subcategorization requirements show that neither overt movement analyses as in Watanabe (1992, 2001) and Richards (2001) nor covert movement analyses as in Huang (1982) and Lasnik and Saito (1992) are not suitable mechanisms of licensing *wh*-in-situ in Indonesian. The difference in morphological composition between *wh*-phrases in English and Indonesian also casts doubts on the general applicability of Pesetsky's (1987) variant of the unselective binding approach to Indonesian. During the course of this demonstration, I have reported judgments from my language consultants that are contradictory with those reported by Saddy with respect to the lack of pair-list readings in multiple questions and the obligatory widest scope reading of in-situ *wh*-phrases.

I have proceeded to show that there are several problems with Saddy's own analysis of *wh*-in-situ in Indonesian as an interrogative definite description. In particular, this analysis is in direct contrast with the assumption since the earliest days of generative grammar that *wh*-phrases are indefinite expressions that contain an individual variable. Of course, the island-insensitive behavior of *wh*-in-situ in Indonesian could follow from Saddy's

treatment of this class of expressions as definite descriptions, but this analysis directly contradicts the motivated assumption about *wh*-phrases stated above; rather, a more natural possibility is to show that the scope-taking property is a special case of existential indefinites, which are independently known to freely take wide scope, in massive violation of known island constraints on syntactic movement. The evidence based on reduplication from Cole and Hermon (1998) constitutes further evidence against the characterization of the in-situ *wh*-phrases in terms of definite description.

In light of these considerations, I have proposed that all the apparent peculiar properties of *wh*-in-situ in Indonesian receives a unified explanation once we allow existential quantification over choice functions along the lines suggested by Reinhart (1992, 1997, 1998, 2006). I have further compared this analysis with a fine-grained analysis of unselective binding developed by Cole and Hermon (1998, 2000). Based on facts regarding on the “Donald Duck” Problem and the intermediate scope reading of *wh*-in-situ in multi-clausal contexts, I have argued that the former approach is superior to the latter approach in that it requires no special technical stipulations that would be required under Cole and Hermon’s analysis. The overall result of this chapter, therefore, provides support that choice function should be admitted as a natural mechanism of *wh*-licensing on the part of the semantic interface that does not depend on syntax.

One question that the proposed analysis leaves unresolved is, why choice function seems not to apply for embedded questions headed by interrogative verbs?¹⁴ We have seen from the contrast between (6a) and (6b) that verbs like *ingin tahu* ‘wonder’ necessarily requires overt movement of a *wh*-phrase into the specifier of its complement CP (see also notes 5 and 12 for relevant discussion). This question could be resolved if “part of the mechanism of quantificational WH construal is derivative from the simple fact of movement having taken place.” (Saddy 1991: 212). Specifically, the fact that overt movement occurs means that the (regular) quantificational *wh*-construal is available for the LF interface. Therefore, starting up a new domain-specific operation (choice function) is more costly than the already activated syntactic operation. Notice that the former option is also prohibited independently because of the minimalist interface thesis that the semantic interface attempts to blindly follow the carve parametrically defined by syntax. In Chinese and Japanese, however, syntactic movement as a way of *wh*-construal is impossible for independent syntactic reasons (such as strength of Q-features). Thus, the LF component activates the choice function option as a case of Interface Economy (Reinhart 2006) to derive a variety of

¹⁴ Thanks to Heidi Harley (personal communication) for raising this important question.

interpretations that other movement languages would derive by regular *wh*-movement (but cf. the discussion in note 5 above for an alternative viewpoint).

The results achieved in this chapter have certain important implications for the way the syntactic computation interacts with the semantic component. As I have shown in this chapter, the choice function analysis along the lines proposed by Reinhart (1992, 1995, 1997, 1998, 2006) provides a straightforward solution to various syntactic and semantic problems in Indonesian with movement-based and unselective binding analyses proposed for other Asian languages such as Chinese and Japanese. The plausibility of the choice function analysis appears to suggest that the semantic component is relatively independent from the needs of syntactic computation *per se* but serves more the needs of the system of concepts (inferences, contexts, concepts) that interacts with the faculty of language. Of course, the fact that the NP vs. non-NP distinction is crucial for the application of choice function indicates that certain purely syntactic information is used for the purposes of semantic interpretation on the part of the linguistic interface between syntax and the conceptual system, but it may well be the other way around; the NP vs. non-NP distinction is a syntactic manifestation of the application of choice function and the lack thereof; syntax happens to provide sufficient tools to work out the semantic interpretation right on the part of the syntax-external interpretive component. One

speculation that would emerge from the present study of *wh*-in-situ in Indonesian is, then, that the relevant interface does whatever it can to meet the needs of communication.

A natural question to ask here, then, is why doesn't this option manifest itself with English *wh*-questions. More specifically, why does Indonesian allow three types of questions (overt *wh*-movement, partial *wh*-movement, and *wh*-in-situ) whereas English allows only overt *wh*-movement and *wh*-in-situ? This difference between Indonesian and English as well as its semantic corollaries may well be captured as another case in which the minimalist interface thesis is at work; it is because the LF interface blindly follows the carve parametrically defined by syntax. Cole and Hermon (1998, 2000) propose a parametric theory of *wh*-questions whereby languages differ in whether a question word consists in an operator and variable combined in a single word or of a variable generated in a thematic position separately from a base-generated Q-operator. Malay/Indonesian can take both feature bundling options. Suppose that the variable is generated separately from the base-generated *wh*-operator. Then, since there is no morphological feature to legitimize the movement of the variable, the Economy of Derivation dictates that there is no movement. To satisfy the communicative needs of the C-I system that would have been met by this movement, the LF interface applies choice functions to this in-situ object. If the variable and the operator are bundled together in a single word, then full syntactic movement must occur to create an operator-variable structure and obtain

appropriate interrogative scope. The partial movement option could arise in Malay/Indonesian either as a type of prolepsis (cf. Davies 2003, 2005) or due to the existence of the null *wh*-expletive (Cole and Hermon 1998) in these languages. In English, on the other hand, the *wh*-operator and variable must be bundled together but there is an additional requirement that only one *wh*-phrase suffices to type a clause as a question; other *wh*-expressions must remain in situ and hence evaluated at the interface via choice functions. In this way, the difference in diversity of *wh*-questions between Malay/Indonesian and English follows from the way that certain morphosyntactic features are bundled in each language and the way LF reacts to it. Whether there is independent morphological/syntactic evidence for this type of lexical parameterization in domains other than *wh*-questions is an important task to be undertaken.

We have seen thus far in this dissertation that the syntax-external linguistic interfaces are subservient, doing the best they can to meet the needs of the universal computational system, syntax, within the parametrically defined range of options available to a particular language. We have seen in the preceding chapters that this characterization is on the right track with respect to the interaction of syntax with phonology and certain aspects of semantics such as the denotation of bare nominals. The results reported in this chapter seem slightly deviant from those reported in the earlier chapters in that they suggest that the semantic component is relatively more independent from the requirements of syntax and serves the

communicative needs imposed from our language-independent conceptual system, as hinted at the outset of section 4. However, it is also possible to draw a different conclusion from the current investigation of *wh*-in-situ in Indonesian, namely, that the choice function approach is developed in the syntax-external semantic component due to the very fact that syntax does not move *wh*-in-situ in tandem with the language-external requirements from the conceptual system. Within the deterministic theory of movement, as adopted in the Minimalist Program, movement occurs only to satisfy the morphological features of the element that moves/is moved to. Thus, what syntax can do is to leave a *wh*-in-situ in its base position and send the derivation that contains this structure to the semantic component. It seems, then, that there might be no way that semantics would interpret the relevant structure. However, under the minimalist interface thesis argued for in this dissertation that semantics is purely interpretive, the very fact that syntax sends *this* configuration means that semantics should be rich enough to be able to accommodate and interpret this configuration. There are several ways in which this can be done in the semantic interface. The approach I have pursued in this chapter is that the semantic component independently applies the interpretive mechanism of choice function to satisfy the communicative needs of the language-external conceptual system. This interaction between the linguistic interface and the system of concepts is not unreasonable on two counts. First, one primary way language can be used is communication from the

perspective of the conceptual system, hence the exact form of the interpretive method might be heavily influenced by language-external requirements imposed from the conceptual system. Second, the system of concepts takes care of information such as concepts, logical reasoning, contexts, beliefs and complex thoughts. Thus, the choice function and its accompanying notions such as existential quantification, set formation, non-empty set, and so can be analyzed as the linguistic manifestation of the interface strategy that the linguistic semantic component happens to have selected on the basis of its interaction with the concepts system.

To the extent that the above noted scenario is tenable, it indicates that syntax is actually an “imperfect” system for the purposes of communication; the faculty of language is poorly designed for this purpose. This is because it does not provide sufficient information that the semantics could access in order to find a unique way to accommodate a multitude of communicative needs as required by the conceptual system. This is hardly surprising, of course, under the minimalist interface thesis, according to which syntax is totally blind to the external linguistic interface such as phonology and semantics, as I have argued for so far in this dissertation. Evidence abounds that this is the correct characterization of syntactic system. Embedding in the syntax as in *Why did you think that you kissed me?* cannot help creating ambiguities that should be ideally avoided for the purposes of communication (as in the Gricean theory of cooperative maxims). See also Hinzen (2006) for many other examples

to illustrate this blind nature of syntactic computation, some of them already discussed in the previous chapter of this thesis. I provide one possible explanation for why syntax has this property of “imperfection” in terms of the evolution of the language faculty.

Based on this consideration, I conclude that syntax is indeed an autonomous computational system that is blind to the consequences it has for the syntax-external semantic component. Syntax sends its final outputs (e.g., structures that contain in-situ *wh*-phrases) to semantics, which assigns an interpretation to them via whatever domain-specific mechanism it happens to have at its disposal (e.g., choice function) from the way it interacts with the language-external conceptual system. However “poor” the result of syntactic computation may be from the perspectives of the conceptual system, the semantic component devises ways to legitimize, repair and improve them to satisfy the requirements of human communication. This scenario, therefore, allows us to maintain the minimalist interface, argued for thus far in this thesis, namely, that the syntax-external components that connect meaning and sound to structure are purely interpretive, doing whatever domain-specific operations it can to make certain syntactic objects legible for language-external systems to use but in a way that it is totally subservient to the principles and parameters active in a given language. I return to further ramifications of this conclusion in chapter 7.

CHAPTER 6 REDUPLICATION ASYMMETRIES IN INDONESIAN: THEIR THEORETICAL IMPLICATIONS FOR THE SYNTAX-MORPHOLOGY INTERFACE ¹

1. Introduction

In this chapter, I explore issues revolving around the morphology-syntax interface with a case study in reduplication in Indonesian. Through detailed discussion of a certain asymmetry between nominal and verbal reduplication in this language, I propose further evidence for the thesis of minimalist interface that the syntax-external phonological component is fundamentally interpretive, assigning phonological interpretation solely based on the input created by the universal syntactic computation.

I couch the examination of the validity of the thesis of minimalist interface within the long-standing debate in the contemporary linguistic theory between lexicalist and non-lexicalist theories of the lexicon-syntax interface. I argue that the minimalist interface thesis leads us to the non-lexicalist view as hold in the recent morphosyntactic framework of Distributed Morphology (Halle and Marantz 1993, 1994) that words are not a triplet of sound, meaning, and their correspondence, as commonly assumed, but rather nothing but the morpho-

¹ This chapter is a substantially expanded version of the paper presented at the 33rd Annual Meeting of the Berkeley Linguistic Society held at the University of California, Berkeley with Bradley McDonnell and reported in Sato and McDonnell (in press) and Sato (2008). I am very grateful to Bradley for our corroboration, in particular, his timely help in conducting a corpus study mentioned in Table 5 that forms the basis of this chapter.

phonological output of the complex objects created by syntactic derivation. According to this view, the syntactic derivation constructs whatever objects it can, based on a particular arrangement of morpho-syntactic features available in a given language; what the syntax-external phonological component can do is to assign a language-particular surface realization to this output object post-syntactically. In this chapter, I argue that this view is correct by showing that reduplication in Indonesian is sensitive to syntactic structure, not just morphemes, thereby providing empirical support for the framework of Distributed Morphology.

With the phonological feature assignment being purely interpretive and post-syntactic and governed by the way syntactic computation unfolds, the notion of minimalist interface leads us to another claim that there is no such thing as the lexicon in the traditional sense in which a word is constructed by processes different from syntactic combinatorial processes and assigned a meaning and sound pre-syntactically. I show that this non-lexicalist approach to the lexicon-syntax interface provides a natural account of reduplication asymmetries in Indonesian. At the same time, I show that several variants of the lexicalist theory have difficulties in accounting for the attested pattern of reduplication in Indonesian precisely because they postulate the traditional lexicon as a pre-syntactic/autonomous generative component. This result, therefore, provides a

morphological piece of evidence for the non-lexicalist view of the lexicon-syntax interface entailed by the thesis of minimalist interface.

The empirical domain I investigate is reduplication in Indonesian. A corpus survey of four popular newspapers published in Indonesia reveals a curious asymmetry between nominal and verbal reduplication that has not been reported in the literature on Indonesian morphology: nominal stems allow both stem and stem-affix reduplication whereas verbal stems allow only stem reduplication. I show how this asymmetry as well as the word-internal stem reduplication pattern pose non-trivial empirical and architectural difficulties for several versions of the lexicalist theory as presented in Chomsky (1970), Anderson (1982, 1992), Kiparsky (1982a, b, c, 1985), Mohanan (1986), and Di Sciullo and Williams (1987)/Williams (2007).² I show that these difficulties arise as the artifact of traditional

² A similar problem has been independently noted in Yaqui reduplication by Haugen and Harley (2006), though with certain properties that distinguish it from the Indonesian reduplication pattern to be discussed in this chapter. Haugen and Harley observe that, in Yaqui, the inflectional process of reduplication targets the head of the verbal head of an N + V compound rather than the compound itself, as shown in (ia-c).

(i) Word-internal head-reduplication in Yaqui

- | | | | | | |
|----|-------------------|--------------|---|-----------------------|-------------------------------|
| a. | kuta-siute | ‘wood-split’ | ⇒ | kuta-siu-siute | |
| | stick-tear | | | stick-Red-tear | |
| b. | chit-wat-te | ‘spitting’ | ⇒ | chit-wat-wat-te | |
| | saliva-throw-Intr | | | saliva-Red-throw-Intr | |
| c. | hiavih-muuke | ‘gasping’ | ⇒ | hiavih-mu-muuke | |
| | breathe-die | | | breathe-Red-die | (Haugen and Harley 2006:6, 7) |

theories of the syntax-lexicon interface as in the lexicalist theory that postulates the lexicon as an autonomous pre-syntactic generative component or system whose relation with respect to syntax is atomic and asymmetric. The thesis of minimalist interface, by contrast, leads us to the alternative view that words as well as sentences are objects created by the sole generative syntactic component. Given the role of the morpho-phonological component as subservient to the needs of syntax, the same thesis also entails that morpho-phonology is located after syntax and assigns interpretation to the objects created by syntax. These two views amount to the core claims that have been independently made in the morphosyntactic theory of Distributed Morphology (Halle and Marantz 1993, 1994; Marantz 1997; Harley and Noyer 1999; Embick and Noyer 2007). I show that this conception of the lexicon-syntax interface provides a principled explanation of the reduplication asymmetry and the word-internal reduplication pattern in Indonesian; these facts are straightforwardly derived as a natural consequence of a particular hierarchical arrangement of morphosyntactic features, including Asp and Num, which is independently motivated by the semantic/selectional properties of certain derivational affixes such as *ber-* and *an-*. This result, therefore, provides a strong piece of evidence against the traditional lexicalist architecture, and, at the same time,

They argue that this word-internal reduplication pattern presents evidence against lexicalist views of the lexicon-syntax interface as in Di Sciullo and Williams (1987) and Anderson (1982, 1992) and present a new analysis of this pattern within the framework of Distributed Morphology.

argues in favor of more recent non-lexicalist theories of the syntax-lexicon interface as in Distributed Morphology that attempt to locate all types of word formation within the sole realm of the syntactic derivation. Accordingly, I construe this result as further empirical evidence for the thesis of minimalist interface that the syntax-external interface component is a fundamentally interpretive system whose role is to assign whatever interpretation it can to the outputs delivered by the universal syntactic computation.

The present chapter is organized in the following manner. In the next section, I provide an overview of the so-called lexicalist theory in the generative tradition and compare its grammatical architecture with more recent, non-lexicalist theories of the lexicon-syntax interaction as in Distributed Morphology. I highlight one primary difference between the two theories concerning the way the lexicon interacts with the syntax. Lexicalist theories predict that no “lexical” processes, defined along a variety of dimensions, can follow syntactic combinatorial processes such as Merge and Move (Chomsky 1995) due to their postulation of the lexicon as a pre-syntactic generative component. In contrast, non-lexicalist theories predict that there is no inherent ordering between the two types of operations because they do not posit such a component, relegating the roles of the lexicon to the sole generative syntactic component. This difference becomes important in later sections. In section 3, I report the results of a corpus survey of four popular newspapers in Indonesian. The most

important among them is my new finding initially made by the pilot study in Sato and McDonnell (in press) that there is an asymmetry between nominal and verbal reduplication in Indoensian, namely, that derivational nominal affixes are productive in reduplicating both the stem-affix combination and the stem alone whereas derivational verbal affixes allow stem reduplication, but never stem-affix reduplication. In the next two sections, I consider implications of this asymmetry for the proper theory of the syntax-lexicon interface. In section 4, I show that this asymmetry as well as the word-internal reduplication pattern poses non-trivial architectural/empirical difficulties for several, well-known versions of the lexicalist theory as in Chomsky (1970), Anderson (1982, 1992), Kiparsky (1982a, b, c 1985)/Mohanan (1986), and Di Sciullo and Williams (1987)/Williams (2007). Based on this consideration, I show, in section 5, that the two observations concerning reduplication in Indoensian can be given a straightforward account under non-lexicalist, morpho-syntactic theories of the syntax-lexicon interface as in Distributed Morphology if we take into account a particular hierarchical arrangement of certain morphosyntactic features such as Asp and Num as well as the underlying syntactic category of input stems for reduplication. In section 6, I provide a summary of the contents of this chapter and discuss their implications for the proper theory of the syntax-lexicon interface and minimalist interfaces.

2. Lexicalist vs. Non-Lexicalist Theories

In this section, I compare two competing theories of the syntax-lexicon interface, restricting attention to the debate within the framework of generative grammar: lexicalist theories and non-lexicalist theories.

2.1. *Lexicalist Theories*

In this subsection, I introduce the main claims of the lexicalist theory. I show that the most important theoretical tenet of this hypothesis for the purposes of this chapter is that the lexicon is postulated as an autonomous generative component prior to the syntax that is responsible for certain morphologically derived complex objects that are identified along several dimensions.

The theory of the syntax-lexicon interface most often termed the *Lexicalist Hypothesis* (Chomsky 1970, 1993, 1995; Lieber 1980; Williams 1981; 2007; Anderson 1982, 1992; Farmer 1982; Lapointe 1980, 1981; Jensen and Stong-Jensen 1984; Di Sciullo and Williams 1987; Pesetsky 1979; Kiparsky 1982a, b, c, 1985; Mohanan 1986) claims that a) there are two independent generative components for the formation of words and phrases and that b) there is a strict division of labor between the two components. Under this traditional architecture of the lexicon-syntax interaction, the products of the operations in the lexical component serve as atomic indivisible units that syntactic processes operate on

as terminal nodes. As a result, the lexicalist theory adopts one or the other version of the so-called *Lexical Integrity Hypothesis*, according to which principles of syntax are not operative in generating the structure of words, the products of the lexical component. This hybrid approach to word formation stems primarily from the time-honored observation that “words” are somehow different from “phrases” along several (somewhat unclear) dimensions, including semantic and phonological idiosyncrasies/compositionality, gaps/productivity, and derivation/inflection. As a natural consequence of the strict division of labor between the lexicon and syntax, under this lexicalist architecture of grammar, there is no reason to expect that the interface of the syntax and the lexicon is direct; rather, the interface between the two components may well be opaque.

Under the standard interpretation of the history of the theories of the lexicon-syntax interface from the early 1970s to the present within the framework of generative grammar, the Lexicalist Hypothesis comes in two varieties, *strong* and *weak* versions. The strong version of the Lexicalist Hypothesis, represented by work as in Lieber (1980), Lapointe (1980, 1981), Williams (1981), Farmer (1982), Pesetsky (1979), Kiparsky (1982a, b, c, 1985), Mohanan (1986), Di Sciullo and Williams (1987) and Chomsky (1993, 1995), holds that all word formation processes occur in the pre-syntactic lexical component. The weak version of the Lexicalist Hypothesis, which has been most often associated with

Chomsky (1970) (“Remark on Nominalizations”) and Anderson (1982, 1992) in the literature, maintains that certain (regular, productive) word formation processes occur in the syntax whereas other (irregular, non-productive) processes occur in the pre-syntactic lexicon in a way that is conditioned by a variety of criteria, including productivity, derivation vs. inflection, and semantic and/or morphological idiosyncrasies. This is one interpretation of Chomsky’s position, for example, that by Spencer (1991):

(1) Spencer’s (1991: 69) interpretation of Chomsky’s (1970) work

Chomsky argued that transformations should capture regular correspondences between linguistic form, and that idiosyncratic information belonged in the lexicon ... derived nominalizations are morphologically, syntactically and semantically idiosyncratic...

Against this interpretation of Chomsky’s work as the birthplace of the (weak) lexicalist theory, Marantz (1997) claims that Chomsky actually argued against a lexicalist treatment of derived nominalizations by showing that such a treatment needs to stipulate a uniform pattern concerning the unacceptability of the transitive use of internally caused change of state predicates (Levin and Rappaport-Hovav 1995) such as *growth* (as in **John’s growth of tomatoes*). In contrast, this pattern would be naturally explained by the lack of (a certain

type of) v in a syntactic approach to word formation. In this chapter, I follow Spencer's interpretation of Chomsky's work for the purposes of discussion.

Scholars working within the Minimalist Program (that do not adopt the theory of Distributed Morphology; see section 2.2) seem to assume the strong version of the hypothesis as defined above, essentially following the idea of Chomsky (1993, 1995) that syntax selects *fully inflected lexical items* from the Numeration and combines them by the recursive process of Merge.³ Within the morphological research, however, the lexicalist hypothesis is split into the weak and strong versions as defined above. Non-Chomskyan declarative frameworks, including Lexical Functional Grammar (Kaplan and Bresnan 1982; Bresnan 1982), Head-Driven Phrase Structure Grammar (Pollard and Sag 1994) and Construction Grammar (Goldberg 1995, 2006), all adopt what can be termed 'Hyper-Lexicalism', according to which *all* operations for word and phrasal formation occur in the lexical component.⁴

Given below are some definitions of the Lexicalist Hypothesis from the lexicalist literature.

See section 4 for detailed discussion of the weak and strong lexicalist theory presented in

³ This remark seems to be true until quite recently, though, in his most recent work (Chomsky 2006), Chomsky also hints at the following non-lexicalist view of "words", which was in fact what Chomsky (1970) meant to adopt, if Marantz (1997) is correct.

"Possibly the functional category v determines the verbal character of the root R that is its complement, along lines suggested by Alec Marantz, in which case verbal phrases are of the form v -RP." (Chomsky 2006: 12)

⁴ As pointed out to me by Andrew Carnie (personal communication).

Chomsky (1970), Kiparsky (1982a, b, c, 1985), Mohanan (1986), and Di Sciullo and Williams (1987)/Williams (2007).

(2) The Generalized Lexical Hypothesis (Lapointe 1981: 125)

No syntactic rule can refer to an element of morphological structure where element of morphological structure here refers to any morphological feature, any morphological category or any element dominated by such a category.

(3) The Thesis of the Atomicity of Words (Di Sciullo and Williams 1987: 48, 49)

Although syntactic rules can access the categorial status and argument structure of a lexical item, they will never depend on how that categorial status or argument structure was arrived at through morphological derivation or on the internal constituency of words. The rules of syntax can see that a word has such and such properties, but they cannot see how it came to have those properties. ...Words are “atomic” at the level of phrasal syntax and phrasal semantics. The words have “features,” or properties, but these features have no structure, and the relation of these features to the internal composition of word cannot be relevant in syntax.

(4) The Lexical Integrity Principle (Bresnan and Mchombo 1995: 181, 182)

...words are built out of different structural elements and by different principles of composition than syntactic phrases. Specifically, the morphological constituents of words are lexical and sublexical categories-stems and affixes-while the syntactic constituents of phrases have words as the minimal, unanalyzable units; and syntactic ordering principles do not apply to morphemic structures. As a result, morphemic order is fixed, even when syntactic word order is free; the directionality of 'headedness' of sublexical structures may differ from supralelexical structures; and the internal structure of words is opaque to certain syntactic processes.

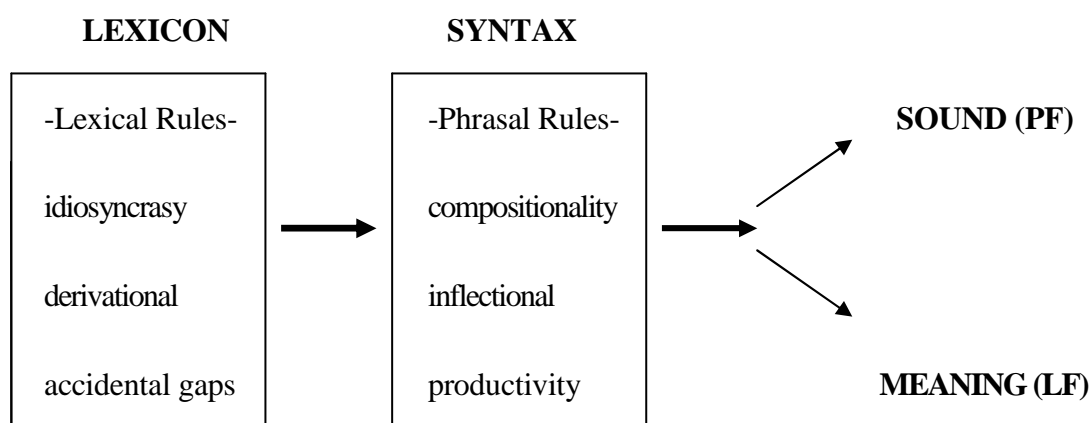
(5) *The Thesis of the Atomicity of Words* (Di Sciullo and Williams 1987: 48, 49)

Although syntactic rules can access the categorial status and argument structure of a lexical item, they will never depend on how that categorial status or argument structure was arrived at through morphological derivation or on the internal constituency of words. The rules of syntax can see that a word has such and such properties, but they cannot see how it came to have those properties. ...Words are "atomic" at the level of phrasal syntax and phrasal semantics. The words have "features," or properties, but these features have no structure, and the relation of these features to the internal composition of word cannot be relevant in syntax.

Although the precise claims about the lexicon-syntax interaction vary from author to author within the lexicalist researchers, the essence of the Lexicalist Hypothesis remains the same, namely, that a) there are two independent components, the lexicon and syntax, and that b) the type of primitives and processes involved in each component are distinct, with the processes of one module being not operative to the structures or operations of the other.

Most commonly, the Lexicalist Hypothesis is implemented in a generative model of grammar (Chomsky 1981, 1986a, 1993, 1995), as shown in (6).

(6) The Lexicalist Hypothesis (embedded within the Government-and-Binding Model)



This architecture entails that the lexical word-building component is sequentially ordered prior to D-structure, the interface of the lexicon and syntax. In other words, the hypothesis endorses a particular sequence of word formation, namely, that all word-building rules in

the Lexicon should precede all phrase-building rules in syntax. Rules in the lexicon are associated with properties such as semantic/phonological idiosyncrasies, derivation and accidental gaps. Rules in the syntax have the complement of these properties such as semantic/phonological compositionality, inflection and productivity. This point is also clarified by Borer's (1998) remark below:

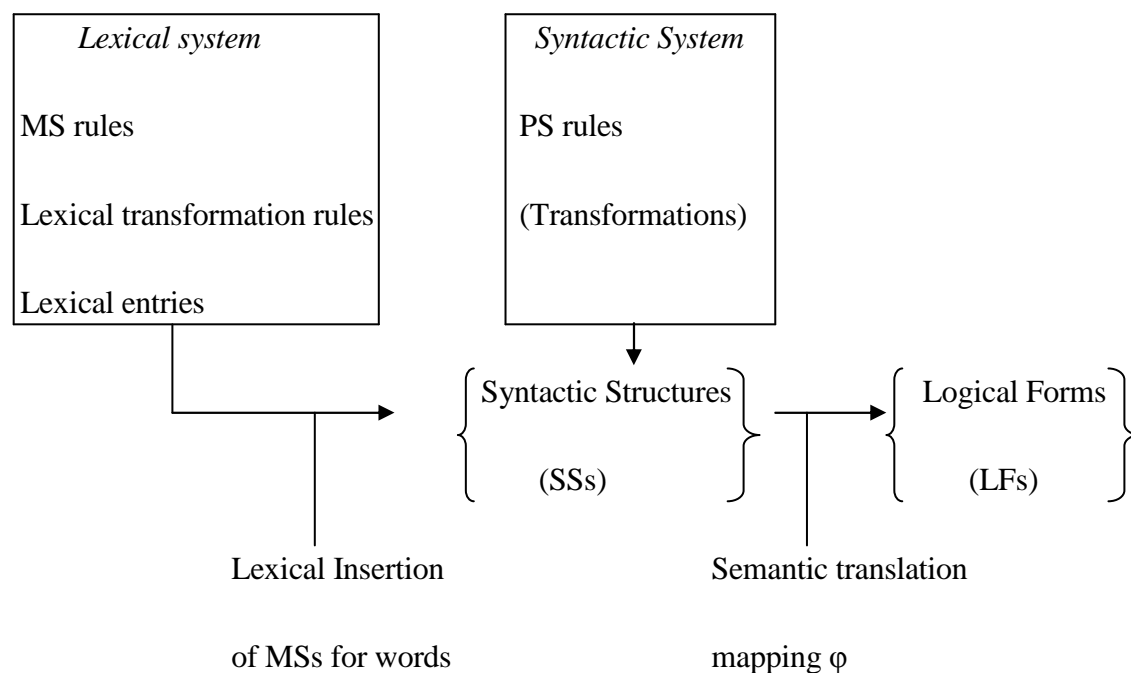
(7) Borer's (1998) Statement of Chomsky's (1970) Weak Lexicalist Hypothesis

The way in which L[exical] I[n]tegrity H[ypothesis] is enforced in many of these models is by assuming that the W[ord] F[ormation] component, as a block of rules, is ordered with respect to the syntax. The WF component and the syntax thus interact only in one fixed point. Such ordering entails that the output of one system is the input to the other. This notion of the autonomy of the syntax and the WF component, and the restricted interaction between them, thus mimics the notion of autonomy developed for the interaction between the syntax and the phonology, where it is the output of the former which interacts with the latter. (Borer 1998: 152, 153)

Lapointe (1981) is one example of work in the lexicalist literature that explicitly mentions the relative sequential ordering of lexical processes with respect to syntactic processes. After the

quote given in (2) above, Lapointe (p. 125) continues as: ‘This framework has the general organization outlined in Figure 1.’ The relevant figure is reproduced in (8), which is one particular instantiation of the general lexicalist architecture in the sense that it posits a pre-syntactic lexical component (lexical System) prior to the generative syntactic component (Syntactic System).

(8) Organization and interactions of the lexical and syntactic systems with the lexical insertion and semantic translation mappings in the grammatical theory of Lapointe 1980.



(Lapointe 1981: 126)

In this architecture, the Lexical System is located prior to the Syntactic System. Based on this kind of remark together with the generative architecture of the generative grammar (e.g. the Government-and-Binding/Minimalist Program of Chomsky 1981, 1986a, 1995) within which the lexicalist hypothesis has been commonly couched, it seems reasonable to make a prediction that lexical processes cannot follow syntactic processes under any type of the lexicalist theory of the syntax-lexicon interface that is tied with the generative model of grammar in which a D-structure/Numeration serves as the interface between the lexicon and syntax.

This characterization is also supported in light of the following consideration. It has been commonly assumed within the lexicalist literature (see Di Sciullo and Williams 1987 and Bresnan and Mchombo 1995; Chomsky 1981, 1986a, 1993, 1995) that principles in the Lexicon operate only on zero-level categories that serve as atomic unanalyzable units which the syntactic derivation uses to create phrasal-level complex objects; in other words, the output of the lexicon seems to be the input for the syntax. It cannot be the case that the output of syntax becomes the input for the lexical processes in the lexicalist sense because this would mean that the lexicon deals with non-zero-level categories, contrary to what lexicalists generally have agreed upon. Therefore, in any lexicalist model of grammar that allows the interaction of lexicon and syntax (in whatever constrained manner it may be), syntax cannot do its job unless the lexicon first provides atomic units. This point is in fact

agreed upon by the lexicalists; for example, Williams (2007: 351) notes that “the channel of communication [between the word system and the phrasal system-YS] is asymmetrical, by virtue of the fact that phrases are made out of words, but not vice versa. Based on this consideration, I assume for the rest of this chapter that, under the lexicalist hypothesis, the lexical component logically precedes the syntactic component, hence, that processes in the lexicon can never follow processes in the syntax.

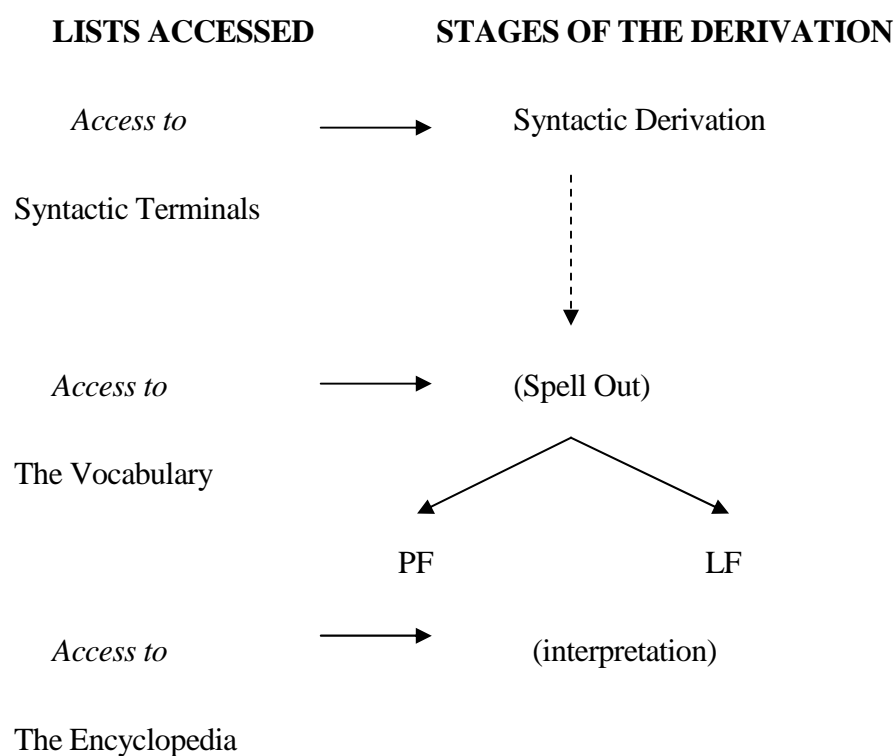
To sum up, I have introduced the general theoretical assumptions of the lexicalist theory as presented in the literature. I have pointed out that one crucial prediction of this hypothesis is that no lexical processes should be able to apply after syntactic processes due to the way that this hypothesis is commonly couched within the generative architecture of grammar where D-structure/Numeration serves as the interface between the lexicon and syntax and the lexicon provides unanalyzable terminal units that syntax operates on to create complex phrasal-level objects. As we will see in the next subsection, this lexicalist view is radically different from more recent non-lexicalist views of the syntax-lexicon interface such as Distributed Morphology.

2.2. Non-Lexicalist Theories

The morphosyntactic framework of Distributed Morphology (Halle and Marantz 1993, 1994; Marantz 1997; Harley and Noyer 1999; Embick and Noyer 2007) claims that there is a

single generative component– syntax – that assembles ‘words’ as well as sentences. This theory of word formation is in direct contrast to the Lexicalist Hypothesis introduced above, which posits two generative components – lexicon and syntax – for word formation. The grammatical architecture of this non-lexicalist framework is given in (9).

(9) The Architecture of Distributed Morphology (adopted from Embick and Noyer 2007: 301)



The model in (9) makes it clear that, within the theory of Distributed Morphology, information that in other theories is assumed to be solely included in the pre-syntactic lexical component is

“distributed” across several components of the grammar: syntax, post-syntactic vocabulary insertion, and the Encyclopedia. The primitive elements in Distributed Morphology that syntax manipulates come in two types. One type is roots, which are atomic unanalyzable elements; the other type is functional heads such as *n/v/a* as well as other ordinally postulated heads such as Asp, Tense, C, Num, etc (*l-morphemes* and *f-morphemes* in the sense of Harley and Noyer 1999, 2000 and *roots* and *abstract morphemes* in Embick and Noyer 2007). Roots are considered acategorial; their syntactic category is contextually specified by combining with category-defining functional heads such as *v*, *n*, and *a*. For example, the root $\sqrt{\text{destr}}$ is realized as the noun *destruction* under the nominalizing environment ($[_{nP} n [\sqrt{\text{destr}}]]$); it is realized as the verb *destroy* the verbalizing environment ($[_{vP} v [\sqrt{\text{destr}}]]$); it is realized as the adjective *destructive* in the adjectivalizing environment ($[_{aP} a [\sqrt{\text{destr}}]]$) (Marantz 1997, in press; cf. Chomsky 1970, 2006). It is claimed within the Distributed Morphology that phonological features are assigned to abstract morphemes (as well as roots under certain views) post-syntactically. The mechanism to assign this feature is Vocabulary Insertion in the technical sense; “the Vocabulary is the list of the phonological exponents of the different abstract morphemes of the language, paired with conditions on insertion. Each such pairing of a phonological exponent with information about the grammatical (i.e. syntactic and morphological) context in which the exponent is inserted is called a vocabulary item.”

(Embick and Noyer 2007: 297). A list of potential candidates compete for the same abstract morpheme position in the morphosyntactic derivation; which candidate is selected as the exponent of that position is determined by the Subset Principle proposed by Halle (1997); see Embick and Noyer (2007: 298, 299) for concise illustrations of how this principle serves to constrain vocabulary insertion; also recall our analysis of *meN-/ng-* deletion in Indonesian/Javanese as failure of phonological insertion presented in chapter 2. The Encyclopedia, whose role is not directly relevant to the present chapter (though see section 5), is defined as a component that “lists the special meanings of particular roots, relative to the syntactic context of the roots, within local domains” (Marantz 1997: 204). Thus, the encyclopedic entry for the verb *kick* specifies that it means ‘kill’ in the environment of “___ the bucket” (Harley and Noyer 1999: 4).

Three ingredients of this theory of the syntax-lexicon interface that will become important in the following sections are as follows. First, the framework of Distributed Morphology claims that all types of word formation, including those that would be treated in the pre-syntactic lexical component in the lexicalist theory, are conducted within the sole realm of the syntactic computation in much the same way as sentences and phrases are. For this reason, there is no sense in which lexical processes must precede syntactic processes because there is no pre-syntactic generative component in the lexicalist sense in the first place; rather,

cases in which “lexical” processes in the lexicalist sense follow or are interleaved with syntactic processes are predicted to be possible.

Second, this framework claims that phonological features for a morphosyntactic head are inserted post-syntactically with reference to syntactic environments that surround the head, as illustrated above in the *destroy/destruction/destructive* alternation. This second claim is a special case of the more general claim, entailed by the thesis of the minimalist interface, that the syntax-external phonological component is interpretive, assigning a language-particular interpretation to the output of the syntax in the way that syntax demands. Accordingly, the thesis does not go well with the lexicalist theory, under which a word constructed in the lexicon is inserted into a terminal node in syntax as a triplet of sound, meaning and their correspondence. In this sense, Distributed Morphology provides a theoretical model that meshes nicely with the expectation of the minimalist interface defended in this thesis.

Finally, in contrast to the lexicalist theory of the lexicon-syntax interface, the theory of Distributed Morphology claims that the syntax-lexicon interface is direct. To borrow the phrase from Embick and Noyer (2007: 302) , “there is no syntax/morphology ‘interface’ because “words and phrases are assembled by the same generative system, and there is thus no sense in which words must ‘interface’ with the syntax; rather they are derived by the rules of syntax.” This view of the syntax-lexicon interface, thus, allows us to derive the effects of the

so-called Mirror Principle of Baker (1985), which stipulates a condition on how syntactic and morphological structures relate to one another, as the automatic consequence of the architecture of the theory.

In the rest of this chapter, I show that the non-lexicalist theory of Distributed Morphology provides a natural explanation of certain facts concerning reduplication in Indonesian that pose difficulties for several variants of the lexicalist hypothesis. This result, thus, constitutes evidence for the interpretive nature of the syntax-external linguistic interface, as entailed by the thesis of minimalist interfaces.

3. Asymmetries between Nominal and Verbal Reduplication in Indonesian

To find out existing patterns in nominal and verbal reduplication in Indonesian, Sato and McDonnell (in press) conducted a corpus survey of four popular newspapers published in Indonesian. The present corpus survey contains approximately 160, 000 words, taken from the archives of the following four newspapers: *Tempo*interaktif (www.tempointeraktif.com), *Suarapembaruan* (www.suarapembaruan.com), *Mediaindo* (www.mediaindo.co.id), and *Kompas* (www.kompas.com). The result of this survey is shown in Table 5. I have included here the results concerning derivational affixes; see Sato and McDonnell (in press) for results that cover inflectional affixes.

Table 5: The Corpus Survey of Four Popular Newspapers in Indonesia (approx.160, 000 words)

		Stem Reduplication		Stem-Affix Reduplication	
		Total Tokens	Unique Forms	Total Tokens	Unique Forms
No Affix		1014	312	N/A	N/A
Verbal Affixes	<i>ber-</i>	89	37	0	0
	<i>meN-</i>	30	23	0	0
	<i>di-</i>	23	20	0	0
	<i>ter-</i>	13	9	0	0
Nominal Affixes	<i>-an</i>	32	22	19	15
	<i>peN-</i>	0	0	8	5
	<i>peN-an</i>	0	0	2	2
	<i>per-an</i>	6	2	9	6
	<i>ke-an</i>	1	1	10	8

The results given in Table 5 reveal that there is a curious asymmetry between nominal and verbal reduplication in Indonesian, which has escaped attention in the literature on the morphology of Indonesian. As is clear from this survey, derivational verbal affixes such as *ber-*, *meN-*, *di-* and *-an* allow only stem reduplication. By contrast, derivational nominal

affixes behave differently from verbal affixes in that they potentially allow both types of reduplication. Specifically, *peN*, *peN-an*, and *ke-an* have strong tendency to feed stem-only reduplication, whereas *-an* and *per-an* allow stem-only and stem-affix reduplication. As is true for the corpus studies in general, it is difficult to know what forms cannot be produced in Indonesian, though the study does provides an indication that the reduplication asymmetry is real.⁵ To address this concern, I have conducted grammaticality judgment tasks with one informant to confirm whether the forms not found in the corpus study are actually unacceptable to the native language speaker of Indonesian.

The following examples show that the corpus study in Table 5 reflects the grammatical intuition of the actual speaker. Consider the reduplication pattern found in the verbal prefix *ber-*. Table 5 above indicates that this prefix only allows stem-reduplication. This result is confirmed by the contrast in acceptability between (10a-c) and (11a-c).

(10) Stem Reduplication with the Derivational Verbal Prefix *ber-*

- a. *belit* ‘twist’ \Rightarrow [*ber* [*belit-belit*]] ‘meander’
- b. *cakap* ‘talk’ \Rightarrow [*ber* [*cakap-cakap*]] ‘chat’
- c. *jalan* ‘walk’ \Rightarrow [*ber* [*jalan-jalan*]] ‘stroll’

⁵ As pointed out by an anonymous *Morphology* reviewer.

(11) Stem-Affix Reduplication with the Derivational Verbal Prefix *ber-*

- a. *belit* ‘twist’ \Rightarrow *[[*ber-belit*]-[*ber-belit*]] ‘meander’
- b. *cakap* ‘talk’ \Rightarrow *[[*ber-cakap*]-[*ber-cakap*]] ‘talk’
- c. *jalan* ‘walk’ \Rightarrow *[[*ber-jalan*]-[*ber-jalan*]] ‘stroll’

(10a-c) show that the prefix *ber-* allows stem reduplication. (11a-c) show that stem-affix reduplication is unacceptable for this prefix. This contrast, therefore, shows that the results given in Table 5 are real. A similar argument can be made for the observation made in Table 5 that derivational nominal affixes allow both stem and stem-affix reduplication. To take the suffix *-an*, we have seen above that this suffix allows the two types of reduplication. That this is correct is evidenced by the grammaticality of both examples in (12a-c) and (13a-c) below.

(12) Stem Reduplication with the Derivational Nominal Suffix *-an*

- a. *sayur* ‘vegetable’ \Rightarrow [[*sayur-sayur*]-*an*]] ‘many types of vegetables’
 \Rightarrow * [[*sayur-an*]-[*sayur-an*]]
- b. *buah* ‘fruit’ \Rightarrow [[*buah-buah*]-*an*]] ‘many types of fruit’
 \Rightarrow * [[*buah-an*]-[*buah-an*]]

- c. biji ‘seed’ ⇒ [[biji-biji]-an]] ‘many types of seeds’
 ⇒ *[[biji-an]-[biji-an]]

(13) Stem-Affix Reduplication with the Derivational Nominal Suffix *-an*

- a. pikir ‘think’ ⇒ [[pikir-an]-[pikir-an]] ‘thoughts’
 ⇒ *[[pikir-pikir]-an]]
- b. tulis ‘write’ ⇒ [[tulis-an]-[tulis-an]] ‘writings’
 ⇒ *[[tulis-tulis]-an]]
- c. masuk ‘enter’ ⇒ [[masuk-an]-[masuk-an]] ‘inputs’
 ⇒ *[[masuk-masuk]-an]]

(12a-c) show that the nominal suffix *-an* allows stem reduplication while (13a-c) show that the same suffix can also feed stem-affix reduplication. It is important to observe that the choice between the two forms of reduplication is not entirely free with this suffix; rather, the choice is affected by the type of stem that it is identified with. Thus, when this suffix is combined with nominal stems as in (12a-c), it only allows stem-reduplication. On the contrary, when this suffix is combined with verbal stems as in (13a-c), it only allows stem-affix reduplication. Thus, it is not the case that a single nominal affix allows both types of reduplication; it allows

both types when an appropriate class of elements combines with a particular nominal suffix. We can observe this effect in the behavior of circumfixes such as *peN-*, *peN-an*, and *ke-an*, whose dominant reduplication pattern is stem reduplication, as shown in the results reported in Table 5. This point will become very important in section 5.

4. Reduplication Asymmetries in Indonesian and Lexicalist Theories

The purpose of this section is to see whether the lexicalist theory might be able to accommodate the existing reduplication patterns in Indonesian. I show that the nominal vs. verbal reduplication asymmetry and the existence of a word-internal reduplication pattern that targets the non-edge of a complex stem cannot be accounted for by several versions of the lexicalist theory as in Chomsky (1970), Anderson (1982, 1992), Kiparsky (1982a, b, c, 1985), Mohanan (1986), and Di Sciullo and Williams (1987)/Williams (2007).

4.1. Chomsky's (1970) Weak Lexicalist Hypothesis

Chomsky (1970) proposes, based on his examination of several syntactic and semantic contrasts between derived nominalization as in *destroy* \Rightarrow *destruction* and gerundive transformations (*destroy* \Rightarrow *destroying*), that non-productive, irregular processes take place in the pre-syntactic lexical component while productive, regular processes take place in the

syntactic/transformational component. This separation of two types of complex word formation in terms of their regularity/productivity has been widely taken in the generative literature to define the classical version of the weak lexicalist theory (though, recall our earlier discussion of Marantz 1997). For example, Fabb (1984) considers productivity as the defining criteria with which to distinguish lexical and syntactic word formation.

If we adopt Chomsky's version of the lexicalist hypothesis, *ber-/an-* affixation as observed in examples like (10-13) counts more as a lexical/pre-syntactic process for the following reasons. First, the literature on the morphology of Indonesian as in McDonald (1967), Sneddon (1996), and others points out that the verbal prefix *ber-* may attach to nominal, numeral, and verbal bases that yield unpredictable/irregular semantic outcomes. First, predicates consisting of this prefix and a nominal base refer to a customary possession of, or to characterization by the referent of the noun, as shown in (14a, b). This type of prefixed predicate can also be used to refer to the act of producing the reference of the noun or making use of it, as shown in (14c, d). If the nominal base refers to a profession or way of life of an animate being, the derived predicate refers to the property of making a living with that possession or by that way of life, as shown in (14e, f).

(14) *ber*-prefixation: Input = Noun/Output= Verb

- a. anak ‘child’ → [ber [anak]] ‘have children’
- b. kaki ‘foot’ → [ber [kaki]] ‘have feet’
- c. kokok ‘cackle’ → [ber [kokok]] ‘produce a cackle’
- d. sepeda ‘bicycle’ → [ber [sepeda]] ‘use a bicycle’
- e. kuli ‘coolie’ → [ber [kuli]] ‘work as a coolie’
- f. tukang ‘artisan’ → [ber [tukang]] ‘work as an artisan’

(MacDonald 1967: 44, 45)

Second, the prefix can combine with a numeral, unreduplicated or reduplicated, to derive the complex noun meaning ‘forming a group of’ and ‘in groups of’, as shown in (15a-c).

(15) *ber*-prefixation: Input = Numeral/Output= Numeral

- a. dua ‘two’ → [ber [dua]] ‘two together’
- b. ratus ‘hundred’ → [ber [ratus]] ‘in hundreds’
- c. karung ‘sack’ → [ber [karung]] ‘in sackfuls’ (MacDonald 1967: 47)

Finally, the prefix may attach to verbal bases to create intransitive verbs; it works as a verbalizer of bases that otherwise do not occur alone, as shown in (16a, b). If the root is reduplicated, an additional meaning of variety, repetition or lack of purpose is added, as in (16c, d).

(16) *ber*-prefixation: Input = Verb/Output= Verb

- | | | | | | |
|----|---------|---------|---|---------------------|------------------|
| a. | -henti- | ‘stop’ | → | [ber [henti]] | ‘come to a stop’ |
| b. | -pikir- | ‘think’ | → | [ber [pikir]] | ‘be cogitating’ |
| c. | belit | ‘twist’ | → | [ber [belit-belit]] | ‘meander’ |
| d. | cakap | ‘talk’ | → | [ber [cakap-cakap]] | ‘have a chat’ |

(MacDonald 1967: 47,48)

The function of the derivational nominal suffix *-an* is no more or less complex. It has the role of nominalization when it attaches to verbal bases. It serves as a kind of classifier, meaning ‘types of’, as reflected in the English translations in examples in (12a-c) when it attaches to nominal bases. These semantic considerations, therefore, suggest that the two affixes constitute an irregular process and that the affixation involved is a lexical/pre-syntactic process in Chomsky’s sense. In section 5, however, I show that the suffix *-an* is polysemous,

but with its dual functions being determined by different attachment sites at which it merges with two types of complements (NumP vs. *v*P).

By contrast, reduplication in Indonesian is a fully productive, hence syntactic process under Chomsky's proposed productivity-based division of two types of word formation. Reduplication of any countable noun produces a grammatical form that is specifically plural; recall our observation made in chapter 4 based on Carson (2000) and Chung (2000). Thus, reduplication in Indonesian is a productive realization of the Num head in the nominal domain. On the other hand, it is not apparently as clear whether the corresponding argument can be made for the verbal domain to show that verbal reduplication is really a productive process. However, the following two considerations show that it is more like a syntactic process rather than a lexical process in the lexicalist sense. First, the literature on the verbal reduplication in Indonesian as in MacDonald (1976) and Sneddon (1996) notes that reduplication of a verb adds emphasis to an action denoted by the base stem and yields outcomes related to variety, multiplicity, and atelicity. Sneddon (1996) gives a variety of meanings as follows:

(17) The Semantic Effects of Verbal Reduplication

- a. With some verbs reduplication gives a connotation of action done in a causal or leisurely way .

Examples: *duduk* 'sit' \Rightarrow *duduk-duduk* 'sit about'

berjalan 'walk' \Rightarrow *berjalan-jalan* 'walk about, go for a stroll'

- b. With many verbs reduplication indicates continued action, either an action done over a period of time or an action performed repeatedly

Example: Bu Yem mengurut-urut rambut anaknya.

Mrs Yem stroked-Red hair child-her

'Mrs.Yem stroked her child's hair.'

- c. With some verbs reduplication gives a meaning somewhat different from that of the single form, usually conveying a sense of intensity.

Examples: *menjadi* 'become' \Rightarrow *menjadi-jadi* 'get worse'

meminta 'request' \Rightarrow *meminta-minta* 'beg'

- d. Accompanied by *tidak* ‘not’ reduplication of the verb can indicate that the action has not occurred, usually implying that this is contrary to expectation.

Example: Sudah dua hari Pak Tanto tidak muncul-muncul.

yet two day Mr Tanto Neg turn up-Red

‘Mr Tanto has not turned up for two days now.’

(Sneddon 1996: 20)

(16a-c) all belong to the type (17a) in Sneddon’s classification. This semantic effect as well as the other three in (16b-d) indeed are in keeping with the general notion of *plurality/emphasized quantity*, a crosslinguistically attested effect of reduplication, as evidenced by the extensive investigation of the function of reduplication conducted by Moravcsik (1978). Though Moravcsik herself concludes (p. 325) that “no explanatory or predictive generalization about the meanings of reduplicative constructions can be proposed,” as Travis (1999, 2003) argues, her extensive cross-linguistic investigation of the functions of reduplication across languages suggests that reduplication has some abstract quantificational function which is diversely instantiated as plural, distributivity, multiple iterative event readings, reciprocals, emphasis, and so on. This argument, thus, suggests that reduplication in Indonesian can be regarded as a syntactic process that relates to quantification. Travis (1999, 2003) also argues, while drawing a

clear separation between phonological and syntactic reduplication, that all types of reduplication are underlyingly caused by their surrounding syntactic configurations.

The second argument to support the syntactic nature of the reduplication in Indonesian comes from work on Madurese reduplication made by Davies (1999, 2000). Davies shows that reduplication in this language forces the multiple event reading of a verb based on his examination of reduplicative constructions in Madurese. There seems to be a general agreement in the lexicalist literature, at least tacitly, that the lexicon creates complex words based solely on lexical categories (N, V, A) but never on functional categories (T, C). This assumption is natural because time or event reference must make crucial reference to the rules of sentence formation. The following examples from Indonesian, modeled after the corresponding examples in Madurese provided by Davies (2000: 127-129), show that reduplication of a verb in Indonesian creates a variety of new interpretations unavailable to its unreduplicated counterpart, such as multiple event readings, interleaved activity readings, and temporally displaced readings.

(18) Reduplication in Indonesian

a. Esti meng-elus(-elus) rambut anak-nya.

Esti AV-stroke-Red hair child-her

‘Esti stroked her child’s hair many times.’

- b. Aini dan Lina me-motong(-motong) kayu selama dua jam dan menanam bibit

Aini and Lina AV-cut-Red wood for two hours and plant seed

‘Aini and Lina cut down trees for two hours and planted seeds.’

- c. Aini dan Lina men-cubit (*-cubit) adik-nya yang lucu. Aini mem-cubit-nya hari Senin

Aini and Lina AV-pinch-Red child-their that cute Aini AV-pinch-her day Monday

Lina hari selasa.

Lina day Tuesday

‘Aini and Lina pinched their cute baby. Aini did so on Monday and Lina did so on Tuesday.’

The example in (18a) illustrates the multiple event reading whereby the telic event of stroking a child’s hair occurred several times. If reduplication does not occur, by contrast, the sentence is ambiguous between the single event reading and the multiple event reading, a pattern that we have also seen to characterize nominal reduplication in chapter 4 of this dissertation. This event-related property caused by reduplication can also be seen in the example in (18b). Although judgments are subtle, according to my two language informants, the example in (18b) with reduplication allows the interpretation where the event of tree-cutting is interspersed with the event of seed-planting; for example, this sentence is true in the situation where Aini and Lina continued the activity of tree cutting for one hour, then did seed-planting for some time,

and then resumed the tree-cutting activity for another hour. This interspersed activity reading is impossible without reduplication of the verb in (18b). Similarly, the acceptability of the example in (18c) shows that the activity of the reduplicated verb can be spaced over time. For example, (18c) is acceptable with reduplication under the reading where Aini pinched her baby on Monday but Lina did so on Tuesday. The acceptability of this example with reduplication is what we predict precisely because the reduplication of a verb feeds multiple event readings. This reading, however, is unacceptable without verbal reduplication, as shown in (18c).

What is important about these examples is that the availability of these three readings, derived by verbal reduplication, makes crucial reference to the notion of time or event. Again, this reference should not be possible in the lexical component to the extent that the implicit but natural assumption holds, namely, that the lexicalist sense of lexicon does not contain functional elements such as T and C. The readings forced by reduplication in Indonesian as in (18a-c), thus, provide an independent argument for treating Indonesian reduplication as a syntactic/non-lexical process.

With the two facts noted above in mind, consider now whether the examples of stem-reduplication and the nominal vs. verbal reduplication asymmetry might be accounted for under Chomsky's theory. Examples of stem-reduplication as illustrated in (10a-c) and (12a-c) instantiate the word-internal reduplication, namely, that an affix (either *ber-* or *-an*) is attached to the complex

stem created by reduplication. In other words, the affixation applies word-internally. This pattern of reduplication poses an inverse ordering problem for Chomsky's version. The formation of stem-reduplicated forms such as *belit-belit* and *sayur-sayur* requires the syntactic process of reduplication because reduplication is a productive process. The *ber-/an* affixation applies to the stem-reduplicated form to yield the grammatical forms such as [[ber-[belit-belit]] and [[sayur-sayur]-an]]. This ordering should be impossible, however, under the lexicalist architecture of the lexicon-syntax interface because the generation of these forms requires that the syntactic process of reduplication be followed by the lexical/pre-syntactic process of affixation. Furthermore, it seems that Chomsky's variant of the weak lexicalist hypothesis does not have anything to say about why there is an asymmetry between nominal and verbal reduplication in Indonesian, as illustrated in the examples in (10-13) and Table 5, where nouns allow both stem and stem-affixation whereas verbs only allow stem reduplication. Based on these considerations, I conclude that Chomsky's (1970) version of the weak lexicalist theory have non-trivial architectural and empirical problems in face of the existing reduplication patterns in Indonesian.

4.2. Anderson's (1982, 1992) Weak Lexicalist Theory

Anderson (1982, 1992) develops a different version of the weak lexicalist theory from Chomsky's that does not depend on the somewhat vague notion of productivity. He argues that

inflectional morphology is treated in the syntax whereas derivational morphology is treated in the lexicon. He defines the inflectional/ syntactic nature of any element as follows:

(19) The Definition of Inflectional Morphology in Anderson (1982, 1992)

Inflectional morphology is what is relevant to syntax. (Anderson 1982: 587)

This definition, thus, allows any affixation that has relevance to syntax such as agreement, tense, event structure to be treated in the syntactic component.⁶ This conception of the weak lexicalist theory is particularly problematic in face of the Indonesian reduplication facts. The affixation of *ber-* counts as a lexical process because it does not seem to have syntactic effects such as agreement, tense, and event structure. However, we have seen in section 3 that reduplication in Indonesian has clearly syntactic outcomes in the form of multiple event readings and discontinuous time-interval reading.

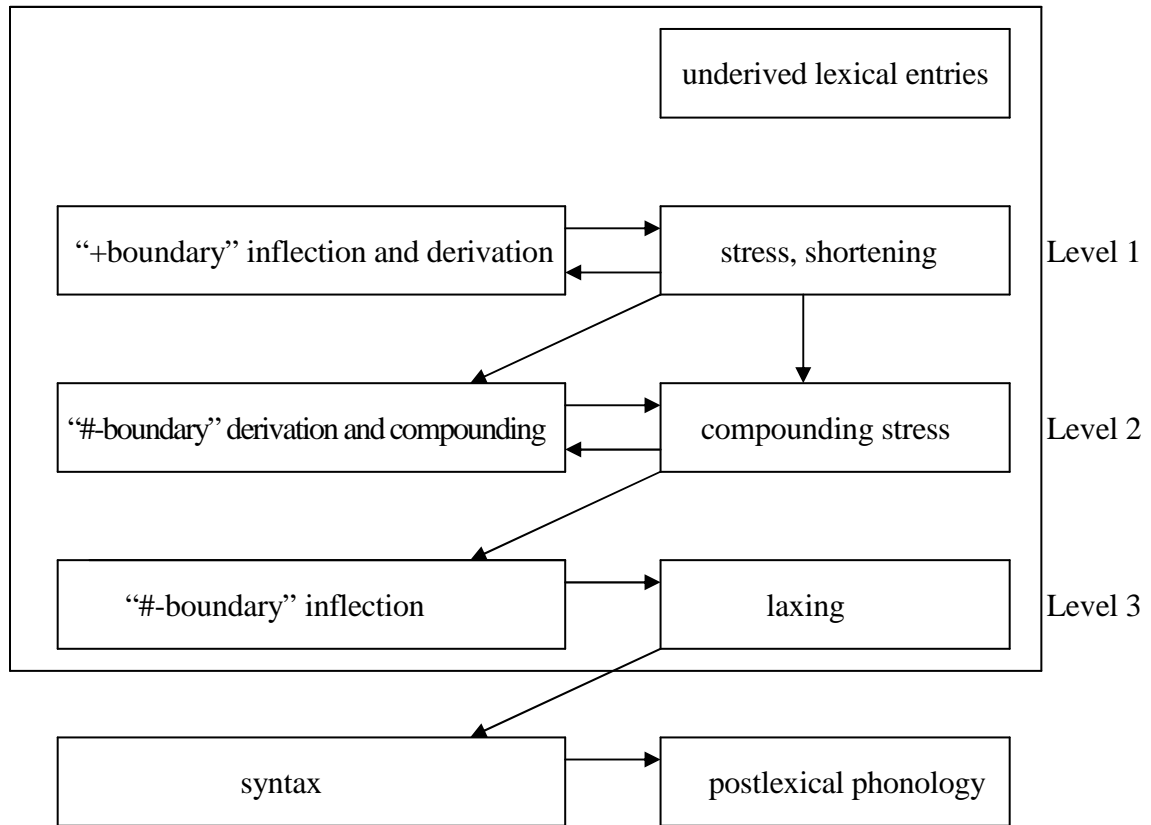
⁶As Di Sciullo and Williams (1987) point out, however, this definition does not work in the way Anderson wants it. If the inflectional nature of a particular affix is determined by its relevance to syntax, then almost all affixes could be considered as syntactic, a state of affairs that Anderson does not want. For example, they note (p. 69) that “this definition would seem to consign all nominalizing affixes, such as *-ion*, to inflection because nouns and verbs have different syntactic properties, and the affix makes the difference.” Similarly, *ber-*suffixation should be treated as syntactic, contrary to what Anderson’s weak lexicalist theory would actually intend, because it creates a related verb from the verbal stem, and the notion of verb is relevant to syntax. The problem is more general. For the purposes of this paper, I maintain the pretense that Anderson’s theory would treat *ber-*affixation as lexical in the intended sense.

This means that reduplication is an inflectional process to be treated in the syntax. Then, the word-internal reduplication pattern illustrated in (10a-c) and (12a-c) should be ungrammatical because the generation of such a pattern requires the application of the syntactic rule to be followed by the application of the lexical rule. Anderson's version of the weak lexicalist theory also has little to say about why the reduplication is asymmetric between nominal and verbal suffixes in Indonesian.

4.2. Kiparsky's (1982a, b, c, 1985)/Mohanan's (1986) Lexical Phonology

The same reduplication asymmetry and the word-internal reduplication pattern also refute one well-known version of the strong lexicalist theory known as Lexical Phonology (Kiparsky 1982a, b, c, 1985; Mohanan 1986; see also Pesetsky 1979). This theory maintains that morphology and phonology interact in tandem, with each stratum/cycle governing operations with certain characteristics. Specifically, affixational/inflectional processes with irregular phonological and morphological consequences occur in Stratum 1 while regular inflectional processes with transparent consequences occur in a later Stratum (Stratum 3 in Kiparsky/Stratum 4 in Mohanan). Kiparsky's (1982a) model of the Lexical Phonology is given in (20). See also Mohanan (1986) for a further development of Kiparsky's original model, which I am not going to discuss here.

(20) Kiparsky's (1982a) Model of Lexical Phonology in English (Kiparsky 1982a: 133)



This model assumes that the word formation rules and the lexical phonological rules are partitioned into an ordered series of levels/strata/cycles. “+boundary” inflectional affixes in Level 1 include the umlaut of *tooth-teeth*, the ablaut of *sing-sang* and other stem-changing morphology whereas “+boundary” derivational affixes includes what have been called Level 1 affixes in the Level-Ordering Hypothesis of Siegel (1973) and Allen (1978) such as *-al*, *-ous*, and *-im*, as in *refusal*, *pious*, and *impotent*. “#-boundary” derivation in Level 2 involves what have been called

Level 2 affixes in the Level-Ordering Hypothesis such as *un-*, *-ness*, and *-er*; also compounding, a process of combining two independent root elements such as *black board*, *nurse shoes*, and *red coat*. Finally, “#-boundary” inflection in Level 3 deals with the affixation involving the rest of the regular inflectional affixes such as plural *-s*, and past tense *-ed* in English. To illustrate, consider the derivation of *codifiers*. The base stem *code* is submitted to the phonological rules of Level 1, where the word formation rule attaches the Level 1 affix *-ify* to the stem. This derived stem is then assigned stress as *códifý* in the same Level. The resulting object is submitted now to the phonological component of Level 2, in which the word formation rule attaches the agentive suffix *-er* to derive the complex form *codifier*. Finally, when the resulting object enters Level 3, the regular plural formation process applies to this object to yield the final output *codifiers*. In this way, a set of phonological and morphological processes that apply to complex word formation is ordered in a series of strata/cycles. Kiparsky assumes that the derivation of all words should go through all these three Levels, even though relevant phonological and morphological processes might apply vacuously to a given form at any of these three levels.

This ordered block of rule application correctly predicts why forms such as **un[ept]* is ill-formed in contrast to *in[ept]*. As we have seen above, the prefix *un-* is a Level 2 affix. Thus, *un-* prefixation occurs in Level 2. To create the form **unept*, however, the bound morpheme *-ept-* must traverse the word formation process in Level 1 that would

assign the appropriate Level 1 affix *in-* to the stem to create the grammatical form *inept*.

The ill-formedness of **unept*, thus follows, as a natural consequence of the ordered series of morphophonological rule application.

One theoretical tenet of Lexical Phonology which is important for the purposes of this chapter lies in the *Bracketing Erasure Convention* (Pesetsky 1979). This convention deletes all brackets at the end of each stratum/level of word formation, thus has the effect of rendering access to the previously available internal structure of complex words opaque in later strata/cycles. This convention, thus, derives a lexicon-internal version of the Lexical Integrity Hypothesis defined in section 2, namely, that word formation processes in Level 2 and 3 cannot look into the morphological makeup of complex morphological objects created by word formation processes in Level 1 and Level 2, respectively. Lexical Phonology, therefore, makes an explicit prediction that no processes in a particular level should be able to apply to a complex object that is derived by word formation processes characteristic of earlier levels. This prediction is clearly falsified by the reduplication pattern attested in Indonesian. We have seen in section 2.1 that reduplication is a fully productive process. Under Kiparsky's model, this process should be located in Level 3 on a par with the regular inflectional affixes such as plural *-s*, and past tense *-ed*: recall that any countable noun and verb can be input for reduplication just as any countable noun and verbs can be affixed by *-s* and *-ed* in English, respectively

(except irregular nouns and verbs, of course). We have also seen in the same section that affixes such as *ber-* and *-an* yield a set of semantic irregularities when attached to a stem. This unpredictable behavior leaves affixation of these pieces in Level 1 on a par with the irregular umlaut and ablaut rules as in *brother-brethren*. Now, to derive the word-internal reduplication pattern as illustrated in (10a-c) and (12a-c) under Kiparsky's model, the Level 1 affixation (*ber-*affixation and *-an* suffixation) must be preceded by the Level 2 inflectional process (reduplication), an ordering that should be impossible in Lexical Phonology due to its central hypothesis that each level/stratum is strictly ordered, hence cannot be traversed. To illustrate the issue at hand with *ber-belit-belit*, the base *belit* is submitted to Level 1, at which the *ber-*prefixation would apply to yield [*ber-belit*]. This complex object is submitted to Level 3, at which reduplication applies to the whole object to create the output [[*ber-belit*]-[*ber-belit*]]. Importantly, this output is ill-formed in Indonesian even though this is the only output that is predicted to be possible under the strict layering of levels in Lexical Phonology.

This type of word-internal reduplication pattern is also problematic for Lexical Phonology in three other respects. First, due to the Bracketing Erasure Convention, Kiparsky's model makes a prediction that reduplication should must target the right or left edge of the whole complex object because at the time this process applies in Level 3, the input transferred from Level 2 enters the Level 3 as an atomic unanalyzable element as the result of the erasure of all

word-internal constituent boundaries. Thus, the existence of forms such as [ber-[belit-belit]] shows that reduplication targets part of the complex stem rather than the left or right edge of it. Second, Kiparsky assumes that the output of each level is itself a full-fledged lexical item. However, the ill-formedness of forms such as *belit-belit shows that this is not always the case. Finally, as in Chomsky's (1970) weak lexicalist theory, Kiparsky's theory also does not seem to provide us with any way of explaining why the asymmetry between nominal and verbal reduplication observed in section 2 obtains in Indonesian.

4.3. *Di Sciullo and Williams' (1987)/Williams' (2007) Strong Lexicalist Theory*

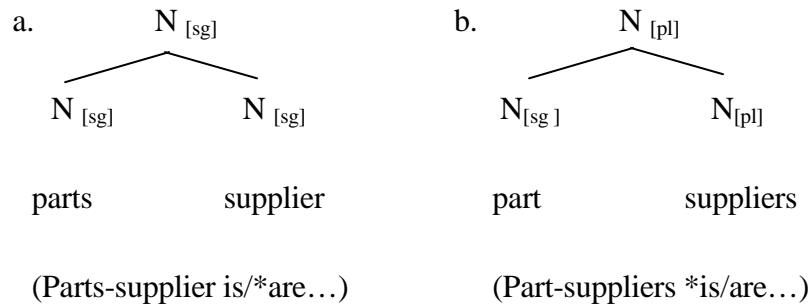
Di Sciullo and Williams (1987) develop the most comprehensive defense of the strong lexicalist theory, which is conceptually very different from other instantiations of the theory as in Kiparsky (1982a, b, c, 1985) and Mohanan (1985). They maintain that morphology and syntax are two different domains of inquiry with two different primes (e.g., stems, affixes, roots vs. NPs, VPs, CPs) and operations (compounding, θ -identification vs. movement, quantification). Thus, for Di Sciullo and Williams, the so-called lexicalist hypothesis/the lexical integrity hypothesis/the lexical atomicity "is not a principle of grammar but rather a consequence of the conception that grammar contains two subparts, with different atoms and different rules of formation" (p.2). The word "lexicon" takes an entirely different sense in their framework from the most common

usage as the generative system that stores words and their formation rules: the lexicon is the storage house for *listemes*, “objects of no single specifiable type (words, VPs, morphemes, perhaps intonation patterns, and so on) that “fail to conform to interesting generalizations.” (p. 3)

Assuming this strict division of labor between the word system and the phrase system, Di Sciullo and Williams maintain that the morphology and syntax can still communicate with one another through a restricted range of shared vocabulary, specifically, the “topmost properties of words, the features and argument structure of the topmost words.” (p. 45), as stated in their proposed *Thesis of Atomicity of Words* defined as in (3). Williams (2007), the most recent update of Di Sciullo and Williams’s lexicalist hypothesis, maintains essentially the same position.

Di Sciullo and Williams illustrate this limited cross-modular communication with compounding in English. Compounding involves the creation of what they call *morphological objects* that derive their agreement features from the percolation of the features of the right-hand head (Williams 1981). Crucially, it is this output agreement recorded on the top-most level of the compound (namely, the topmost N in (21a, b)) that is used for the purposes of syntactic subject-verb agreement, as the contrast between (22a, b) shows.

(21) English N + V compounds



(22) a. Parts-supplier is/*are mean to me.

b. Part-suppliers *is/are mean to me.

This agreement pattern correctly falls out from Di Sciullo and William's system because the feature specification for the non-head member of the compound is invisible from the perspective of syntax. Thus, this pattern is one way in which the syntax and morphology can communicate in a restricted range of share vocabulary though the Thesis of Atomicity of Words still blocks the syntax from accessing the internal morphological composition of compounds.

At this point of the present research, it is not clear whether any aspect of Indonesian reduplications facts could prove Di Sciullo and William's (1987) version of the lexicalist theory untenable. For Di Sciullo and William's lexicalist theory and other variants of the

hyper-lexicalist approach as in HPSG, it is extremely difficult, if not impossible, to find certain morphological patterns that could tease apart the predictions of their theory and some other theories; see Haugen and Harley (2006), though, who provide a detailed discussion of how certain compound-internal reduplication patterns in Yaqui pose a problem for their lexicalist theory; see also Harley to appear for relevant discussion). I suspect, however, that Di Sciullo and Williams's hypothesis is not incorrect; when carried to its extreme, it boils down to the syntactic approach to reduplication to be proposed in the next section in a different module of grammar (that they call *the Word System*).

4.4. *The Lexicon as the Source of the Ordering Paradox*

To summarize I have shown that reduplication within lexically/pre-syntactically derived complex stems in Indonesian poses non-trivial empirical and architectural problems for a number of well-known versions of the weak and strong lexicalist theory as presented in Chomsky (1970), Anderson (1982, 1992), Kiparsky (1982a, b, c, 1985), Mohanan (1986), and Di Sciullo and Williams (1987)/Williams (2007). I have also shown that those lexical approaches would have little to say about how the asymmetry between nominal and verbal reduplication arises in this language. Thus, those facts on Indonesian reduplication provide strong arguments against certain versions of the weak/strong lexicalist theory.

It is important to point out that this type of inverse ordering is a problem only when we postulate the lexicon/morphology as the pre-syntactic generative component that is responsible for certain types of word formation characterized by productivity, semantic/phonological compositionality, the relevance of morphological primes to the syntax, and so on, as assumed in Chomsky (1970), Anderson (1982, 1992), and Kiparsky (1982a, b, c), and Mohanan (1986). In other words, this problem does not (or cannot) arise in non-lexicalist theories of the lexicon-syntax interface that do not posit such an independent component prior in addition to the generative system of syntax. In light of this consideration, in the next section, I pursue an alternative, non-lexicalist analysis of the reduplication in Indonesian within the more recent morphosyntactic framework of Distributed Morphology as reviewed in section 2.2.

5. A Distributed Morphology Approach to Reduplication Asymmetries in Indonesian

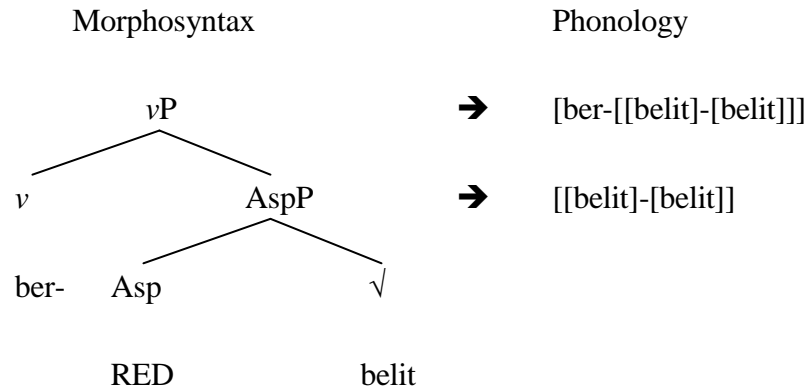
In this section, I show that the asymmetry between nominal and verbal reduplication and the word-internal reduplication pattern receive a straightforward account within the theory of Distributed Morphology (Halle and Marantz 1993, 1994; Harley and Noyer 1999; Embick and Noyer 2007). Specifically, I propose that these facts are explained as a natural consequence of a particular hierarchical arrangement of morphosyntactic features such as Aspect and Number in Indonesian. I assume, in line with much recent work on reduplication

in a number of different theoretical frameworks, that this process consists in affixation of the reduplicative null morpheme RED (UPLICATION) that triggers copying on a stem on its local environment; see Marantz (1982), McCarthy and Prince (1986, 1993, 1995); see also Travis (1999) and Haugen (2004) for a syntactic approach to reduplication. I argue that this particular analysis has an important bearing on the validity of the minimalist interface that has been defended and substantiated in the preceding chapters.

5.1. *Verbal Reduplication*

Consider first verbal reduplication. As we have seen in section 2, derivational verbal affixes can only allow stem reduplication. This pattern is naturally explained if verbal reduplication is mediated by the Inner Aspect head (Travis 1999) that dominates the reduplicative null morpheme and if *ber*—prefixation represents the addition of the higher *v* projection to the Inner Aspect structure. This assumption is supported by the fact that, as noted in section 2, verbal reduplication has various effects related to the Aktionsart of the predicate. Under this set of assumptions, then the morphosyntactic derivation for the example in (10a), [ber-[belit-belit]], will be as in (23).

(23) The Morphosyntactic Derivation of the Stem-Reduplication in (10a)



In this derivation, the Asp head merges with the acategorical root *belit* ‘twist’. The object that results from this merger is phonologically realized as the reduplicative form, $[[\text{belit}]-[\text{belit}]]$, because the only stem that the RED morpheme in the Asp head triggers copying of is the root *belit* on its local c-commanding environment. The Asp head undergoes further merger with the verbalizing prefix *ber-*. The complex morphosyntactic object, then, is interpreted at the syntax-external phonological component as $[\text{ber}-[[\text{belit}]-[\text{belit}]]]$, as desired.

It is important to note that the reduplicative morpheme intervenes between the ν head and the root in this derivation. Accordingly, the RED morpheme cannot reach up to the position of the ν head to include the verbalizing prefix in its domain for reduplication to yield the ungrammatical form as in $[[\text{ber-belit}]-[\text{ber-belit}]]$. This derivation, thus, correctly predicts the unavailability of the stem-affixation reduplication pattern for derivational

verbal affixes such as *ber-*. In this way, the fact that verb stems only allow stem reduplication naturally falls into place by assuming a particular hierarchical arrangement of morphosyntactic features/heads.

It is also to be stressed here that the state of affairs observed above in which the functional heads are linearized in the direction predicted by the hierarchical alignment of morphosyntactic features is exactly what is expected under the theory of Distributed Morphology. As we have seen, the word formation of all kinds is conducted by the single generative procedure as the sentence formation of any kind is in this non-lexicalist theory. Accordingly, the verbal reduplication pattern in Indonesian is simply the direct consequence of the grammatical architecture of the Distributed Morphology. On the contrary, under non-lexicalist views of the syntax-lexicon interface, there is no reason to expect that the syntactic structure and the morphological structure match in this manner, as the interface between the lexicon and syntax is indirect for reasons mentioned in section 2.1. Thus, the reduplication for verb stems in Indonesian can be construed as one good testing ground to tease apart the predictions of the two competing theories.

The proposed analysis of verbal reduplication in Indonesian also supports the locality of post-syntactic phonological feature assignment at the syntax-external interface. The proposed analysis crucially rests on the idea that the post-syntactic late insertion of phonological material

at the interface closely mirrors the way the syntactic derivation proceeds; *ber-* cannot be included as part of input for verbal reduplication because it is merged in a structurally higher position than the object (AspP) that becomes the target for reduplication. The root must be included for reduplication because it is in the c-commanding domain of the RED morpheme. By contrast, the stem-affix reduplication pattern as in [[*ber-belit*]-[*ber-belit*]] is simply underivable under the syntax-driven interpretive nature of the phonological component. The proposed analysis, thus, can be considered another particular instantiation of the theoretical position entailed by the thesis of minimalist interface that the syntax-external sound component is fundamentally interpretive, doing the best it can to satisfy the structural requirements imposed on it by the syntactic derivation. This point also holds true for the analysis of nominal reduplication presented in the next subsection.

5.2. *Nominal Reduplication*

Let us now turn to nominal reduplication. We have seen in section 2 that derivational nominal suffixes allow both stem and stem-affix reduplication. At the same time, we have also noted that the choice between the two types of reduplication is not entirely free but rather is governed by the syntactic category of the input stem. This latter point is crucial for the account presented below.

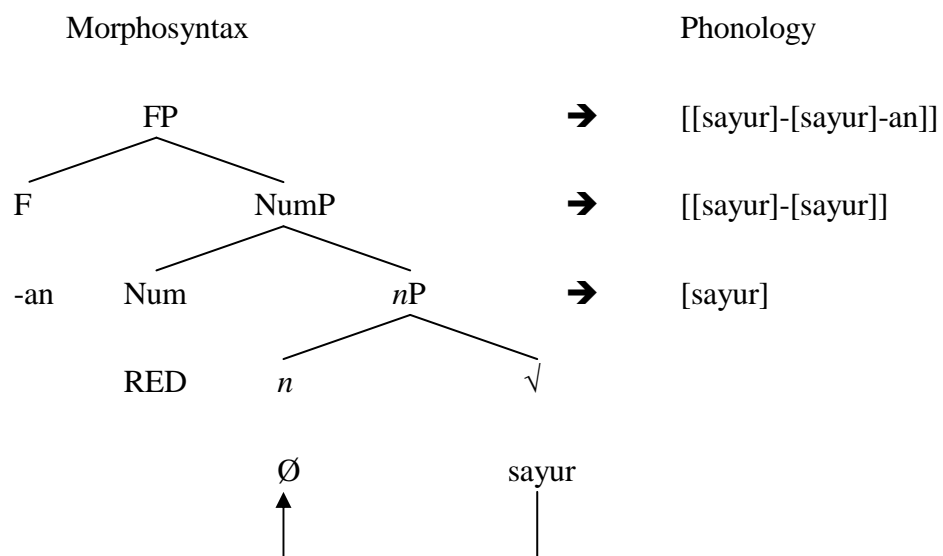
The input nominals in (24a-c) that allow only stem reduplication are all simplex nominals (i.e. *sayur* ‘vegetable’, *buah* ‘fruit’, and *biji* ‘seed’) whereas the input nominals in (25a-c) that allow only stem-affix reduplication are all complex deverbal nominals (i.e. *pikir* ‘think’ \Rightarrow *pikir-an* ‘thought’, *tulis* ‘write’ \Rightarrow *tulis-an* ‘writing’, and *masuk* ‘enter’ \Rightarrow *masuk-an* ‘input’). This difference, I claim, holds a key to a full understanding of why nominal derivational affixes in principle allow two types of reduplication unlike their verbal counterparts.

Let us assume that nominal reduplication consists in the copying of a nominal stem by the reduplicative null morpheme located in the Num head. The Num head selects a nominal stem as its complement, a rather natural assumption provided that reduplication of a nominal element yields the form that is specifically plural in Indonesian, as we have seen in section 3 and chapter 4 of this dissertation. This assumption also meshes nicely with the analysis of the denotation of bare nominals presented in chapter 4, under which bare uninflected nominals in Indonesian project up to the Num P that selects an NP as its complement. Alternatively, we might assume that a nominal stem is selected by the Q(uality) Phrase in light of the fact that one function of the nominal reduplication is to yield an emphasized quantity reading for the reference of the noun stem. For the

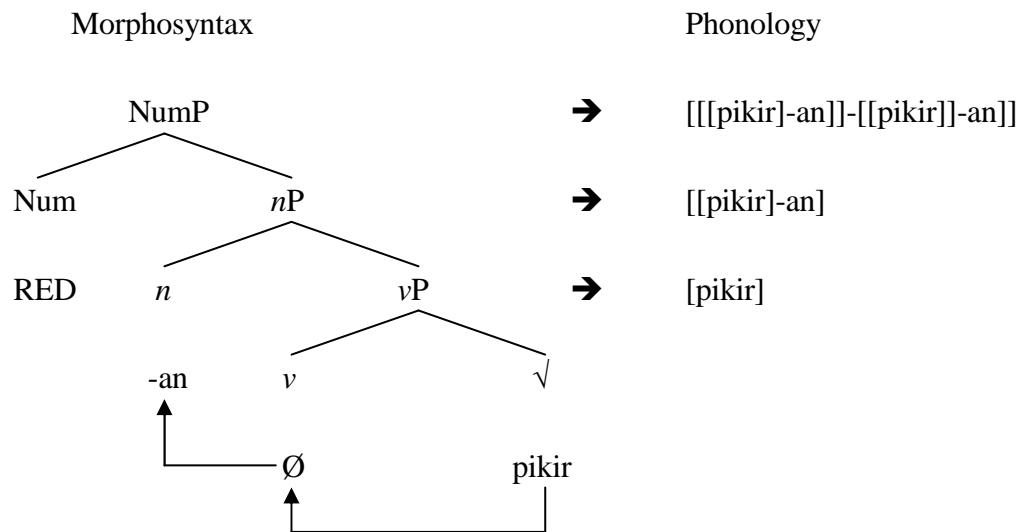
purposes of this chapter, however, I assume, following the analysis in chapter 4, that the RED morpheme is hosted by the Num head.

Under this analysis, then, simplex nominal stems as in (24a-c) can directly merge with the Num head. Verbal stems as in (25a-c), by contrast, cannot merge with the Num head this way because this head only selects a nominal stem as its complement, as stated in the previous paragraph. Thus, they are nominalized by the suffix *-an* before they can merge with the Num head. The morphosyntactic derivations for the examples in (24a) and (25a), then, will be as in (26) and (27), respectively. I designate the head that host the suffix *-an* in (26) as the F, returning to its precise nature shortly.

(26) The Morphosyntactic Derivation of the Stem-Reduplication in (24a)



(27) The Morphosyntactic Derivation of the Stem-Reduplication in (25a)



In the derivation in (26), the root *sayur* ‘vegetable’ is instantiated as a noun by adjoining to the null nominalizing head. This stem, being a nominal, can directly merge with the Num head as input for reduplication, which creates the form [[pikir]-an] at the syntax-external sound component. The derivation further continues by merging the *nP* with the Num head that hosts the RED morpheme. Since the RED morpheme can have access to the *nP* in its local c-commanding domain, the output of the *nP* must be included for reduplication at the Num head level. This information is realized at the post-syntactic phonological component as [[[pikir]-an]]-[[pikir]]-an]]. Note that the form [[pikir]-[pikir]-an]] cannot be created at the interface because the syntactic derivation here dictates that late insertion of phonological material at the interface must mirror the history of syntactic derivation in a cyclic manner, as

expected under the thesis of minimalist interface. The derivation in (27) is crucially different from that in (26), in that the base stem is verbal. Accordingly, the stem must undergo zero-derivation into nominal stems by the suffixation of the nominalizing suffix *-an* to serve as the complement that can satisfy the categorial restriction imposed by the Num head. Since the RED morpheme contained in this head includes the nominalizing suffix as well as the base stem in its local c-commanding domain, the syntactic derivation dictates that the phonological component include both elements as input for reduplication, thereby closely following the path curved by syntactic derivation in a local manner. Under this derivation, then, the stem reduplication pattern as in the hypothetical *[[[pikir]-[pikir]]-an] is simply underivable due to the way syntactic derivation proceeds and the way a particular set of morphosyntactic features is organized as shown in the derivation in (27). In this way, the proposed non-lexicalist, morphosyntactic analysis provides a straightforward explanation for the fact that the choice between the stem and stem-affix reduplication correlates with the underlying category of the input stem (verbal vs. nominal).

There is one important question that remains under the proposed account of nominal reduplication, which concerns the status of the suffix *-an*.⁷ As is clear from the position of this suffix in the two morphosyntactic derivations in (26) and (27), the proposed analysis entails

⁷ Thanks to Heidi Harley (personal communication) and an audience member at the 33rd Meeting of the Berkeley Linguistic Society for asking this question.

that this single suffix has two separate functions in the two cases. It is clear that the function of *-an* is that of nominalization in the stem-affix reduplication. The question is what role the suffix plays in the stem reduplication. It is important to recall in this connection that, in examples such as (10a-c), *-an* yields the reading that can be roughly “many types of”. Thus, *buah-buah-an*, derived from *buah* ‘fruit’, means ‘many types of fruits’. This observation is related to the remark made in MacDonald Dardjowidjojo (1967: 66), who observes that “when added to a noun root (usually reduplicated) the suffix *-an* forms a noun which refers to a collection of the referents of the simple noun, or of related referents,” as shown in (28a, b).

(28) The Function of *-an* in Stem Reduplication

- a. kaleng ‘tin can’ \Rightarrow [[kaleng]-[kaleng]-an] ‘canned goods’
 b. pohon ‘tree’ \Rightarrow [[pohon]-[pohon]-an] ‘trees, the vegetable kingdom’

(MacDonald Dardjowidjojo 1967: 66)

Interestingly, however, the same suffix cannot be attached to stems like *jeruk* ‘lemon’ to derive *jeruk-jeruk-an*; according to my native language consultants, the form itself is not ungrammatical but means something entirely different, something like “fake orange.” This is presumably related to the kind-denoting nature of bare nominals in Indonesian, as

argued extensively in chapter 4.⁸ Since the name of a kind refers to an amorphous, mass-like property of that kind under Carlson's (1977)/Chierchia's (1998a, b) sense of the term, it is natural that *-an* attached to a bare nominal yields readings such as "many types of X".

However, this line of thinking still leaves unexplained why bare nominals such as *jeruk* 'lemon' do not yield "many types of lemons"-readings when the stem is reduplicated and suffixed by *-an*. I maintain that the suffixability of *-an* is determined by the world-knowledge of what constitutes the most natural kind in classificatory terms in the mind of Indonesian speakers. The same language informants point out that the noun *buah* admits *-an* when reduplicated because it refers to what they would take to be a natural classificatory term; on the other hand, the noun *kejur* 'lemon' does not in their opinion because it is a specific instance of the classificatory term *buah* and does not presumably serve a natural classificatory function in the light of their encyclopedia knowledge. This observation is highly reminiscent of Chierchia's (1998a) discussion of the nature of a kind. For example, he notes (p. 348) makes the following remark, citing Carlson (1977: 26ff) and Krifika et al. (1995).

⁸ Thanks to an anonymous *Morphology* reviewer for suggesting this possibility.

(29) Chierchia's (1998a) discussion on the nature of a kind (Chierchia 1998a: 348)

By 'natural' kinds, we do not necessarily mean, in the present context, just biological ones or even 'well-established' ones. Artifacts (like chairs or cars) or complex things (like intelligent students or spots of ink) can qualify as kinds, to the extent that we can impute to them a sufficiently regular behavior...What counts as kind is not set by grammar, but by the shared knowledge of a community of speakers.

Note that this line of thought provides somewhat interesting support for the Distributed Morphology view of roots as assigned idiomatic meanings in a language-particular manner in the post-syntactic Encyclopedic component (see Harley 2008 for much relevant discussion on the nature of roots in Distributed Morphology from a case study in English compounds). The decision on whether a particular root such as *sayur* 'vegetable' constitutes a natural kind or not depends on the speaker's perception of whether the root makes a natural classification in his/her mind. The Distributed Morphology model predicts precisely this state of affair because the relevant decision is a matter of the encyclopedic component that deals with linguistically unpredictable meanings for expressions such as *cats*, *kick the bucket*, and so on. Therefore, I conclude that a) *-an* is an exponent of what native speakers of Indonesian take it to a natural kind in a classificatory sense and that b) it is generated in the head of

whatever projection that selects the phrase that denotes a kind. This conclusion is also independently supported by the morphosyntax of bare nominals in Indonesian discussed in chapter 4, where I have provided independent evidence that whatever value the Num head takes (specifically, be it plural or neutral), the denotation of the NumP is computed at the interface as a kind in languages such as Indonesian, Javanese, and Japanese. Thus, these considerations thus further support my conclusion that *-an* occupies the head that selects the NumP as its complement as in the derivation in (26).

6. Conclusions

Let us now take stock of what I have done so far in this chapter. I have started this chapter by introducing the results of my corpus study of four popular newspapers published in Indonesian. This study has revealed that a) nominal derivational affixes such as *-an* in principle allow both stem and stem-affixation reduplication whereas verbal derivational affixes such as *ber-* allow only stem reduplication and that b) both nominal and verbal stems may allow reduplication to target part of a morphologically/lexically derived complex word rather than its left or right edge. I have also shown that these results of the corpus study are indeed verified by native speakers' intuition by conducting grammaticality judgment tasks.

Then, I have demonstrated that these two facts concerning Indonesian reduplication pose non-trivial architectural and empirical challenges for a number of well-known versions of the weak and strong lexicalist theory as in Chomsky (1970), Anderson (1982, 1992), Kiparsky (1982a, b, c, 1985)/Mohanan (1986), and Di Sciullo and Williams (1987)/Williams (2007). I have also emphasized that the inverse ordering paradox caused by the word-internal reduplication pattern only arises in a theory of the lexicon-syntax interface that postulates the generative lexicon as an autonomous pre-syntactic component. Accordingly, the inverse ordering problem ceases to be a problem under non-lexicalist theories of the interface because we do not have any pre-syntactic word-formation component prior to the syntactic component in the first place.

Based on this theoretical consideration, I have argued that the two facts about Indonesian reduplication noted above receive a straightforward explanation within the more recent, non-lexicalist, morphosyntactic theory of Distributed Morphology outlined in Halle and Marantz (1993, 1994), Harley and Noyer (1999), and Embick and Noyer (2007) once we take seriously a particular hierarchical arrangement of certain morphosyntactic features/heads such as Asp and Num as well as the underlying syntactic category of input stems for reduplication. The key assumption of the proposed analysis is that the post-syntactic phonological feature assignment closely mirrors the bottom-up derivation of

morphosyntactic structures; the phonological component requires the reduplicative morpheme to target only the constituent within its c-commanding domain and the assignment of phonological feature applies in a bottom-up fashion, much in the way that syntax works. According to this analysis, the stem-affix reduplication in cases such as *[[[sayur]-an]-[[sayur]-an]]] and the stem reduplication in cases such as *[[[pikir]-[[pikir]-an]]] are simply underivable.

The overall result in this chapter, therefore, provides a strong piece of evidence against the traditional lexicalist architecture of the syntax-lexicon interface, and, at the same time, argues in favor of non-lexicalist theories as in the recent Distributed Morphology framework that attempt to locate all types of word formation within the sole realm of the syntactic derivation. In the next section, I show that a similar argument can be made against the lexicalist model of the syntax-lexicon interface from the phenomenon of what I term “phrasal inclusion within lexical words” in Indonesian and many other languages.

In the rest of this chapter, I briefly discuss further implications of the results achieved in this chapter for the proper theory of the syntax-lexicon interface and, most importantly, for the thesis of minimalist interface.

6.1. *Implications for the Proper Theory of the Lexicon-Syntax Interface*

Let me first note that minimalist considerations of parsimony, elegance, and simplicity, known as the *methodological minimalism*, “the drive for simple and nonredundant theories of the world” or “seeking the best way to theorize about a particular domain of inquiry” (Martin and Uriagereka 2001: 1), favor the non-lexicalist hypothesis in which the single generative syntactic component is responsible for all types of word formation rather than the lexicalist hypothesis in which a separate pre-syntactic component is postulated for certain types of word formation in addition to the syntactic component. That is, the non-lexicalist position is the null hypothesis in light of this formal consideration. In terms of the empirical question of which theory is right, as Embick and Noyer (1999: 291, 321, 322) note, the lexicalist claim that the generative lexical component be posited in addition to the syntax can only be motivated to the extent that a certain important generalization *cannot* be stated in a syntactic approach to word formation but instead *must* be treated in a lexicalist approach to word formation; the simple observation that a lexicalist approach works is simply irrelevant to the debate between lexicalism and non-lexicalism. Without this type of demonstration, the methodological parsimony noted above always tells us that the single-engine hypothesis of the non-lexicalist theory is the null hypothesis. A similar argument holds for the other

part of the lexicalist claim that ‘words’ are different from ‘phrases’. Again, the burden is on lexicalist theorists to show that certain patterns of “words” cannot be accommodated within the syntactic approach to word formation and that rules of word formation must be treated differently from rules of phrase formation.

Importantly, the results achieved in this chapter provide a new type of empirical evidence against the lexicalist theory from another perspective; the postulation of a pre-syntactic level such as lexicon actually would lead us to miss certain important generalizations that could be most naturally statable within the morphosyntactic structure. In this sense, the patterns of reduplication in Indonesian provide a unique empirical demonstration that the general architecture of the lexicalist theory is untenable inasmuch as it postulates a pre-syntactic lexical component prior to the syntactic component. In other words, the proper theory of the lexicon-syntax interface must be one of the variants of the grammatical models such as Distributed Morphology that posits a single level, namely, the syntactic component, for all types of word formation that has been independently shown to be a generative system in the framework of generative grammar. This point is to be kept in mind in the next chapter as well, where I provide another empirical argument of the same type against the lexicalist view of the lexicon-syntax interaction based on the phenomenon of “phrasal inclusion within lexically derived words” in Indonesian.

6.2. *Minimalist Interfaces*

The conclusion reached at the end of the previous subsection is quite in line with the idea of minimalist interface that the syntax-external components are purely interpretive, doing the best they can to assign interpretation to the output of the generative syntactic system in the manner required by the system. This chapter has shown that the thesis of minimalist interface is substantiated in three domains. First, we have seen that the late insertion model as in Distributed Morphology is what the minimalist would lead us to adopt, given that the phonological component is purely interpretive. Second, we have seen that the way that late phonological feature assignment works in the syntax-external interpretive component crucially depends on the structural notions such as locality and c-command that have been shown to be important theoretical ingredients in the generative syntactic research. Finally, the proposed non-lexicalist analysis within the Distributed Morphology framework leads to the theoretical consequence that there is no such thing as the syntax-lexicon interface, as also pointed out by Embick and Noyer (2007), because words are constructed in the same way as sentences and phrases are by the same set of rules and principles. This is an optimal consequence of the thesis of the minimalist interface, which views the role of syntax-external linguistic interface as autonomous systems that interpret whatever objects are sent out by the sole generative engine of syntax and add certain modifications to make them usable for the

language-external conceptual and articulatory systems within the range of options set up by syntax. In this way, close examination of the existing reduplication patterns in Indonesian conducted in this chapter provides further substantiation for the thesis of minimalist interface defined as above. I discuss further ramifications of this thesis in the next chapter.

CHAPTER 7 MINIMALIST INTERFACES

In this final chapter, I summarize the high points of the contents of my investigation in the previous chapters and identify potentially significant conclusions to be drawn from this dissertation on the interface between syntax and its neighboring phonological and semantic interfaces.

7.1. Summary of the Dissertation

The leading idea in this whole dissertation has been that of minimalist interfaces, namely, that the syntax-external linguistic components conduct a handful of modular-specific operations to arrange complex objects created by the universal law of syntax to be usable for language-independent articulatory and conceptual systems. All the apparently disparate phenomena drawn from the sizable portion of the grammars of Indonesian and Javanese discussed in detail in the preceding chapters have been shown to provide support for this notion of interface strategies. This idea has been instantiated in a variety of ways. The active voice deletion caused by the movement of an NP (chapter 2) is one clear case where the syntax-external phonological component closely mirrors the local computation of syntax as required by the Phase Theory. The denotation and morphosyntax of bare nominals in Indonesian and Javanese

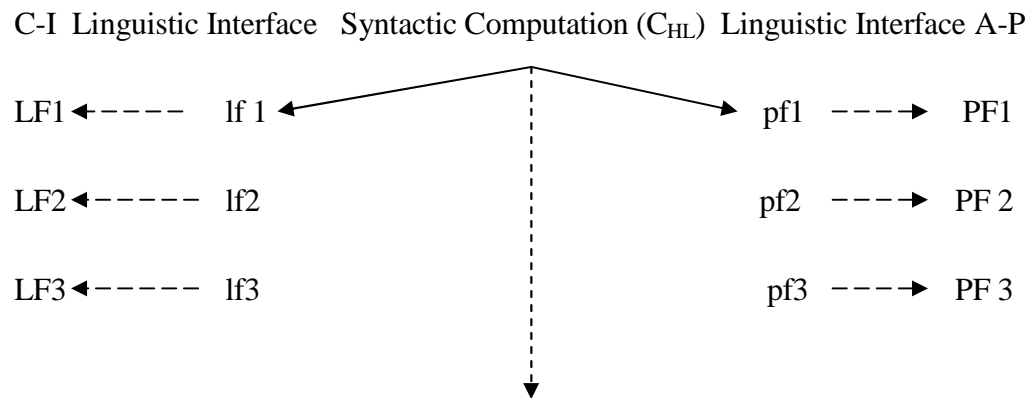
as well as several other languages (chapter 4) led to a relativized parametric theory of nominals that proposes that there is only one semantic interpretation that can be conducted at the semantic interface; what differs in the height of functional structures above the lexical projection and the possible set of values that the Num feature can take in each language. The cross-classification of these two parameters yields different outcomes at the semantic interface. This result, therefore, shows that what the interface can do is to accept whatever objects syntax sends off and give an interpretation to the objects in a way that is compatible with our encyclopedic/world knowledge of how nouns are conceptualized. Investigation of nominal vs. verbal reduplication (chapter 6) also suggests that the phonological component is fundamentally interpretive, doing whatever operations it can to mirror the way syntactic derivation proceeds, but within the range of options set up by syntax, such as c-command, locality, cyclic assignment of phonological features and so on. P-stranding under sluicing in Indonesian (chapter 3) constitutes one clear case where the role of the syntax-external phonological component is clearly at work. I have argued that certain violations created by syntax such as percolation failure can be repaired at the interface by deletion. At the same time, I have also shown that other violations such as the failure of D-to-P incorporation in syntax cannot be repaired at the interface. This analysis, if correct, is a natural outcome of the proposed thesis of minimalist interfaces: interfaces can conduct operations to legitimize

otherwise illegitimate syntactic objects but only within the range of options admitted by language-particular parameters. Similarly, the semantic component behaves as an autonomous component that applies choice functions to *wh*-in-situ questions in Indonesian (chapter 5). Particular details of the choice function analysis, I showed, provides further confirmation for the minimalist interface thesis defined above in that the semantic component does conduct the domain-specific operation of choice function but crucially because syntax is not fine-tuned to create a multitude of interpretations that would be required by the communicative demands from the language-independent conceptual system.

7.2. Minimalist Interfaces

The thesis of minimalist interface leads to a particular understanding of the way syntax interfaces with its neighboring linguistic components, semantics and phonology. Couched within the recent derivational theory of syntax and its correspondence with phonology and semantics (see Chomsky 1995, 2000, 2001, 2004, 2005, Epstein et al. 1998, Uriagereka 1999, and Grohmann 2003), the thesis of minimalist interface defined above yields the architecture of the interface between syntax, syntax-external interfaces, and language-independent AP and CI systems, as shown in Figure 1 from chapter 1.

Figure 1. The Architecture of the Interfaces under the Thesis of the Minimalist Interface



The small *lf* and *pf* should be read as the chunk of syntactic structures that are transferred to syntax-external phonological and semantic interfaces. The big *LF* and *PF* should read as whatever representation that the C-I and A-P systems use based on their corresponding *lf* and *pf*.

One crucial character of the dynamic view of the linguistic interface embedded within the derivational model in Figure 1, which has been a recurrent theme throughout this dissertation, was that syntax is not entirely crash-proof; it could make certain derivational mistakes; some of them can be repaired by domain-specific operations at the interfaces such as deletion or choice functions so that they may still become usable at the language-independent articulatory and conceptual systems. This was clearly demonstrated in our investigation of the repair-by-deletion analysis for the P-stranding and the choice function analysis for *wh*-in-

situ questions. What is crash-proof is the linguistic system as a whole including the syntactic computation plus the linguistic interfaces.

The present model of linguistic interface driven by considerations of minimalist interfaces suggests a partial return to the earlier model of grammar as in the Government-and-Binding Theory, which reduces core syntactic computations to the single operation of *Move- α* (Chomsky 1981; Chomsky and Lasnik 1977; Lasnik and Saito 1992; see also Prince and Smolensky 1993). Under this generation + filter approach, the syntactic computation itself creates a multitude of syntactic objects and all ungrammatical sentences are filtered out by domain-/level-specific conditions such as government, binding, case theory, control theory, and bounding theory. The present model, however, is conceptually quite different from the GB conception of the “overgenerate +filter approach” in that the sole role of interfaces is not to constrain the forms of objects created by syntactic derivation but rather to improve them so that they become legible to the language-independent thought and production modules.

The thesis of minimalist interfaces serves not only as a hypothesis about the way syntactic computation networks with its neighboring linguistic interfaces but also opens a new line of inquiry on the proper understanding of the nature of syntactic derivation. For example, recall that this dissertation adopts the view that syntax is a functionally blind

combinatorial system (Uriagereka 1998, 2002; Hinzen 2006; Chomsky 2004, 2006) that cannot know about the fate of its own generated objects; they may or may not be convergent derivations. This means that it is the task of linguistic interfaces to determine the status of such objects in terms of usability on the part of language-external sound and concept systems and do whatever they can to modify/ornament/repair/remedy them so that they may become legible and usable on the part of the A-P/C-I systems. This view of convergence yields a new insight on the current debate on what syntactic objects constitute a phase domain. As stated in chapter 2, it has been currently proposed in Chomsky (2000, 2001, 2004, 2005, 2006) that certain syntactic objects such as ν Ps and CPs form a phase head; once these objects have been arranged in syntax, the complements of the phase heads ν and C undergo cyclic Spell-Out to the linguistic interfaces for semantic and phonological interpretation. Chomsky draws on considerations of computational efficiency in support of the phasehood of these functional heads; syntactic derivation can forget about what it constructed in the past if derivation works in this way. Under this conceptual view of phase theory, then, specific syntactic objects have privileged status in that they serve as a unit of objects not only for syntax but also for purposes of phonological and semantic interpretation. The present thesis of minimalist interface, however, leads us to the view quite different from the view just mentioned: it is a property of the linguistics interfaces

that requires these syntactic objects to be chunked into phases. This is because syntax is just about combining a subset of morphosyntactic features culled into numeration via the recursive process of Merge and create a hierarchical object out of them. There is nothing within syntax *per se* to tell us why *v*Ps and CPs are phases, not others (see Epstein and Seely 2002 for critical discussion of Chomsky's view of phases). Then, we naturally expect that interfaces are actively "invasive": they actively participate in linguistic computation in such a way that *v*Ps and CPs turn out to behave as phases; in this way, *v*Ps and CPs do not need to be arbitrarily identified as phases, but rather the fact that they are phases falls out as an epiphenomenon of the interface computations. This interface-driven view has also been recently argued for in a different context by Boeckx (2007), who proposes to let interfaces determine the convergence of objects created by syntax. This view is conceptually natural, as mentioned several times in this dissertation, in the light of the fact that what actually directly interacts with the language-independent sound and conceptual modules are semantic and phonological interfaces in Figure 1; to put it differently, by virtue of its place in Figure 1, syntax cannot worry about what will happen in the "negotiation" between these components and the A-P/C-I systems. It is in this respect that Chomsky's characterization of these objects in terms of interface properties is correct. *v*Ps and CPs are not inherent phases in the sense that their construction

automatically entails Spell-Out; rather, they are determined as phases contextually by the syntax-external semantic interface because they instantiate what have been variously characterized as having “full argument structure” (Chomsky 2000, 2001, 2004, 2005, 2006), “complete functional complex” (Chomsky 1986b), or “propositional content”; they happen to create good reconstruction sites that are necessary for proper interpretation of binding relations at the interface (Legate 2003). The same story holds for the other phonological interface: *v*Ps and CPs are phases not by virtue of their inherent privilege within the syntax but by virtue of the fact that they happen to demarcate possible domains for phonological rules such as pauses and parentheticals (Uriagereka 1999), phonological phrasing (Dobashi 2003), nuclear sentence stress (Legate 2003; Kahnemuyipour 2004; Wagner 2005), and sandhi/mutation phenomena (Sato in press c; Carnie in preparation); they happen to provide good escape hatches for various sorts of agreement such as WH-agreement in Chamorro (Chung 1982, 1994, 1998), the realis/irrealis alternation in Palauan (Georgopoulos 1985, 1991), the voice-movement interaction in Tagalog and Malagasy (Rackowski and Richards 2005; Pearson 2001, 2005) and the active voice deletion in Indonesian/Javanese/Madurese and many other phenomena briefly mentioned in chapter 2. The present interface-based approach to determination of phasal constituents, of course, also opens a new possibility that other syntactic objects than *v*Ps and CPs can be phases

depending on external requirements from linguistic interfaces. For example, the current minimalist interface is expected to derive the recent claim that DPs are one such candidate, as argued for independently in recent work such as Svenonius (2004) and Hiraiwa (2005), because certain types of DPs (action nominals, for example) could encode the same amount of information as their verbal counterparts in terms of argument structure. Other categories such as TP may well be strong phases given that the information contained within ν Ps forms a proper subset of that contained by TPs. Category-defining derivational morphemes such as *-al*, *-ous*, *-ful* and many others also may form phase heads, as proposed independently by Marantz (in press) within the framework of Distributed Morphology (see also Arad 2003 for potential arguments from the denominal verb formation in Hebrew), because the determination of syntactic category has direct relevance on the linguistic interfaces, as evidenced by categorial selection (# *John broke black*) and category-sensitive stress contours (PROduce_N vs. proDUCE_V). This line of thought, of course, does not take into consideration the recent conceptual argument made by Richards (2007) for the phasehood of ν Ps and CPs based on the computational efficiency and feature inheritance. The issue, however, is still on the jury.

One could maintain a slightly different view of the connection between syntax and its linguistic interfaces.¹ For example, Chomsky (2006:8-9) argues that “the relation of the generative procedure to the interfaces is asymmetrical, CI taking precedence: optimization is primarily to the CI interface” (see also Boeckx 2008). Chomsky illustrates this primacy of the CI interface from the way natural language deals with Internal Merge. The initial copy of an item is created by External Merge whereas all other copies of the same item are created by Internal Merge. Studies on reconstruction (recall our discussion of Legate’s (2003) study in chapter 2) show that all these copies play a role in semantic interpretation at the C-I interface. This state of affairs does not obtain at the A-P interface because in normal cases, all the copies except the highest one are deleted when the derivation is externalized at the interface. However, as is well known, this deletion pattern at the phonological interface causes serious problems in language processing (e.g., garden paths), a difficulty that would be easily overcome if all copies were pronounced at the interface. Chomsky takes this conflict between computational efficiency and communicative needs to support the primacy of the semantic interface in language design over the phonological interface. This primacy of the C-I interface in language design may also be reflected in the phasehood of particular syntactic objects such as *v*Ps and

¹ The primacy of the C-I interface in language design, expressed in the present paragraph, is also shared by Heidi Harley (personal communication, May 1, 2008), though in a different context related to the phasehood of particular syntactic objects. The comments that follow after my exposition of Chomsky’s (2006) position are based on my interpretation of her written comments on an earlier draft of this chapter.

CPs. Recall that Chomsky (2000, 2001, 2004, 2005, 2006) proposes that these particular objects constitute a natural characterization in terms of “full argument structure”, “proposition”, and other conceptual-semantic notions. In other words, this particular chunking receives independent justification in terms of what little is known about the way meaning is calculated in natural language. This characterization is not easy to come by at the phonological interface because there is nothing inherently special about these particular syntactic objects in terms of phonology that warrant their special treatment; other chunks such as AspP, TP, and VP may well count as phases as are CPs and TPs. True, one could make a case for the independent role of the phonological interface in the determination of phasehood of *v*Ps and CPs by their relative phonological isolability such as VP-fronting, pseudoclefting, intonational boundary, fragments, etc. (Chomsky 2004; see Bošković (2001) for a critical discussion on this point) but this behavior of these particular chunks may well be an epiphenomenon that arises from the fact that CPs and *v*Ps have conceptually identifiable propositional/argument-structural properties that are intrinsic to the way semantic computation works in parallel with narrow syntactic computation.

A similar interface-driven approach also sheds a new light on the status of the PIC discussed in chapter 2 of this dissertation.² This condition states that the complement of a

² I am grateful to Heidi Harley (personal communication), Cecile McKee (personal communication), and Mosa Hulden (personal communication) for sharing their perspectives on the content of this paragraph with me.

phase head undergoes Transfer to the interface components once the higher phase head is introduced into the syntactic workspace. Chomsky's only argument for this condition is computational complexity, namely, that this condition allows syntax to "forget about" material it deals with at earlier cycles of the syntactic derivation. A natural question to ask is, why do we need to worry about whether syntax forgets about material if the phase theory is intended to be an abstraction of syntactic computation? Of course, one could devise an argument for this particular condition on theory-internal grounds. The reductionist approach to syntax entailed by the minimalist interface thesis actually forces us to take a different approach to this condition: it is deeply rooted in the way parsers work. Let us suppose that human language parsers are top-down local parsers in that they cannot wait too long to calculate filler-gap dependencies. Let us also suppose with a leap of faith that "too long" means a filler-gap dependency that crosses either a vP and CP . Then, the local movement of a phrase required by the phase theory could receive a principled syntax-external motivation in terms of local processing; the parser must locate as fast as possible where the real gap is. If the filler-gap dependency is created by one-fell-swoop movement in a tri-clausal environment, parsers are unable to process this dependency. This is essentially the idea by Givón (1979). In a similar vein, developing a comprehensive principle-based approach, Pritchett (1991: 302) argues that island violations not

ungrammatical but rather “exceed the capacity of the human sentence processor to perform certain structural analyses during parsing.” This potential worry on the part of human language processor is easily eliminated if the edge serves to provide a local signal (either morphological or semantic) in a step-wise fashion that the real gap is still to be expected down below in a more deeply embedded clause. To take the distribution of the active voice morphology in Indonesian and Javanese, the PIC dictates that the movement is phase-dependent, with its reflexes manifested in these languages in the form of the *meN-/ng-* deletion. This analysis is couched in the syntax-phonology interaction. However, a deeper motivation for why this phase-based derivation is enforced in the computational component of human language might lie elsewhere; it is a good solution provided by the formal architecture of the interface to the local parser. If this speculation is real, the computational constraint such as the PIC might turn out to provide quite an important intersection for theoretical syntacticians and psycholinguists alike.

The present reductionist approach to various known properties of syntactic computation goes even deeper. Consider the notions of *phrase* and *word*. These notions have been considered as primitives of syntax since the advent of generative grammar. The current framework of Minimalist Program also seems to implicitly adopt this premise. Though the Bare Phrase Structure Theory of Chomsky (1995) and Speas (1990) have contributed to a

substantial simplification of the phrase structure in syntax, the $X^{\text{Max}} = \text{phrase}/X^{\text{Min}} = \text{word}$ equations are still assumed in much of the current research, at least tacitly. Carnie (1995, 2000) argues that the notions of XP/phrase and X/word are not primitives of syntax; rather, they behave as they are because of the properties they exhibit in various components of grammar and their interactions. Under this view, the verb *kick*, for example, is specified as what we normally call an X element not because it is inserted into a terminal node in syntactic derivation but because the other components of grammar require such a particular status for various (language-particular) reasons related to θ -marking, tense and agreement features, complement selection, reference, among many others. Carnie's theory of phrasality, thus, is quite interface-oriented in that it proposes to reduce the alleged phrasality in syntax to independently necessary conditions to be satisfied at the PF and LF components. This theory, therefore, allows for linguistic interfaces to choose and specify the status of an object that is left phrasally ambiguous within syntax.

7.3. Questions for Future Research and Conjectures about Linguistic Interfaces

I conclude this chapter by mentioning several important theoretical questions the model of linguistic interfaces in this dissertation brings to light. First, given the view adopted in this model that syntax is, as it were, a functionally blind computational system solely consisting of

abstract processes such as Internal/External Merge, Agree, Spell-Out that apply to a language-particular set of morphosyntactic features, then at what point do we know whether a particular syntactic object created by syntax converges or not? The proposed model suggests that it is the task of linguistic interfaces in Figure 1 to make sure that they will be legible to their neighboring C-I and A-P systems. This position seems natural under the particular conception of syntax as a generative system which does not care about the fate of its own syntactic objects. This conception is relatively easy to support at the syntax-semantics interface from numerous cases, some of them discussed in chapter 4, where semantics tries to give a reasonable interpretation (sometimes a coarse interpretation) to whatever objects syntax sent off to the interface. Evidence for the comparable position at the syntax-phonology interface is not easy to come by due to the ill-understood nature of the phonological interface at this point, but my guess is that this interface takes the form of syntactic representation just as the syntactic structure is because we have seen that domain-specific operations such as deletion can target only syntactic constituents such as TPs and *v*Ps.

The second question is what type of operations the syntax-external linguistic interfaces can conduct to repair syntactic failures. I have discussed only two such operations, deletion and choice functions, as two good candidates that these interfaces employ to save syntax. The thesis of minimalist interface should at least lead us to find many other operations.

Several operations on the phonological side of interface immediately come to mind: resumption (Sells 1984; Aoun and Li 2003; Boeckx and Lasnik 2006), copy spell-out (Bošković 2001; Landau 2007), focus intonation (Neelman and Reinhart 1998; Reinhart 2006), and phonological restructuring (Dobashi 2003). Some candidates for the other side of interface that might serve the role of repair for the purposes of communicative needs from the C-I system include topic-focus articulation (Kiss 1998; Reinhart 2006), metaphors (cf. Chomsky 2004), and so on. Detailed investigation of what other linguistic phenomena can be brought to bear on the ameliorating role of linguistic interfaces is an important task to undertake in future research.

The third question is what determines that a particular object, that is created by syntax and passed through domain-specific operations of its neighboring interface to the A-P/C-I systems, is grammatical or not? Under one standard view of the minimalist program, syntax never makes superfluous derivational steps (Economy of Derivation) or creates unnecessary representations (Economy of Representation); derivations and representations that involve mistakes are deemed simply underivable, hence ungrammatical. Under another view of the syntactic derivation, as in the Crash-Proof Syntax of Frampton and Guttman (1999, 2002), syntax is so constructed that every object generated by this perfect component is geared to be grammatical at the A-P and

C-I systems. We have seen that this view seems untenable to the extent that my analysis of the P-stranding under sluicing presented in chapter 3 and the choice function analysis of *wh*-in-situ questions given in chapter 5 is on the right track.

The minimalist interface thesis, of course, forces the conclusion that the convergence of a particular syntactic object is entirely for the A-P and C-I systems to decide because these and only these components use the object for purposes of communication but we have quite a limited understanding of what is in these systems. We do have some intuitive ideas about both of these domains such as the two dimensional nature of sound strings, the world knowledge, how things are naturally categorized and our understanding seems more clear in the C-I system than the A-P system. For example, we have seen a bit about how our C-I systems are supposed to work in chapter 4 where the denotation of a bare noun is sometimes coerced in service of the denotation of a functional element that goes with it or in contexts that force definite interpretation to the bare noun in languages such as Indonesian, Javanese, and Japanese that arguably lack DP projections. Things are less clear about the other system. For example, why is the failure of D-to-P incorporation, but not the failure of feature percolation, so critical a mistake so that the resulting string that contains the former violation is judged ungrammatical? The minimalist interface thesis could allow us to answer this

question as follows.³ Let us suppose that there are two types of violations in the syntax-phonology interface. One is a strictly syntactic/derivational “violation” that cannot be simply created in the syntactic computation. To take D-to-P incorporation, this operation is conducted in the syntax immediately once the preposition is introduced into the workspace and serves as a probe to attract the D head within its minimal search domain (its complement domain); whatever uninterpretable/unvalued feature of the P (e.g. strong D-feature in languages with D-P coalescence) needs to be checked must be checked, since that is the sole driving force for mechanical computation. The failure of the D-to-P incorporation, therefore, is simply an impossible scenario in the minimalist vision of syntactic computation. Thus, it is fully expected that there is no sense in which the failure of D-to-P incorporation could ever be repaired at the PF interface. The situation could be different with the other type of violation, interface violation. To take [+wh] feature percolation, failure of this process is a representational violation whose severity for linguistic computation could vary from language to language. Therefore, it is possible, in principle, that the failure of this percolation in languages with the obligatory value of this percolation mechanism (as in Indonesian) could be tolerated within syntax *per se* but rather is checked later at the PF interface. Under this view, syntactic representations that contain failures of percolation could still have

³ I owe the exposition of the following answer to this question to the written suggestions provided by Heidi Harley (personal communication, May 1, 2008) on an earlier draft of this chapter, which I paraphrase here.

chances to converge at the interface depending on what happens at this interface. If PF does not do anything about it, then this type of representation would persist at the interface: the representational constraint then applies to this representation and rules it out as ungrammatical. That was seen to be the case with P-stranding under *wh*-movement in Indonesian. If PF does conduct its domain-specific operation to the otherwise ill-formed object by deleting the offending part of the representation, then the representational constraint has nothing to apply to. As a result, the derivation can still continue to converge to yield a grammatical output at the A-P system. That was seen to be the case with P-stranding under sluicing/pseudogapping (*wh*-movement + TP/vP deletion) in Indonesian. Therefore, under this bipartite conception of violations, the contrast in “reparability” between the failure of D-to-P incorporation and the failure of feature percolation falls out naturally from the very architecture of the syntax-phonology interface proposed in this dissertation. One could think of a similar bifurcation in the interaction of syntax with the semantic interface, but examination of this possibility is left for another occasion.

In a similar vein, our present thesis also allows us to reach a better understanding of the distinction between *grammaticality* and *acceptability* in a way that the parsimonious (Chomsky 1995, 2000, 2001, 2004, 2005, 2006)/ crash-proof view (Frampton and Guttman 1999, 2002) of minimalist syntax cannot achieve. To take one example, it has

been widely known (see Chomsky 1986b, Lasnik and Saito 1984, 1992, Chomsky and Lasnik 1993, and many references cited therein) that there is a contrast between extraction of an argument and that of an adjunct from a syntactic island. Many papers have been written that attempt to come up with technical mechanisms to capture this such as gamma-marking, star-marking convention, and so on. The thesis of minimalist interfaces leads us to expect that this line of inquiry is misguided; rather, this contrast has nothing to do with syntactic mechanisms (in fact, a possibility that cannot be entertained anymore due to the Inclusiveness Condition) but tells us something about the way the C-I system independently works. One might entertain the idea that an inherently referential nominal object (such as *who*, *what*) is easier to retrieve than adjunct expressions (such as *why* and *how*), which do not have independent reference. This is not entirely an unreasonable possibility because I have provided evidence in chapter 5 that only those *wh*-phrases in languages such as English and Indonesian that contain an N-set can (or must) interpreted in situ by the choice function whose theoretical properties are not unlike those of the conceptual module.⁴ This possibility, in turn, invites another idea that often-varying degrees of unacceptability among speakers for island-violating examples are not a matter of syntax but of language-independent considerations about reference,

⁴ Thanks to Juan Uriagereka (personal communication) for useful discussion in this regard.

basic templates for conversation, frequency of forms, contexts of utterance, and many related semantic factors. Thus, the present thesis may even provide quite a unique interdisciplinary corroboration between pure theoretical linguists, psycholinguists, and philosophers/external semanticists.

However, a fundamental question still remains. What determines that a particular violation is repairable or not? The failure of feature percolation was argued to be repairable whereas the ECP-type/constituency-violation was not. Ross (1969) and Merchant (2001) argue that certain violations such as the *that*-Trace Effect, the Left Branch Condition, the Complex NP Constraint are repairable whereas Boeckx and Lasnik (2006) argue that superiority violations are not repairable. Where is this type of discrepancy rooted? As I have argued in chapter 2, the derivational vs. representational violation provides a first approximation in seeking an answer to this question. One might entertain the hypothesis that the former cannot be generated in syntax hence are never violable whereas the latter can be tolerated in syntax but cannot be processed at linguistic interfaces. This hypothesis thus relegates “repairable” violations to general cognitive limitations on language parsing. Finding the right cutting-off point between repairable and irreparable violations is a quite important research agenda for minimalist researchers.

The final and the most important question that the present dissertation brought to light is, why would language be organized in the way suggested in this dissertation? Why would a language not be a truly “perfect” system?⁵ This dissertation argued in several places, especially in chapters 3, 4, and 5, that the objects created by the core syntactic computation are quite imperfect and need to be remedied/modified/repared to be legible and usable on the part of the language-external A-P and C-I systems. One possible answer may be sought in the recent observation, made within the context of the Minimalist Program, that recursivity, the fundamental property of natural language ensured by the combinatorial process of Merge, is not a language-specific computational procedure; its workings could be seen in other human activities such as mathematics (number quantification), music, tools, spatial navigation, foraging, tracking, social interaction, among other relevant abilities (Hauser et al. 2002). This observation suggests that the general recursive system has been recruited into language, in much the same way it has been into other aspects of human activity such as those just mentioned above. Thus, it is reasonable to expect that this general system is not well adapted/fine-tuned to the particular way natural language should work in its actual use as communicative tool, even though this is the primary purpose of language,

⁵ I thank Andrew Carnie (personal communication) for asking this question and Dave Medeiros (personal communication) for useful comments. See also Uriagereka (1998, 2002), Hinzen (2006), and many works cited therein for much relevant discussion.

viewed from the perspective of the A-P and C-I systems. Once Merge has been wired into the faculty of language, however, the functionally blind purely mechanical combinatorial machine got connected to the A-P and C-I systems. The objects created by this component are so crude, hence are not tailored for the A-P and C-I systems to use. At the same time, those language-external systems are not general enough to be able to create representations solely based on the objects created by syntactic derivation. I conjecture that the linguistic interfaces PF and LF are developed in the language faculty, as broadly construed (including the sensory-motor and conceptual-intentional system; see Hauser et al. 2002 for the distinction between the *FLB* (the Faculty of Language in the Broad sense) and the *FLN* (the Faculty of Language in the Narrow sense), to solve this two-way discrepancy and to connect the syntax and the A-P/C-I modules *in the best possible and most economical way*. Specifically, if syntax creates an object that requires a minimum amount of work, then the PF/LF adds a correspondingly minimum amount of modification to send it off to the A-P/C-I systems. This was the case in the proposed analysis of the distribution of active voice morphology in Indonesian and Javanese (chapter 2), of the denotation and morphosyntax of bare nominals across languages (chapter 4), and of reduplication asymmetries between nominal and verbal affixes in Indonesian (chapter 6). If syntax creates an object that requires substantive modification for legibility, on the other hand, then the PF and LF conduct a

handful of domain-specific operations such as deletion and choice functions to the object to make it usable for the external systems. This was the case in the proposed analysis of the P-stranding pattern under sluicing in Indonesian (chapter 3) and of in-situ *wh*-questions in Indonesian (chapter 5). This conjecture leads us to adopt a particular minimalist vision of linguistic interfaces: PF and LF do only a minimum amount of work to enable the connection the purely mechanical computation and the language-external A-P/C-I systems for convergence. Whence came Minimalist Interfaces, the title of this dissertation.

The view of the language design expressed in the previous paragraph depends on a certain conception of “perfect language”, as often used in Chomsky’s minimalist work (Chomsky 1995, 2000, 2001, 2004, 2005, 2006).⁶ One of the important agendas of the minimalist inquiry since its inception (Chomsky 1995) has been to clarify and substantiate the question of “how well-designed/perfect language could be”. If we take “perfect language” to mean “how well-designed language could be *for interfaces*”, then the answer would be in the negative because the syntax is poorly designed for the purposes of actual language use; it creates failure of percolation (chapter 2), creates type mismatches (chapter 4), and lacks communicative diversities required for actual communication (chapter 6).

⁶ I am very grateful to Heidi Harley (personal communication, May 1, 2008) for extensive written comments on the content of this paragraph and clarifying two interrelated notions of “perfect language” here. The content in this paragraph is based on my understanding of her comments on an earlier draft of this chapter.

This is the view expressed in the last paragraph. In addition to this interface-based conception of language design, however, there is another sense of “perfect language” in which minimalist researchers ask “how well-designed language is *in terms of computational elegance, mathematical beauty, and simplicity*.” This seems to be the view that many minimalist/biolinguistic researchers (Medeiros 2006b), including Chomsky himself, currently adopt. Under this conception of “perfect language”, then, one could say that all imperfections actually lie in the linguistic interfaces, not within syntax. The AP and CI interfaces are so biologically specialized for the purposes of actual communication, parsing/processing, etc., that the objects created by such a mathematically elegant syntax would be too perfect to be usable for externalization at all. Resolution of conflicts between the two competing conjectures of language design, of course, requires a large-scale corroboration in comparative research, neuroscience, psycholinguistic, and related fields.

What has been said above in this section is all interesting but quite premature. However, there seems to be some converging evidence, as is clear from the results of earlier chapters, that this view of linguistic interfaces is not terribly a wrong idea. Whether this idea turns out to be a good heuristic of the further exploration of linguistic interfaces and their networking with the language-external A-P and C-I systems, of course, only time will tell.

REFERENCES

- Abdullah bin Abdul Kadir Munshi (1963 [1849]). *Hikayat Abdulla*, Vol. 2. Singapore: Pustaka Antara.
- Abels, Klaus (2003). *Successive cyclicity, anti-locality, and adposition stranding*. Doctoral dissertation, University of Connecticut.
- Adger, David (2007). Stress and phasal syntax. *Linguistic Analysis* 33 (2003): 238-266.
- Aldridge, Edith (In press). Phase-based account of extraction in Indonesian. *Lingua*.
- Allen, Margaret (1978). *Morphological investigations*. Doctoral dissertation, University of Connecticut.
- Almeida, Diogo and Masaya Yoshida (2007). A Problem for the preposition stranding generalization. *Linguistic Inquiry* 38: 349-362.
- An, Duk-Ho (2007). Clauses in non-canonical positions at the syntax-phonology interface. *Syntax* 10: 38-79.
- Anderson, Stephen (1976). On the notion of subject in ergative languages. *Subject and topic*, ed. Charles Li, 1-24. New York: Academic Press.
- Anderson, Stephen (1982). Where's morphology. *Linguistic Inquiry* 13: 571-612.

Anderson, Stephen (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.

Aoun, Joseph and Audrey Li (1990). Minimal disjointness. *Linguistics* 28: 189-203.

Aoun, Joseph and Audrey Li (2003). *Essays on the representational and derivational nature of grammar: The diversity of wh-constructions*. Cambridge, MA: MIT Press.

Aoun, Joseph, Norbert Hornstein, David Lightfoot, and Amy Weinberg (1987). Two types of locality. *Linguistic Inquiry* 18: 537-577.

Arad, Maya (2003). Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 21: 737-778.

Arka, Wayan and Christopher Manning (1998). Voice and grammatical relations in Indonesian: A new perspective. *The Proceedings of the LFG'98 Conference*, ed Miriam Butt and Tracy King. Stanford: CSLI Publications.

Arregi, Karlos (2002). *Focus on Basque movement*. Doctoral dissertation, MIT.

Baker, Carl. L. (1970). Notes on the description of English questions: The role of an abstract question morpheme. *Foundations of Language* 6: 197-219.

Baker, Mark (1985). The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373-415.

Baker, Mark (1988). *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Baker, Mark (1996). *The polysynthesis parameter*. New York: Oxford University Press.

Baker, Mark (2003). *Lexical categories: Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.

Baker, Mark, Kyle Johnson, and Ian Roberts (1989). Passive arguments raised. *Linguistic Inquiry* 20: 219-251.

Barss, Andrew (2000). Minimalism and asymmetric *wh*-interpretation. Step by step: *Essays on minimalism in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 31-52. Cambridge, MA: MIT Press

Bašić, Monika (2004). *Nominal subextractions and the structure of NPs in Serbian and English*. Master's thesis University of Tromsø.

Beerman, Dorothee (1990). *Semantic and syntactic constraints on German determiner movement within the prepositional phrase*. Master's thesis, University of Texas, El Paso.

Bejar, Susana and Diane Massam (1999). Multiple case checking. *Syntax* 2: 65-79.

Bernstein, Judy (1997). Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102: 87-113.

Bobaljik, Jonathan (1995). *Morphosyntax: The syntax of verbal inflection*. Doctoral dissertation, MIT.

Boeckx, Cedric (2007). Eliminating spell-out. *Linguistic Analysis* 34 (2003): 414-425.

- Boeckx, Cedric (2008). Approaching parameters from below. Ms., Harvard University.
- Boeckx, Cedric and Norbert Hornstein (2006). Control in Icelandic and theories of control. *Linguistic Inquiry* 37: 591-606.
- Boeckx, Cedric and Howard Lasnik (2006). Intervention and repair. *Linguistic Inquiry* 37:150-155.
- Borer, Hagit (1984). *Parametric syntax*. Dordrecht: Foris.
- Borer Hagit (1998). Morphology and syntax. *The handbook of morphology*, ed. Andrew Spencer and Arnold Zwicky, 151-190. Oxford: Blackwell.
- Borer, Hagit (2005). *Structuring sense volume 1: In name only*. New York: Oxford University Press.
- Borer, Hagit and Kenneth Wexler (1987). The maturation of syntax. *Parameter setting*, ed. Thomas Roeper and Edwin Williams, 23-172. Dordrecht: Leidel.
- Bošković, Željko (1997). *The syntax of nonfinite complementation: An economy approach*. Cambridge, MA: MIT Press.
- Bošković, Željko (2001). *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: North-Holland/Elsevier.

- Bošković, Željko (2003). On left branch extractions. *Investigations into formal Slavic linguistics: Contributions of the Fourth European Conference on Formal Descriptions of Slavic Languages-FDSL IV*, ed. Peter Kosta, Joanna Błaszczak, Jens Frasek, Ljudimila Geits, and Marzena Żygis, 543-577. Frankfurt an Main: Peter Lang.
- Bošković, Željko (2004). Topicalization, focalization, lexical insertion, and scrambling. *Linguistic Inquiry* 35: 613-638.
- Bošković, Željko (2005). On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59: 1-45.
- Bošković, Željko and Howard Lasnik (2003). On the distribution of null complementizers. *Linguistic Inquiry* 34: 528-546.
- Branigan, Phil and MacKenzie Maguerite (2002). Altruism, A-movement and object agreement in Innu-Aimûn. *Linguistic Inquiry* 33: 385-407.
- Bresnan, Joan (1971a). Sentence stress and syntactic transformations. *Language* 47: 257-281.
- Bresnan, Joan (1971b). Contraction and the transformational cycle. Ms., MIT.
- Bresnan, Joan (1982). The passive in lexical theory. *The mental representation of grammatical relations*, ed. Joan Bresnan, 3-86. Cambridge, MA: MIT Press.
- Bresnan, Joan, and Sam Mchombo (1995). The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13: 181-254.

- Carlson, Gregory (1977). *Reference to kinds in English*. Doctoral dissertation, University of Massachusetts, Amherst.
- Carnie, Andrew (1995). *Non-verbal predication and head movement*. Doctoral dissertation, MIT.
- Carnie, Andrew (2000). On the definitions of X^0 and XP. *Syntax* 5: 59-106.
- Carnie, Andrew (in preparation). Consonant mutation as evidence for a vP phase. Ms., University of Arizona.
- Carson, Jana (2000). *The semantics of number in Malay noun phrases*. Master's thesis, University of Calgary.
- Cartier, Alice (1989). De-voiced transitive verb sentences in formal Indonesian. *Serpih-serpih telaah pasif Bahasa Indonesia* [Studies on Passives in Indonesian], ed. Purwo Bambang Kaswanti, 84-145. Yogyakarta: Penerbit Kanisius.
- Cheng, Lisa (1991). *On the typology of wh-questions*. Doctoral dissertation, MIT.
- Cheng, Lisa and Rint Sybesma (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30: 509-542.
- Chierchia, Gennaro (1998a). Reference to kinds across languages. *Natural Language Semantics* 6: 339-405.
- Chierchia, Gennaro (1998b). Plurality of nouns and the notion of "semantic parameter. *Events and grammar*, ed. Susan Rothstein, 53-103. Dordrecht: Kluwer.

Choe, Jae-Woong (1987). LF movement and pied-piping. *Linguistic Inquiry* 18: 348-353.

Chomsky, Noam (1964). *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, Noam (1970). Remarks on nominalizations. *Readings in English transformational grammar*, ed. Roderick Jacobs and Peter Rosenbaum, 184-221. Waltham, MA: Blasdell.

Chomsky, Noam (1972). Some empirical issues in the theory of transformational grammar. *Goals of linguistic theory*, ed. Stanley Peters, 63-130. Englewood Cliffs: Winston.

Chomsky, Noam (1973). Conditions on transformation. *A festschrift for Morris Halle*, ed. Stephen Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.

Chomsky, Noam (1977). On *wh*-movement. In *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrianne Akmajian, 71-131. New York: Academic Press.

Chomsky, Noam (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam (1986a). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, Noam (1986b). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam (1991). Some notes on economy of derivation and representation. *Principles and parameters in comparative grammar*, ed. Robert Freidin, 417-454. Cambridge, MA: MIT Press. [Reprinted as chapter 2 in Chomsky (1995)]

Chomsky, Noam (1993). A minimalist program for linguistic theory. *A view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam (2000). Minimalist inquiries: The framework. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels & J. Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam (2001). Derivation by phase. *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam (2004). Beyond explanatory adequacy. *Structure and beyond: The cartography of syntactic structures, volume 3*, ed. A. Belletti, 104–131. Oxford: Oxford University Press.

Chomsky, Noam (2005). On phases. Ms., MIT. [To appear in *Foundational issues in linguistic theory*, ed. Robert Freidin, Carlos Otero, and Maria-Luisa Zubizarreta. Cambridge, MA: MIT Press.

Chomsky, Noam (2006). Approaching UG from below. Ms., MIT.

Chomsky, Noam and Howard Lasnik (1977). Filters and Control. *Linguistic Inquiry* 8: 425-504

- Chomsky, Noam and Howard Lasnik (1993). Principles and parameters theory. *Syntax: An international handbook of contemporary research*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506-569. [Reprinted as chapter 1 in Chomsky (1995)]
- Chung, Sandra (1976). An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7: 41–87.
- Chung, Sandra (1982). Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry* 13: 39–77.
- Chung, Sandra (1994). Wh-agreement and “referentiality” in Chamorro. *Linguistic Inquiry* 25:1–44.
- Chung, Sandra (1998). *The design of agreement: Evidence from Chamorro*. Chicago: University of Chicago Press.
- Chung, Sandra (2000). On reference to kinds in Indonesian. *Natural Language Semantics* 8: 157-171.
- Chung, Sandra (In press). Indonesian clause structure from an Austronesian perspective. *Lingua*.
- Chung, Sandra, William Ladusaw, and James McCloskey (1995). Sluicing and Logical Form. *Natural Language Semantics* 3: 239-282.
- Cinque, Guglielmo (1990). *Types of \bar{A} -dependencies*. Cambridge, MA: MIT Press.

- Cinque, Guglielmo (1999). *Adverbs and functional heads-A crosslinguistic perspective*. Oxford: Oxford University Press.
- Clahsen, Harold (1991). Constraints on parameter setting: A grammatical analysis of some acquisition stages in German child language. *Language Acquisition* 1: 361-391.
- Clahsen, Harald and Penke Martina (1992). The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-Corpus. *The acquisition of verb placement*, ed. Jürgen Meisel, 5-22. Dordrecht: Kluwer.
- Clements, George (1984). Binding domains in Kikuyu. *Studies in the Linguistic Sciences* 14: 37-56.
- Clements, George, James McCloskey, Joan Maling and Annie Zaenen (1983). String-vacuous rule application. *Linguistic Inquiry* 14: 1-17.
- Cole, Peter and Gabriella Hermon (1998). The typology of WH movement: WH questions in Malay. *Syntax* 1, 221-258.
- Cole, Peter and Gabriella Hermon (2000). Partial *wh*-movement: Evidence from Malay. *Wh-scope marking*, ed. Uli Lutz, Gereon Müller, and Arnim. von Stechow, 101-130. Amsterdam: John Benjamins.
- Cole, Peter and Gabriella Hermon (2005). Subject and non-subject relativization in Indonesian. *Journal of East Asian Linguistics* 14: 59-81.
- Cole, Peter and Gabriella Hermon (In press). Voice in Malay/Indonesian. *Lingua*.

- Cole, Peter, Jonczyk Elizabeth., & Lilley Jason (1999). A note on extraction from object position in Javanese and other Javanic languages. *Proceedings of AFLA VI*, ed. Carolyn Smallwood & Catherine Kitto, 87–93. Toronto: University of Toronto.
- Cole, Peter, Gabriella Hermon, and Norhaida Aman (To appear). Clefted questions in Malay. *Malay/Indonesian Linguistics*, ed. David Gil. London: Curzon Press.
- Contreras, Heles (1986). Spanish bare NPs and the ECP. *Generative studies in Spanish syntax*, ed. Ivonne Bordelois, Heles Contreras, and Karen Zagana, 25-49. Dordrecht: Foris.
- Corver, Norbert (1992). On deriving left branch extraction asymmetries: A case study in parametric syntax. *Proceedings of NELS 22*, ed. by Kimberly Broderick, 67-84. University of Massachusetts, Amherst: GLSA.
- Croft, William (1994). Semantic universals in classifier systems. *Word* 45: 145-171.
- Dardjowidjojo, Soenjono (1978). *Sentence patterns of Indonesian*. Honolulu: University of Hawaii Press.
- Davies, William (1999). Reduplication, multiple events, and Madurese reciprocals. In: Carolyn Smallwood and Catherine Kitto (eds.) *Toronto Working Papers in Linguistics: Proceedings of the Sixth Meeting of the Austronesian Formal Linguistics Association*, 95-106. Toronto: University of Toronto.
- Davies, William (2000). Events in Madurese reciprocals. *Oceanic Linguistics* 39: 123-143.
- Davies, William (2003). Extreme locality in Madurese WH-questions. *Syntax* 6: 237–259.

- Davies, William (2005). Madurese prolepsis and its implications for a typology of raising. *Language* 81: 645-665.
- Den Dikken, Marcel. (2001). 'Plurilinguals', pronouns and quirky agreement. *The Linguistic Review* 18, 19-41.
- Den Dikken, Marcel. (2006). A reappraisal of vP being phasal: A reply to Legate. Ms., CUNY.
- Déprez, Viviane (1989). *On the typology of syntactic positions and the nature of chains: Move α to the specifier of functional projections*. Doctoral dissertation, MIT.
- Di Sciullo, Anna-Maria and Edwin Williams (1987). *On the definition of words*. Cambridge, MA: MIT Press.
- Dobashi, Yoshihito (2003). *Phonological phrasing and syntactic derivations*. Doctoral dissertation, Cornell University.
- Duffield, Nigel (1995). *Particles and projections in Irish syntax*. Dordrecht: Kluwer.
- Du plessis, Hans (1977). *Wh-movement in Afrikaans*. *Linguistic Inquiry* 8: 723-726.
- Dyen, Isidore (1964). *Beginning Indonesian: Lessons 1-24*, Vol. 1-3. U.S. Department of Education: Language Research Section.

- Embick, David and Rolf Noyer (2007). Distributed morphology and the syntax/morphology interface. *The Oxford handbook of linguistic interfaces*, ed. Gillian Ramchand and Charles Reiss, 289-324. Oxford: Oxford University Press.
- Engdahl, Elisabet (1986). *Constituent questions*. Dordrecht: Reidel.
- Englebretson, Robert (2003). *Searching for structure: The problem of complementation in colloquial Indonesian conversation*. Amsterdam/Philadelphia: John Benjamins.
- Epstein, Samuel D. and Daniel Seely (2002). Rule applications as cycles in a level-free syntax. *Derivation and explanation in the minimalist program*, ed. Samuel D. Epstein and Daniel Seely, 65-89. Oxford: Blackwell.
- Epstein, Samuel D., Erich. M. Groat, Kawashima Ruriko & Kitahara Hisatsugu (1998). *A derivational approach to syntactic relations*. Oxford: Oxford University Press.
- Erteschik-Shir, Nomi (1977). *On the nature of island constraints*. Bloomington: Indiana University Linguistics Club.
- Fabb, Nigel (1984). *Syntactic affixation*. Doctoral dissertation, MIT.
- Farmer, Ann (1982). *On the interaction of morphology and syntax*. Doctoral dissertation, MIT.
- Felser, Claudia (2004). Wh-copying, phases and successive cyclicity. *Lingua* 114: 543–574

- Finer, Daniel (1998). V to D raising in Sulawesi relatives. *Arguments and structure in Austronesian languages: Thematic issue of the Canadian Journal of Linguistics* 43, ed. Diane Massam and Barry Miller, 283-306.
- Fodor, Janet, and Ivan Sag (1982). Referential and quantificational indefinites. *Linguistics and Philosophy* 5: 355-398.
- Fortin, Catherine (2007a). Indonesian sluicing. Paper presented at the 81st Annual Meeting of the Linguistic Society of America, Anaheim, CA. January 2007.
- Fortin, Catherine (2007b). *Indonesian sluicing and verb phrase ellipsis: Description and explanation in a minimalist framework*. Doctoral dissertation, University of Michigan.
- Fortin, Catherine (In press). Reconciling *meng-* and NP movement in Indonesian. *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistic Society*. University of California, Berkeley: Berkeley Linguistics Society.
- Fox, Danny (1998). *Economy and semantic interpretation*. Doctoral dissertation, MIT.
- Fox, Danny (2000). *Economy and semantic interpretation*. Cambridge, MA: MIT Press.
- Fox, Danny and Howard Lasnik (2003). Successive-cyclic movement and island repair: The difference between sluicing and VP-ellipsis. *Linguistic Inquiry* 34: 143-154.
- Frampton, John and Samuel Guttman (1999). Cyclic computation, a computational efficient minimalist syntax. *Syntax* 2: 1-27.

- Frampton, John and Samuel Guttman (2002). Crash-proof syntax. *Derivation and explanation in the minimalist program*, ed. Samuel Epstein and Daniel Seely, 90-105. Oxford: Blackwell.
- Franko, Jon (1993). *On object agreement in Spanish*. Doctoral dissertation, University of Southern California.
- Franko, Jon and Alazne Landa (2006). Preverbal N-words and anti-agreement effects. *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 34-42. Somerville, MA: Cascadilla.
- Franks, Steve and Asya Pereltsvaig (2004). Functional categories in the nominal domain. *Proceedings of FASL 12: The Ottawa Meeting 2003*, ed. Olga Arnaudova, Wayles Browne, Maria Luisa Rivero and Danijela Stojanović, 109-128. Ann Arbor, MI: University of Michigan Press.
- Fromkin, Vicki, Robert Rodman, and Nina Hyams (2003). *An introduction to language, seventh edition*. Boston: Heinle.
- Fukui, Naoki (1986). *A theory of category projections and its application*. Doctoral dissertation, MIT.
- Fukui, Naoki (1993) A note on improper movement. *The Linguistic Review* 10: 111-126.
- Fukui, Naoki (1995). *Theory of projection in syntax*. Stanford, CA: CSLI Publications.

- Fukui, Naoki and Margaret Speas (1986). Specifiers and projections. *Papers in Linguistics* 8: *Papers on Theoretical Linguistics*, ed. Naoki Fukui, Tova Rapoport and Elisabeth Sagey, 128-172. Department of Linguistics and Philosophy: MITWPL.
- Georgopoulos, Carol (1985). Variables in Palauan syntax. *Natural Language & Linguistic Theory* 3: 59–94.
- Georgopoulos, Carol (1991). *Resumptive pronouns and A'-binding in Palauan*. Dordrecht: Kluwer.
- Ghomeshi, Jila (1996). *Projection and inflection: A study of Persian phrase structure*. Doctoral dissertation, University of Toronto.
- Gil, David (2002). The prefixes *di-* and *N-* in Malay/Indonesian dialects. *The history and typology of Western Austronesian voice systems*, ed. Fay Wouk and Malcolm Ross, 241-283. Canberra: Pacific Linguistics.
- Gil, David (2005). Word order without syntactic categories: How Riau Indonesian does it. *Verb first: On the syntax of verb initial languages*, ed. Andrew Carnie, Heidi Harley, and Sheila Dooley, 243-263. Amsterdam: John Benjamins.
- Givón, Talmy (1979). *On understanding grammar*. New York: Academic Press.
- Goldberg, Adele (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, Adele (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Grimshaw, Jane (1991). Extended projections. Ms., Brandeis University.

Grimshaw, Jane (1993). Minimal projection, heads and optimality. Ms. Rutgers University Center for Cognitive Science.

Grimshaw, Jane (2005). *Words and structure*. Stanford, CA: CSLI Publications.

Groat, Erich and John O'Neil (1996). Spell-out at the LF interface. *Minimal ideas: Syntactic studies in the minimalist framework*, ed. Abraham Werner, Samuel Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart, 113-139. Amsterdam/Philadelphia: Benjamins.

Groenendijk, Jeroen and Martin Stokhof (1997). Questions. *Handbook of logic and language*, ed. Johan van Benthem and Alice ter Meulen, 1055-1124. Amsterdam/Cambridge, MA, Elsevier/MIT Press.

Grohmann, Kleanthes (2003). *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins.

Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis (1992). Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory* 10, 375-414.

Guilfoyle, Eithne and Maire Noonan (1992). Functional categories and language acquisition. *Canadian Journal of Linguistics* 37: 241-272.

Haegeman, Liliane (1992). *Theory and description in generative syntax: A case study in West Flemish*. Cambridge: Cambridge University Press.

- Hale, Kenneth and Samuel J. Keyser (1993). On argument structure and the lexical expression of syntactic relations. *The view from building 20: Essays on linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel J. Keyser, 53-109. Cambridge, MA: MIT Press.
- Hale, Kenneth and Samuel J. Keyser (2002). *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.
- Halle, Morris (1997). Distributed morphology: Impoverishment and fission. *MIT Working Papers in Linguistics 30: Papers at the Interface*, ed. Benjamin Bruening, Yoonjung Kang, and Martha McGinnis, 425-449. Department of Linguistics and Philosophy: MITWPL.
- Halle, Morris and Alec Marantz (1993). Distributed morphology and the pieces of inflection. *A view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111-176. Cambridge, MA: MIT Press.
- Halle, Morris and Alec Marantz (1994). Some key features of Distributed Morphology. *MIT Working Papers in Linguistics 21: Papers on phonology and morphology*, ed. Andrew Carnie and Heidi Harley, 275-288. Department of Linguistics and Philosophy: MITWPL.
- Harley, Heidi (2005). One-replacement, unaccusativity, acategorial roots and bare phrase structure. *Harvard Working Papers in Linguistics 11*, ed. Slava Gorbachov and Andrew Nevins. Harvard University: Department of Linguistics.
- Harley, Heidi (2006). *English words: A linguistic introduction*. Oxford: Blackwell.

- Harley, Heidi (2008). Compounding in distributed morphology. Ms., University of Arizona, Tucson. [To appear in *The Oxford handbook of Compounding*, ed. Rochelle Lieber and Pavol Stekauer, Oxford: Oxford University Press]
- Harley, Heidi and Rolf Noyer (1999). State-of-the-Article: Distributed morphology. *Glott International* 4: 3-9.
- Harley, Heidi and Rolf Noyer (2000). Formal versus encyclopedia properties of vocabulary: Evidence from nominalizations. *The lexicon-encyclopedia interface*, ed. Bert Peters, 349-374. Amsterdam: Elsevier.
- Hartman, Jeremy (2005). Sluicing in Finnish. Ms., Harvard University.
- Haugen, Jason (2004). *Issues in comparative Uto-Aztecan morphosyntax*. Doctoral dissertation. University of Arizona.
- Haugen, Jason and Heidi Harley (2006). Reduplication in Yaqui (Hiaki) compound verbs and the lexicalist hypothesis. Ms., University of Arizona.
- Hauser, Mark, Noam Chomsky, and Tecumseh Fitch (2002). The language faculty: What is it, who has it, and how did it evolve? *Science* 298: 1569-1579.
- Heim, Irene (1982). *The semantics of definite and indefinite nouns phrases*. Doctoral dissertation, University of Massachusetts, Amherst.
- Heim, Irene and Angelika Kratzer (1998). *Semantics in generative grammar*. Oxford: Blackwell.

- Hersulund, Michael (1984). Particles, prefixes and preposition stranding. *Topics in Danish syntax*, ed. Finn Sørensen and Lars Heltoft, 34-71. Copenhagen: Akademisk Forlag.
- Higginbotham, James (1985). On semantics. *Linguistic Inquiry* 16: 547-594.
- Higginbotham, James and Robert May (1981). Questions, quantifiers, and crossing. *The Linguistic Review* 1: 41-79.
- Hinzen, Wolfram (2006). *Mind design and minimal syntax*. Oxford: Oxford University Press.
- Hiraiwa, Ken (2005). *Dimensions of symmetry in syntax: Agreement and clausal architecture*. Doctoral dissertation, MIT.
- Hopper, Paul (1983). Ergative, passive and active in Malay narrative. *Discourse perspective on syntax*, ed. Flora Kellin-Andreu, 67-88. New York: Academic Press.
- Hopper, Paul (1986). Some discourse functions of classifiers in Malay. *Noun classes and categorization*, ed. Chaudron Craig, 309-325. Amsterdam: John Benjamins.
- Hornstein, Norbert and Amy Weinberg (1981). Case theory and preposition stranding. *Linguistic Inquiry* 12: 55-91.
- Horvath, Julia (1997). The status of 'wh-expletives' and the partial movement construction of Hungarian. *Natural Language & Linguistic Theory* 15: 509-572.
- Huang, J. C.-T. (1982). *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT.

- Iljic, Robert (1994). Quantification in Mandarin Chinese: Two markers of plurality. *Linguistics* 32: 91-116.
- Ionin, Tania and Ora Matushansky (2006). The composition of complex cardinals. *Journal of Semantics* 23: 315-360.
- Jackendoff, Ray (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jackendoff, Ray (1990). *Semantic structures*. Cambridge, MA: MIT Press.
- Jayaseelan, Karattuparambil (1990). Incomplete VP deletion and gapping. *Linguistic Analysis* 20: 64-81.
- Jensen, John, and Stong-Jensen Margaret (1984). Morphology is in the lexicon! *Linguistic Inquiry* 15: 474-498.
- Johns, Alana (1999). The lexical basis of noun incorporation in Inuktitut. Paper presented at the Workshop on the Structure and Constituency in the Languages of the Americas, University of British Columbia, March 1999.
- Kahnemuyipour, Arsalan (2004). *The syntax of sentential stress*. Doctoral dissertation, University of Tronto.
- Kana, Vamarasi M. (1986). *Grammatical relations in Bahasa Indonesia*. Doctoral dissertation, Cornell University.

- Kaplan, Ronald M., and Joan Bresnan (1982). Lexical-functional grammar: A formal system for grammatical representation. *The mental representation of grammatical relations*, ed. Joan Bresnan, 173-281. Cambridge, MA: MIT Press.
- Karttunen, Lauri (1977). Syntax and semantics of question. *Linguistics and Philosophy* 1: 3-44.
- Kats, Jerrold and Paul Postal (1964). *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.
- Kayne, Richard (1981). On certain differences between English and French. *Linguistic Inquiry* 12: 349-371.
- Kayne, Richard (1989). Facets of Romance past participle agreement. *Dialect variation in the theory of grammar*, ed. Paola Benincà, 85-103. Dordrecht: Foris.
- Kayne, Richard (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Kayne, Richard (1998). Overt vs. covert movement. *Syntax* 1: 128-191.
- Kayne, Richard and Jean-Yves Pollock (1979). Stylistic inversion, successive cyclicity and Move NP in French. *Linguistic Inquiry* 9: 595-621.
- Keenan, Edward (1972). Relative clause formation in Malagasy (and some related and some not so related languages). *The Chicago which hunt: Papers from the Relative Clause Festival*, ed. Paul Peranteau, Judith Levi, and Gloria Phares, 169-189. Chicago: Chicago Linguistic Society.

- Keenan, Edward and Bernard Comrie (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 1: 63–99.
- Kiparsky, Paul (1982a). From cyclic phonology to lexical phonology. *The structure of phonological representations*, ed. Hulst H. van der and Norval Smith, 131-175. Dordrecht: Foris.
- Kiparsky, Paul (1982b). Lexical morphology and phonology. *Linguistics in the morning calm*, ed. The Linguistic Society of Korea, 3-91. Seoul: Hanshin.
- Kiparsky, Paul (1982c). Word formation and the lexicon. *Proceeding of the 1982 Mid-America Linguistics Conference*, ed. Frances Ingemann, 3-22. University of Kansas: Department of Linguistics.
- Kiparsky, Paul (1985). Some consequences of lexical phonology. *Phonology Yearbook* 2, ed. Kaisse, Ellen M. and Patricia A. Shaw, 83-138.
- Kiss, Katalin (1998). Identificational Focus versus Information Focus. *Language* 74: 245-268.
- Kitagawa, Chisato and Claudia Ross (1982). Prenominal modification in Chinese and Japanese. *Linguistic Analysis* 9: 19-53.
- Kizu, Mika (2005). *Cleft Constructions in Japanese Syntax*. New York: Palgrave Macmillan.
- Klima, Edward (1964). Negation in English. *The structure of language: Readings in the philosophy of language*, ed. Jerry Fodor and Jerrold Katz, 246-323. Englewood Cliffs, NJ: Prentice-Hall.

- Koopman, Hilda and Dominique Sportiche (1983). Variables and the bijection principle. *The Linguistic Review* 2: 139-160.
- Koopman, Hilde and Dominique Sportiche (1991). The position of subjects. *Lingua* 85: 211-258.
- Kratzer, Angelika (1996). Severing the external argument from its verb. *Phrase structure and the lexicon*, ed. Johan Rooryck and Laurie Zaring, 109-137. Dordrecht: Kluwer.
- Kratzer, Angelika (1998). Scope or pseudoscope? Are there wide-scope indefinites? *Events in grammar*, ed. S. Rothstein, 163-196. Dordrecht: Kluwer.
- Krifka, Manfred, Francis Pelletier, Gregory Carlson, Alice ter Meulen, Gennaro Chierchia, and Godehard Link (1995). *Genericity: An introduction*, ed. Gregory Carlson and Francis Pelletier. 1-124. London/Chicago: University of Chicago Press.
- Kuroda, Shigeyuki (1965). *Generative grammatical studies in the Japanese language*. Doctoral dissertation. MIT.
- Landau, Idan (2007). Constraints on partial VP-fronting. *Syntax* 10: 127-164.
- Lapointe, Steven (1980). *A theory of grammatical agreement*. Doctoral dissertation, University of Massachusetts, Amherst.
- Lapointe, Steven (1981). General and restricted agreement phenomena. *The scope of lexical rules*, ed. Michael Moortgat, Harry van der Hulst and Teun Hoekstra, 125-159. Dordrecht: Foris.

Larson, Richard (1994). The projection of DegP and DP structure. Ms., SUNY, Stony Brook.

Lasnik, Howard (1972). *Analysis of negation in English*. Doctoral dissertation, MIT.

Lasnik, Howard (1995). A note on pseudogapping. *MIT Working Papers in Linguistics 27: Papers on Minimalist Syntax*, ed. Robert Pensalfini and Hiroyuki Ura, 143-163. Department of Linguistics and Philosophy: MITWPL.

Lasnik, Howard (1999). On feature strength: Three minimalist approaches to overt movement. *Linguistic Inquiry* 30: 197-217.

Lasnik, Howard (2001). When can you save a structure by destroying it? *Proceedings of NELS 31*, ed. Min-Joo Kim and Uri Strauss, 301-320. University of Massachusetts, Amherst: GLSA

Lasnik, Howard (2005) "Speculations on P-Stranding," Class lecture handout, Linguistic Society of America Summer Institute, MIT/Harvard University, July 18-August 3 [available at <http://www.ling.umd.edu/~lasnik/LSA%20course/Lasnik%20LSA%20HO%20Sect%202%20Appendix%20P-stranding.pdf>]

Lasnik, Howard (2007). On ellipsis: The PF approach to missing constituents. *University of Maryland Working Papers in Linguistics 15*, ed. Anastasia Conroy, Chunyuang Jing, Chizuru Nakao, and Eri Takahashi, 143-153. Department of Linguistics: UMWPL.

Lasnik, Howard and Mamoru Saito (1984). On the nature of proper government. *Linguistic Inquiry* 15: 235-289.

- Lasnik, Howard and Mamoru Saito (1992). *Move α : Conditions on its application and output*. Cambridge, MA: MIT Press.
- Law, Paul (1991). *Effects of head movement on theories of subjacency and proper government*. Doctoral dissertation, MIT.
- Law, Paul (1998). A unified analysis of P-Stranding in Romance and Germanic. *Proceedings of NELS 28*, ed. Pius Tamanji and Kiyomi Kusumoto, 219-234. University of Massachusetts, Amherst: GLSA.
- Law, Paul (2006). Preposition stranding. *The Blackwell companion to syntax, volume III*, ed. Martin Everaert, Henk van Riemsdijk, Rob Goedemans, and Bart Hollebrandse, 631-684. Oxford: Blackwell.
- Lebeaux, David (1988). *Language acquisition and the form of grammar*. Doctoral dissertation, University of Massachusetts, Amherst.
- Legate, Julie (2003). Some interface properties of the phase. *Linguistic Inquiry* 34: 506–516.
- Leko, N (1999). Functional categories and the structure of the DP in Bosnian. *Topics in south Slavic syntax and semantics*, ed. Mila Dimitrova-Vulchanova and Lars Hellan, 229-252. Amsterdam: John Benjamins.
- Levin, Beth and Malka Rappaport H (1986). The formation of adjectival passives. *Linguistic Inquiry* 17: 623-661.
- Levin, Beth and Malka Rappaport Hovav (1995). *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.

- Li, Yafei (1990). *Conditions on X-movement*. Doctoral dissertation, MIT.
- Lin, Jo-Wang (2004). Choice functions and scope of existential polarity *wh*-phrases in Mandarin Chinese. *Linguistics and Philosophy* 27: 451-491.
- Longobardi, Giuseppe (1994). Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25: 609-655.
- López, Luis (2001). On the (non)complementarity of θ -theory and checking theory. *Linguistic Inquiry* 32: 694-716.
- MacDonald, Ross (1976). *Indonesian reference grammar*. Washington, DC: Georgetown University Press.
- MacDonald, Ross and Soenjono Dardjowidjojo (1967). *Indonesian reference grammar*. Washington, DC: Georgetown University Press.
- Mahajan, Anoop (1989). Agreement and agreement phrases. *MIT Working Papers in Linguistics 10: Functional heads and clause structure*, ed. Itziar Laka and Anoop Mahajan, 217-252. Department of Linguistics and Philosophy: MITWPL.
- Mahajan, Anoop (1990). *The A/A-bar distinction and movement theory*. Doctoral dissertation, MIT.
- Manning, Christopher (1996). *Ergativity: Argument structure and grammatical relations*. Stanford, CA: CSLI Publications.

Marantz, Alec (1982). Re reduplication. *Linguistic Inquiry* 13: 435-482.

Marantz, Alec (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *Upenn Working Papers in Linguistics 4: Proceedings of the 21st Annual Penn Linguistics Colloquium*, ed. Alexis Dimidiates, Laura Siegel, Clarissa Sure-Clark, and Alexander Williams, 201-225. Philadelphia: University of Pennsylvania, Department of Linguistics.

Marantz, Alec (In press). Phases and words. Ms., New York University.

Massam, Diane (2001). Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19: 153-197.

Matthewson, Lisa (1999). On the interpretation of wide-scope indefinites. *Natural Language Semantics* 7: 79-134.

May, Robert (1985). *Logical form*. Cambridge, MA: MIT Press.

McCarthy, John, and Alan Prince (1986). Prosodic morphology. Ms. University of Massachusetts, Amherst & Rutgers University.

McCarthy, John, and Alan Prince (1993). Prosodic morphology I: Constraint interaction and satisfaction. Ms. University of Massachusetts, Amherst & Rutgers University.

McCarthy, John, and Alan Prince (1995). Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers in Linguistics 18: Papers in optimality theory*, ed. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 249-384. University of Massachusetts, Amherst: GLSA.

- McCloskey, James (1979). *Transformational syntax and model theoretic semantics: A case study in Modern Irish*. Dordrecht: Reidel.
- McCloskey, James (2001). The morphosyntax of WH-extraction in Irish. *Journal of Linguistics* 37, 67–100.
- McDaniel, Dana (1989). Partial and multiple *wh*-movement. *Natural Language & Linguistic Theory* 7: 565–604.
- Medeiros, David (2006a). Pied-piping with inversion and labeling. Paper presented at the 80th Annual Meeting of the Linguistic Society of America, Albuquerque, New Mexico.
- Medeiros, David (2006b). A case of ontological minimalism: X-bar as an optimal packaging solution. Ms., University of Arizona, Tucson.
- Merchant, Jason (1998). Pseudosluicing: Elliptical clefts in Japanese and English. *ZAS Working Papers in Linguistics* 10, ed. Artemis Alexiadou, Nanna Fuhrhop, Paul Law, and Ursula Kleinhenz, 88–112. Zentrum für Allgemeine
- Merchant, Jason. (2001) *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Merchant, Jason (2008). An asymmetry in voice mismatches in VP-ellipsis and pseudogapping. *Linguistic Inquiry* 39: 169–179.
- Miller, Wick and Susan Ervin-Tripp (1973). The development of grammar in child language. *Studies in child language development*, ed. Dan Slobin and Charles Ferguson, 355–380. New York: Holt, Rinehart and Winston.

- Mohanan, K.P. (1986). *The theory of lexical phonology*. Dordrecht: Reidel.
- Moravcsik, Edith (1978). Reduplicative constructions. *Universals of human language* volume 3, ed. Joseph H. Greenberg, 297-334. Stanford, CA: Stanford University Press.
- Musgrave, Simon (2001). *Non-subject arguments in Indonesian*. Doctoral dissertation, University of Melbourne.
- Nakamura, Masanori (1994). An economy account of *wh*-extraction in Tagalog. *The Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, ed. Erin Duncan, Donka Farkas, and Phillip Spaelti, 405-420. Stanford: CSLI Publications.
- Nakanishi, Kimiko and Satoshi Tomioka (2004). Japanese plurals are exceptional. *Journal of East Asian Linguistics* 13: 113-140.
- Ndayiragije, Juvénal (2006). The ergativity parameter: The view from antipassive. *Ergativity: Emerging issues*, ed. Alana Johns, Diane Massam, and Juvénal Ndayiragije, 271-292. Dordrecht: Kluwer.
- Neeleman, Ad and Tanya Reinhart (1998). Scrambling and the PF interface. *The projection of arguments: Lexical and compositional factors*, ed. Miriam Butt and Wilhelm Geuder, 309-353. Stanford: CSLI Publications.
- Nichols, Lynn (1997). DP and polysynthesis. *Proceedings of NELS 27*, ed. Kiyomi Kusumoto, 305-320. University of Massachusetts, Amherst: GLSA.
- Nishigauchi, Taisuke (1986). *Quantification in syntax*. Doctoral dissertation, University of Massachusetts, Amherst.

- Nishigauchi, Taisuke (1990). *Quantification in the theory of grammar*. Dordrecht: Kluwer.
- Nishiyama, Kunio, John Whitman, and Eun-Young Yi (1996). Syntactic movement of Overt *wh*-phrases in Japanese and Korean. *Japanese/Korean Linguistics* 5, ed. Noriko Iwasaki, Susan Strauss, and Shoichi Iwasaki, 337-351. Stanford: CSLI Publications.
- Noguchi, Tohru (1997). Two types of pronouns and variable binding. *Language* 73: 770-797.
- Nunes, Jairo (2004). *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.
- Ouhalla, Jamal (1993). Subject extraction, negation and the anti-agreement effect. *Natural Language & Linguistic Theory* 11: 477-518.
- Ouhalla, Jamal (2005). Agreement features, agreement and antiagreement. *Natural Language and Linguistic Theory* 23: 655-686.
- Pearson, Matthew (2001). *The clause structure of Malagasy: A minimalist approach*. Doctoral dissertation, University of California, Los Angeles.
- Pearson, Matthew (2005). The Malagasy subject/topic as an A'-element. *Natural Language & Linguistic Theory* 23: 381-457.
- Pereltsvaig, Asya (2007). The universality of DP: A view from Russian. *Studia Linguistica* 61: 59-94.

- Pérez-Leroux Ana and Thomas Roeper (1999). Scope and the structure of bare nominals: Evidence from child language. *Linguistics* 37: 927-960.
- Pesetsky, David (1979). Russian morphology and lexical theory. Ms. MIT.
- Pesetsky, David (1987). *Wh-in-situ: Movement and unselective binding. The representation of (in)definites*, ed. Eric. Reuland and Alice. ter Meulen, 98-129. Cambridge, MA: MIT Press.
- Pesetsky, David (1995) *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
- Pesetsky, David (1998). *Phrasal movement and its kin*. Cambridge, MA: MIT Press.
- Platzack, Krister (1991). A grammar without functional categories: A syntactic study of early child language. *Nordic Journal of Linguistics* 13: 107-126.
- Poedjosoedarmo, Soepomo (1982). *Javanese influence on Indonesian*. Materials in Languages of Indonesia, No. 7. *Pacific Linguistics*, Series D., No. 38.
- Pollard, Carl, and Ivan Sag (1994). *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
- Postal, Paul (1971). *Cross-over phenomena*. New York: Holt, Rinehart, and Winston.
- Postal, Paul (1972). On some rules that are not successive cyclic. *Linguistic Inquiry* 3: 211-222.

- Postman, Whitney (2002). *Thematic role assignment in Indonesian: A case study of agrammatic aphasia*. Doctoral dissertation, Cornell University.
- Potsdam, Eric (2003). Evidence for Semantic Identity under Ellipsis from Malagasy,” *Proceedings of NELS 33*, ed. Makoto Kadowaki and Shigeto Kawahara, 285-302. University of Massachusetts, Amherst: GLSA.
- Potsdam, Eric (2007) “Malagasy Sluicing and its Consequences for the Identity Requirement on Ellipsis,” *Natural Language and Linguistic Theory* 25, 577-613.
- Prince, Alan and Paul Smolensky (1993). Optimality theory: Constraint interaction in generative grammar. Rutgers University Center for Cognitive Science.
- Pritchett, Bradley (1991). Subjacency in a principle-based parser. *Principle-based parsing: Computation and psycholinguistics*, ed. Robert Berwick, Steven Abney, and Carol Tenny, 301-345. Dordrecht: Kluwer.
- Provovac, Ljiljana (1998). Determiner phrase in a language without determiners. *Journal of Linguistics* 34: 165-179.
- Purwo, Bambang Kaswanti (ed.) (1989). *Serpih-serpih telaah pasif bahasa Indonesia*. [Studies on Passives in Indonesian]. Yogyakarta: Penerbit Kanisius.
- Rackowski, Andrea & Norvin Richards (2005). Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36: 565–599.
- Radford, Andrew (1990). *Syntactic theory and the acquisition of English syntax*. Oxford: Blackwell.

- Radford, Andrew (2004). *Minimalist syntax: Exploring the structure of English*. Cambridge: Cambridge University Press.
- Rappaport, Gilbert (2001). Extraction from nominal phrases in Polish and the theory of determiners. *Journal of Slavic Linguistics* 8: 159-198.
- Reinhart, Tanya (1981). A second COMP position. *Theory of markedness in generative grammar*, ed. Adrianna Belletti, Luciana Brandi, and Luigi Rizzi, 517-557. Pisa: Scuola Normale Superiore.
- Reinhart, Tanya (1991). Elliptic conjunctions-non-quantificational LF. *The Chomskyan turn*, ed. Asa Kasher, 360-384. Oxford: Blackwell.
- Reinhart, Tanya (1992). *Wh*-in-situ: An apparent paradox. In *Proceedings of the Eighth Amsterdam Colloquium*, ed. Paul Dekker and Martin Stokhof, 483-491. University of Amsterdam: LLC.
- Reinhart, Tanya (1995). Interface strategies. *OTS Working Papers in Linguistics*. Research Institute for Language and Speech, Utrecht University.
- Reinhart, Tanya (1997). Quantifier scope: How labor is divided between QR and choice function. *Linguistic and Philosophy* 20: 335-397.
- Reinhart, Tanya (1998). WH-in-situ in the framework of the minimalist program. *Natural Language Semantics* 6: 29-56.
- Reinhart, Tanya (2006). *Interface strategies: Reference-set computation*. Cambridge, MA: MIT Press.

- Richards, Marc (2006). Weak pronouns, object shift, and multiple spell-out: Evidence for phases at the PF-interface. *Minimalist essays*, ed. Cedric Boeckx, 160-181. Amsterdam: John Benjamins.
- Richards, Marc (2007). On feature inheritance: An argument from the phase impenetrability condition. *Linguistic Inquiry* 38: 563-572.
- Richards, Norvin (2001). *Movement in language: Interactions and architectures*. Oxford: Oxford University Press.
- Rizzi, Luigi (1990). *Relativized minimality*. Cambridge, MA: MIT Press.
- Rizzi, Luigi (1997) The fine structure of the left periphery. *Elements of grammar: Handbook of generative syntax*, ed. Liliane Haegeman, 281-337. Dordrecht: Kluwer.
- Robson, Stuart (2002). *Javanese grammar for students*. Victoria: Monash University Press.
- Rodrigues, Cilene, Andrew Nevins, and Luis Vicente (2007). Cleaving the Interactions between Sluicing and Preposition Stranding,” ms., Unicamp, Harvard, and University of California, Santa Cruz. [To appear in *Proceedings of Going Romance 20*, Vrije Universiteit.]
- Rogayah, Razak (1995). *The syntax and semantics of quantification in Malay: A government and binding approach*. Doctoral dissertation. Universiti Sains Malaysia, Penang, Malaysia.
- Ross, John. R. (1967). *Constraints on variables in syntax*. Doctoral dissertation, MIT.

- Ross, John (1969). Guess Who. *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*, ed. Robert Binnick, Alice Davison, Georgia Green, and Jerry Morgan, 252-286. Chicago: Chicago Linguistic Society.
- Rutkowski, Paweł (2002). Noun/pronoun asymmetries: Evidence in support of the DP hypothesis in *Polish*. *Jezikoslovlje* 3: 159-170.
- Rutkowski, Paweł and Hanna Maliszewska (To appear). On prepositional phrases inside numeral expressions in Polish. *Lingua*.
- Ruys, Eddy (1992). *The scope of indefinites*. Doctoral dissertation, Utrecht University.
- Sabel, Joachim (2000). Partial *wh*-movement and the typology of *wh*-questions. *Wh-scope marking*, ed. Uli Lutz, Gereon Müller & Arnim von Stechow, 409–446. Amsterdam: John Benjamins.
- Saddy, Douglas (1991). WH-scope mechanisms in Bahasa Indonesia. *MIT Working Papers in Linguistics* 15: *More Papers on Wh-Movement*, ed. Lisa Cheng and Hamida Demirdache, 183-218. Department of Linguistics and Philosophy: MITWPL.
- Safir, Kenneth (1993). Perception, selection, and structural economy. *Natural Language Semantics* 2: 47-70.
- Saito, Mamoru and Keiko Murasugi (1990). N'-deletion in Japanese: A preliminary study. *Japanese/Korean Linguistics* 1, ed. Hajime Hoji, 285-301. Stanford: CSLI Publications.

- Salles, Heloisa Maria Moreira-Lima (1997). *Prepositions and the syntax of complementation*. Doctoral dissertation, University of Wales, Bangor
- Sato, Yosuke (2007). P-stranding generalization and Bahasa Indonesia: A myth? *Snippets* 16: 17-18.
- Sato, Yosuke (2008). Reduplication asymmetries in Bahasa Indonesia and the organization of the lexicon-syntax interface. Ms., University of Arizona, Tucson.
- Sato, Yosuke (In press a). The distribution of the active voice morphology in Javanese and *v*P phases. *Proceedings of the 2007 Western Conference on Linguistics*. University of California, San Diego: Department of Linguistics.
- Sato, Yosuke (In press b). The distribution of the active voice morphology in Bahasa Indonesia and *v*P phases. In: Yoshiaki Kaneko, Akira Kikuchi, Daiko Takahashi, Yoshiki Ogawa, and Etsuro Shima (eds.) *The State of the Art in Linguistic Research, Interface of Form and Meaning: Festschrift for Professor Masaru Nakamura*. Tokyo: Kaitakusha.
- Sato, Yosuke (In press c). Spelling-out prosodic domains: A multiple spell-out account. *InterPhases: Phase-Theoretic Investigations of Linguistic Interfaces*, ed. Kleanthes Grohmann. Oxford: Oxford University Press.
- Sato, Yosuke (In press d). The denotation and morphosyntax of bare nouns in Javanese: Against the nominal mapping parameter. *Proceedings of the 2007 Mid-America Linguistics Conference: Kansas Working Papers in Linguistics* 30. University of Kansas: Linguistic Graduate Student Association.

- Sato, Yosuke (Forthcoming a). Sluicing in Bahasa Indonesia, P-stranding, and interface repair. *Proceedings of the Thirty-Fourth Annual Meeting of the Berkeley Linguistics Society*. Department of Linguistics: Berkeley Linguistics Society
- Sato, Yosuke (Forthcoming b). The denotation and morphosyntax of NPs in Javanese: A problem for the nominal mapping parameter. *Proceeding of the First Meeting of the Arizona Linguistics Circle*. University of Arizona: Arizona Linguistics Circle.
- Sato, Yosuke (In preparation). A choice function approach to English Jeopardy game questions. Ms., University of Arizona, Tucson.
- Sato, Yosuke and Bradley McDonnell (In press). Reduplication in Indonesian and the lexicalist hypothesis. *Proceedings of the Thirty-Third Annual Meeting of the Berkeley Linguistics Society*. University of California, Berkeley: Department of Linguistics [with Bradley McDonnell]
- Sato, Yosuke and Dwi Hesti Yuliani (In press). *Wh*-in-situ in Bahasa Indonesia and choice function. *Proceedings of the 2007 Mid-America Linguistics Conference: Kansas Working Papers in Linguistics* 30. University of Kansas: Linguistic Graduate Student Association.
- Schmitt, Cristina and Alan Munn (1999). Against the nominal mapping parameter: Bare nouns in Brazilian Portuguese. *Proceedings of NES 29*, ed. Pius Tamanji, Masako Hirotani, and Nancy Hall, 339-353. University of Massachusetts, Amherst: GLSA.
- Schmitt, Cristina and Alan Munn (2002). The syntax and semantics of bare arguments in Brazilian Portuguese. *Linguistic Variation Yearbook* 2: 185-216.
- Seiter, William (1980). *Studies in Niuean syntax*. New York: Garland.

- Sells, Peter (1984). *Syntax and semantics of resumptive pronouns*. Doctoral dissertation, University of Massachusetts, Amherst.
- Sie, Ing Djiang (1988). *The syntactic passive in Bahasa Indonesia*. Doctoral dissertation, University of Amsterdam.
- Siegel, Dorothy (1973). *Topics in English morphology*. Doctoral dissertation, MIT.
- Sneddon, James (1996). *Indonesian: A comprehensive grammar*. Routledge: London.
- Soh, Hooi-Ling (1998). Certain restrictions on A-bar movement in Malay. *UCLA Occasional Papers in Linguistics 21: Proceedings of the third and fourth meetings of the Austronesian Formal Linguistics Association (AFLA)*, ed. Matthew Pearson, 295–308. Department of Linguistics: UCLA.
- Speas, Magaret (1990). *Phrase structure in natural language*. Dordrecht: Kluwer.
- Speas, Margaret (1994). Null arguments in a theory of economy of projection. *University of Massachusetts Occasional Papers in Linguistics 17: Functional Projections*, ed. Ellena Benedicto and Jeffrey Runner, 179-208. University of Massachusetts, Amherst: GLSA.
- Spencer, Andrew (1991). *Morphological Theory*. Oxford: Blackwell.
- Stepanov, Arthur (2001). Late adjunction and minimalist phrase structure. *Syntax* 4: 94-125.
- Stepanović, Sandra (1999). Scrambling: Overt movement or base generation and LF movement. *Journal of Slavic Linguistics* 7: 305-324.

- Stepanović, Sandra (2008). P-stranding under sluicing in a non-P-stranding language? *Linguistic Inquiry* 39:179-190.
- Stowell, Timothy (1981). *Origins of phrase structure*. Doctoral dissertation, MIT.
- Stowell, Timothy (1982). Conditions on reanalysis. *MIT Working Papers in Linguistic 4: Papers on Syntax*, ed. Alec Marantz and Timothy Stowell, 245-269. Department of Linguistics and Philosophy: MITWPL.
- Stowell, Timothy (1989). Subjects, specifiers, and X-bar theory. *Alternative conceptions of phrase structure*, ed. Mark Baltin and Anthony Kroch. 232-262. Chicago: University of Chicago Press.
- Suñer, Margarita (1988). The role of agreement in clitic-doubled constructions. *Natural Language and Linguistic Theory* 6: 391-434.
- Svenonius, Peter (2004). On the edge. *Peripheries*, ed. David Adger, Cecile de Cat and George Tsoulas, 259-287. Dordrecht: Kluwer.
- Szabolcsi, Anna (1987). Functional categories in the noun phrase. *Approaches to Hungarian*, ed. Istvan Kenesei, 167-190. JATE: Szeged.
- Szabolcsi, Anna (1994). The noun phrase. *Syntax and Semantics 27: The syntactic structure of Hungarian*, ed. Ferenc Kiefer and Katalin Kiss, 170-274. San Diego, CA: Academic Press.
- Szabolcsi, A. and Frans Zwarts (1993). Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1: 235-285.

- Szczegielniak, Adam (2006). All sluiced up, but no alleviation in sight,” Ms., Harvard University.
- Takami, Ken-ichi (1988). Preposition stranding: Arguments against a syntactic analysis and an alternative functional explanation. *Lingua* 76: 299-335.
- Tanaka, Hidekazu (2007). Ellipsis under syntactic isomorphism. Ms., University of York.
- Tenny, Carol (1994). *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer.
- Thornton, Rosalind (1995). Referentiality and *wh*-movement in Child English: Juvenile D-linkuency. *Language Acquisition* 4, 139–175.
- Torrego, Esther (1984). More effects of successive cyclic movement. *Linguistic Inquiry* 14: 361-365.
- Toyoshima, Takashi (2004). *Wh*-Tideland. Paper presented at the Second International Theoretical East Asian Linguistic Workshop. National Tsing Hua University, Taiwan. June 12-13, 2004.
- Travis, Lisa (1999). A syntactician’s view of reduplication. In: Carolyn Smallwood and Catherine Kitto (eds.) *Toronto Working Papers in Linguistics: Proceedings of the Sixth Meeting of the Austronesian Formal Linguistics Association*, 312-331. Toronto: University of Toronto
- Travis, Lisa (2001). The syntax of reduplication. In: Min-Joo. Kim and Uri Strauss (eds.) *Papers from the Thirty-First Annual Meeting of the Northeastern Linguistic Society*. 455-469. Amherst, MA: GLSA.

- Travis, Lisa (2003). Reduplication feeding syntactic movement. In: Sophie Burelle and Stanca Somesfalean (eds.) *Proceedings of the 2003 Annual Conference of the Canadian Linguistic Association* 236-247. Montreal: UQAM.
- Trenkic, Danijela (2004). Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua* 114: 1401-1427.
- Trugman, Helen (2005). *Syntax of Russian DPs and DP-internal agreement phenomena*. Doctoral dissertation, Tel-Aviv University.
- Tsai, Dylan. W.-T (1999). *On economizing the theory of A-bar dependencies*. New York: Garland.
- Uriagereka, Juan (1998) *Rhyme and reason: An introduction to minimalist syntax*. Cambridge, MA: MIT Press.
- Uriagereka, Juan (1999). Multiple spell-out. *Working minimalism*, ed. Samuel Epstein & Norbert Hornstein, 251-282. Cambridge, MA: MIT press.
- Uriagereka, Juan (2002). *Derivations: Exploring the dynamics of syntax*. London/New York: Routledge.
- Vainikka, Anne (1993/1994). Case in the development of English syntax. *Language Acquisition* 3: 257-324.
- van Riemsdijk, Henk (1978). *A case study in syntactic markedness: The binding nature of prepositional phrases*. Dordrecht: Foris.

- van Riemsdijk, Henk (1998). Head movement and adjacency. *Natural Language and Linguistic Theory* 16: 633-678.
- Verhaar, Jolen (1989). Syntactic ergativity in contemporary Indonesian. *Serpih-Serpih Telaah Pasif Bahasa Indonesia* [Studies on Passives in Indonesian], ed. Purwo, Bambang Kaswanti, 200-281. Yogyakarta: Penerbit Kanisius.
- Voskuil, Jan (2000). Indonesian voice and A-bar movement. *Formal issues in Austronesian linguistics*, ed. Ileana. Paul, Vivianne Phillips, and Lisa Travis, 195-213. Dordrecht: Kluwer.
- Wagner, Michael (2005). *Prosody and recursion*. Doctoral dissertation, MIT.
- Wang, Chyan-An (2006). Sluicing and Resumption. Paper presented at the 18th Conference of the Northeast Linguistic Society. [To appear in *Proceedings of NELS 37*, University of Massachusetts, Amherst: GLSA]
- Wasow, Thomas (1977). Transformations and the lexicon. *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 327-360. New York: Academic Press.
- Wasow, Thomas (1979). *Anaphora in generative grammar*. Ghent: E. Story-Scientia.
- Watanabe, Akira (1992). Subjacency and S-structure movement of *wh*-in-situ. *Journal of East Asian Linguistics* 1: 255-291.
- Watanabe, Akira (2000). Feature copying and binding: Evidence from complementizer agreement and switch reference. *Syntax* 3: 159-181.

- Watanabe, Akira (2001). Wh-in-situ languages. *The handbook of contemporary syntactic theory*, ed. Mark Baltin and Chris Collins, 203-225. Oxford: Blackwell.
- Watanabe, Akira (2006). Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language and Linguistic Theory* 24: 241-306.
- Wexler, Kenneth and Rita Manzini (1987). Parameters and learnability in binding theory, *Parameter setting*, ed. Thomas Roeper and Edwin Williams, 41-76. Dordrecht: Reidel.
- Williams, Edwin (1981). On the notions “lexically related” and “head of a word.” *Linguistic Inquiry* 12: 234-274.
- Williams, Edwin (2007). Dumping lexicalism. In *The Oxford Handbook of Linguistic Interfaces*, Gillian Ramchand and Charles Reiss (eds.), 353-381. Oxford: Oxford University Press.
- Willim, Ewa (1998). On the DP-hypothesis in Polish, an article-less language. *Projections and mappings: Studies in syntax*, ed. Piotr Stalmaszczyk, 137-158. Lublin: Folium.
- Willim, Ewa (2000). On the grammar of Polish nominals. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels & J. Uriagereka, 319-346. Cambridge, MA: MIT Press.
- Winter, Yoad (1997). Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20: 399-467.
- Wolff, John, Dede Oetomo, and Daniel Fietkiewicz (1992). *Beginning Indonesian through self-instruction*, vol. 2-3. Ithaca, NY.: Cornell University Southeast Asia Program.

- Zagona, Karen (2007). "Adjunct theta-roles" and the configurational determination of roles. *Phrasal and clausal architecture: Syntactic derivation and interpretation*, ed. Simin Karimi Vida Samiian and Wendy K. Wilkins, 396-411. Amsterdam: John Benjamins.
- Zlantić, Larisa (1997). *The structure of the Serbian noun phrase*. Doctoral dissertation, University of Texas, Austin.
- Zubizarreta, Maria-Luisa (1998). *Prosody, focus, and word order*. Cambridge, MA: MIT Press.
- Zwart, Jan-Wouter (1993). Verb movement and complementizer agreement. *MIT Working Papers in Linguistics 18: Papers on Case and Agreement I*, ed. Jonathan Bobaljik and Colin Phillips, 297-340. Department of Linguistics and Philosophy: MITWPL.
- Zwart, Jan-Wouter (1997). *Morphosyntax of verb movement*. Dordrecht: Kluwer.