ABSTRACT

Title of Document: CONSTRAINTS AND MECHANISMS IN LONG-DISTANCE DEPENDENCY FORMATION.

Masaya Yoshida, Ph.D. 2006

Directed By: Professor Colin Phillips, Department of Linguistics

This thesis aims to reveal the mechanisms and constraints involving in long-distance dependency formation in the static knowledge of language and in real-time sentence processing. Special attention is paid to the grammar and processing of island constraints. Several experiments show that in a head-final language like Japanese global constraints like island constraints are applied long before decisive information such as verb heads and relative heads, are encountered. Based on this observation, the thesis argues that there is a powerful predictive mechanism at work behind real time sentence processing. A model of this predictive mechanism is proposed.

This thesis examines the nature of several island constraints, specifically Complex NP Islands induced by relative clauses, and clausal adjunct islands. It is argued that in the majority of languages, both relative clauses and adjunct clauses are islands, but there is a small subset of languages (including Japanese, Korean and Malayalam) where extraction out of adjunct clauses seems to be allowed. Applying

well-established syntactic tests to the necessary constructions in Japanese, it is established that dependencies crossing adjunct clauses are indeed created by movement operations, and still the extraction is allowed from adjuncts.

Building on previous findings, the thesis turns to the investigation of the interaction between real time sentence processing and island constraints. Looking specifically at Japanese, a head-final language this thesis ask how the structure of sentences are built and what constraints are applied to the structure building process. A series of experiments shows that in Japanese, even before high-information bearing units such as verbs, relative heads or adjunct markers are encountered, the structural skeleton is built, and such pre-computed structures are highly articulated. It is shown that structural constraints on long-distance dependencies are imposed on the pre-built structure. It is further shown that this finding support the incrementality of sentence processing.

CONSTRAINTS AND MECHANISMS IN LONG-DISTANCE DEPENDENCY FORMATION

By

Masaya Yoshida

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:
Professor Colin Phillips, Chair
Professor Norbert Hornstein
Professor Robert De Keyser
Professor Jeffrey Lidz
Professor Amy Weinberg

© Copyright by
Masaya Yoshida
2006

Acknowledgements

At last, I have reached the stage where I can write the "Acknowledgements" for my thesis. Since the very beginning of my graduate years, I wanted to get to this point so badly (so many times, I thought I could not), and I believed that writing this part of the thesis would be the great relief (and indeed it is!) Now that I am given a chance to write this part and I realize an enormous number of people to who I owe a lot for writing this dissertation and surviving through my graduate years. I am most pleased by the fact that I have so many good friends, colleagues, mentors, or whatever you may call them, who had been there for me whenever I needed them. Here, I would like to express my deepest gratitude to those who helped me come this far.

First and foremost, I would like to thank my advisor Colin Phillips. I cannot express how thankful I am to him in this limited space. I am sure that I could not have reached this point without him. Since I wrote to him the first e-mail message ten years ago, he is the one who has always been there for me. For me, he is the best mentor without a doubt, and I am so grateful that I have a chance to work with him. After every meeting with him, I had a sense that my work was getting better and better. He always showed me how to see my scattered and seemingly trivial ideas from various different points of view, how to give them shape and what the real questions that I should be asking were. The many discussions with him over years at Maryland, which I greatly enjoyed, form the very foundation of this thesis.

I would like to thank the members of my thesis committee, Norbert Hornstein, Amy Weinberg, Jeffrey Lidz and Robert DeKeyser. It was a great pleasure that I could attend courses taught by Norbert and I could discuss various topics in

linguistics with him. Through his courses and several discussions I had with him, I learned a lot about how to construct good syntactic arguments. He is also always accessible and helpful. Whenever I brought small ideas in syntax and psycholinguistics to him, he came up with brilliant new ideas that I could never think about (he is like an alchemist). The syntax part of my thesis has been greatly improved thanks to the discussions with him and his razor sharp comments and criticisms. Having Amy as one of the members of my thesis committee was critical. Whenever I had trouble making progress she was the one who I would go talk to. She was always encouraging, cheerful and showed me that I pointed in the right direction. Thanks to her cheerful mood, I greatly enjoyed writing this thesis. I was really lucky that Jeff kindly accepted to be one of my committee members. Even though I gave him my huge defense draft at the latest moment, he read through all the details of it and gave me important comments. The discussions with him were also really valuable for writing both psycholinguistics and syntax portions of this dissertation. I am also grateful to Robert DeKeyser. He also received the defense draft at the final stage. Even though he has just 10 days or so to read through the draft, and he is not a specialist in the topic explored in this thesis, he read the draft and gave me significant comments. I thank him for his warm encouragements through e-mails too.

Even though other professors in the department could not be on my thesis committee, they played crucial roles at every stage of my thesis writing and completing the Ph.D. requirements. In writing this thesis and developing my skills in syntax, I profited from Howard Lasnik so greatly. Through his courses and discussions with him, I established the basis of my syntax skills. Even though, he

could not be on my thesis committee, a great part of the syntax chapter would not exist without his help and advice. Juan Uriagereka has contributed a great deal of the development of my syntax skills too. His first syntax course (the very first course in syntax that I took at Maryland) provided me many of the tricks and tools that I used in the syntax chapter of this dissertation. I really enjoyed his course as well as discussions with him. I also would like to show my gratitude to Jacek Witkos who was the other teacher of my first maryland syntax course. Thanks to his positive and encouraging attitude, the first year of my graduate life was really enjoyable. For me the course requirements were torturously hard in the first year. However, his encouraging mood always helped me to get through the hardest times. I will not forget the days we played soccer together. I also benefited from discussion of this material with Tonia Bleam. Her enthusiasm and encouragement made talking to her really an enjoyable experience. Unfortunately we didn't have much time to discuss my thesis but I enjoyed just talking to her very much. I am grateful to other professors, although we couldn't have many chances to talk about my work. I would like to thank Margaret Antonisse, Stephen Crain, Bill Idzardi, Paul Pietroski, David Poeppel, Phillip Resnik, Rozz Thornton, Alexander Williams and Andrea Zukowski. All of them made my graduate life really enjoyable.

Maryland is a unique environment where there are always a lot of Japanese linguists. This high concentration of the density of Japanese population was always an essential part of my life at Maryland. Among the Japanese gang, I owe tremendously huge debts to Takuya Goro and Tomohiro Fujii. Takuya has been one of my oldest linguistics-friends and *sempais* (senior friends) for me. Discussing things

about linguistics with him always benefited me. Sometimes his super sharp comments killed my premature ideas, but discussing linguistics with him for years has been really good training. Aside from linguistics, he was and he is always a good friend and good *sempai*. I am indeed really lucky to have him as a friend. I owe him a lot. As well as Takuya, the presence of Tomohiro Fujii is really important for me. As a friend, as a roommate and as a teacher, I learned a lot of things from him. I really enjoyed ever lasting discussion on linguistics at the apartment (sometime we stayed up over night!!) very much. He was always generous, patient and accessible for casual chit chatting as well as serious linguistics discussion. I am most grateful having Takuya and Tomo as friends and as (unofficial) mentors.

I would like to say thanks to Chizuru Nakao, Hajime Ono, Akira Omaki and Kaori Ozawa too. I greatly enjoyed working with Chizuru and Hajime as well as casually chit chatting with them. It was really a valuable experience for me to work with them. While, I did not have the chance to work with Kaori and Akira, talking about linguistics or talking about casual topics over coffee or beer with them was really a valuable time for me. I am really fortunate to have them as coauthors as well as close friends.

If Sachiko Aoshima and Mitsue Motomomura and Hirohisa Kiguchi had not been at Maryland, my graduate life and my thesis itself would have been so different. The impact of Sachiko's work on this thesis is too obvious to even mention. The psycholinguistic part of this thesis has been built up on what Sachiko has been doing for years. If she were not there, I could not come up with even the most basic idea of this thesis. Additionally, I learned a lot of skills in psycholinguistics from her. I am

really grateful for her generosity and patience. Her cheerful mood and encouragements made discussing linguistics with her an enjoyable and important experience for me. I would like to show my deep gratitude to Mitsue Motomura. From the first time I met her, Mitsue was really friendly and helpful. In the first few weeks or so in Maryland, she did not hesitate using her time to help me out. Without her generous accommodation, the first part of my graduate life would have been much harder. Discussing linguistics with her was also a great joy. She read many of my badly written manuscripts, and she was really patient and generous to give me comments and suggestions. I received a lot of input from her. Hirohisa Kiguchi is another *sempai* who helped me a lot at various stages in my life at Maryland. Whenever I had a hard time, he was always willing to help me. Also through showing his papers, he taught me how to do syntax, and how to develop an idea. The input from him is also a crucial ingredient for me.

I would also like to show my gratitude to my classmates, Robert Fiorentino, Utako Minai, and Avi Derrow. The first year of the graduate life is really hard. However, I was lucky enough to have them as my classmates. The weekly discussion group that we had for the syntax courses helped me very much to survive through the first year. Without their help, and friendship, I do not think I could come to this far. Aside from linguistics, Rob and Utako are really good friends. Going out to drink with them was one of the most enjoyable times for me through all the graduate years.

I am glad to have Scott Fults as my friend. He is also one of my friends whose office I visited from time to time and spent a lot of time discussing linguistics and other many things. Here I would also like to say thanks to his family and friends in

Michigan who provided me with a warm family-like environment: Danielle Hayot, Frankie Jeanne Hayot Fults, Fran Fults, Sarah Fults, Bill Fults, Sandy Fults, Ruth Anna Fults, Dave Rigo, Mary and A.J. Casper.

For friendship and great company over the years, many thanks to Nina Kazanina and Leticia Pablos. They showed me constant friendship and warm encouragement. I really appreciate the time we shared with them (at the office or at the bar). Without their encouragement this dissertation would not exist. I would like to show my special thnks also to Andrea Gualmini and Luisa Meroni who also showed me constant and long-lasting friendship.

I am indebted to all of my fellow students, post-docs, RAs, visitors and people at University of Maryland. Special thanks go to, Diogo Almeida, Tim Hunter, Ellen Lau, Phillip Monahan, Ivan Ortega Santos, John Sprouse, who were willing to spend time to chit chat with me and encouraged me. Speicifically, working with Diogo was really an enjoyable and valuable experience for me. Dave Booth deserves a special mention. I was really happy to have him as my friend and as my roomamte. His support and encoragement are really essential for my life in Maryland. I would like to say thanks also to Ann Adelson, Ilya Arakelyan, Hee-Don Ahn, Rebecca Bayer, Anita Bowles, Ilhan Cagri, Maria Chait, Pritha Chandra, Ray Cho, Joseph Conroy, Stacey Conroy, Maia Duguine, Brian Dillon, John Drury, Chris Dyer, Annie Gagliardi, Silvia Gennari, Sarah Goodman, Ana Gouvea, Lydia Grevenyova, Maximiliano Guimarães, Martin Hackl, Ryuichiro Hashimoto, So-One Hwang, Yu Iwasaki, Julian Jenkins, Youngmi Jeong, Chunyuan Jin, Johanes Jurka, Eric Haag, Susannah Hoffman, Masaaki Kamiya, Ed Kenschaft, Maki Kishida, Jaiva Larsen, Huan Luo,

Shiti Malhotra, Yuval Marton, Jobin Mathew, Rebecca McKewon, Jairo Nunes, Miki Obata, Ilknur Oded, Lisa Pearl, Silke Urban Plesch, Ariane Rhone, Josh Riley, Cilene Rodrigues, Graziar Russo-Lassner, Lisa Sanders, Itziar San Martin, Michael Sabotin, Rachel Shorey, Usama Soltan, Clare Stroud, Eri Takahashi, Shin-Ichi Tanigawa, Heather Taylor, Graciela Tesan, Yi-Min Tien, Virginie Van Wassenhove, Matt Wagers, Jeff Walker, Yadong Wang and Ming Xiang. I also thank Kathi Faukingham and Robert Magee, the great secretaries at department of linguistics. Without their help, my graduate life would have been much much harder.

I owe a great debt to people outside of Maryland. I would especially like to thank Tetsuya Sano, Yukio Otsu, Masaru Kajita and Yasuhiko Kato. Without their help and encouragement studying at Maryland could not have happened. Especially, I owe great debts to Tetsuya who is my first teacher of linguistics and who cared me all the time. I learned a lot of things from him. I also thank him for his constant friendship. This work has been supported by the help from Nobutaka Tazawa. Whenever I come to Japan, he was the one who was willing to help me. I thank him for his friendship and help. I would like to show my special thanks to the friends from Hiroshima: Hiromu Sakai, Megumi Yoshimura and Chao Zhang. Their help for running experiments and valuable discussions with them are essential parts of this dissertation. I would also like to show my gratitude to those who helped me at various stages of my graduate life. Among all the people who helped me over the years, there are four who deserve special mention: Koji Hoshi, Edson Miyamoto, Asako Uchibori and Akira Watanabe. Whenever I needed help, they did not hesitate spending their time. Without their help, advice, and input, this dissertation would have been in much

different shape. My thanks also go to, among others, Sachie Akiba, Pranav Anand, Ken-Ichi Ariji, Angel Jesús Gallego Bartolomé, Cedric Boeckx, Željko Bošković, Suzanne Dikker, Chris Davis, Mika Endo, Sam Epstein, Janet Fodor, Aniela Improta França, Lyn Frazier, Naoki Fukui, Jon Gajewski, Ted Gibson, Arseni Gosev, Kleanthes Grohmann, Valentine Hacquard, Hiroko Hagiwara, Yurie Hara, Naomi Harada, Kaori Hayashi, Ken Hiraiaw, Tsuneaki Hirooka, Yuki Hirose, Masako Hirotani, Hajime Hoji, Natalie Hsu, Akira Ikeya, Atsu Inoue, Miwa Isobe, K. A. Jayaseelan, Kyle Johnson, Kohji Kamada, Ruth Kempson, Jieun Joe Kiaer, Hisatsugu Kitahara, Ivona Kucerova, Shinichiro Ishihara, Hironobu Kasai, Takaomi Kato, Masakazu Kuno, Miriam Lemry, Moti Lieberman, Yoshimi Maeda, Reiko Mazuka, Jason Merchant, Keiko Miyagawa, Satoru Muraoka, Michiko Nakamura, Kentaroo Nakatani, Mineharu Nakayama, Fumikazu Niinuma, Taisuke Nishigauchi, Masashi Nomura, Lance Nathan, Andrew Nevins, Reiko Okabe, Masaki Ohno, Koichi Otaki, Bun-Sik Park, Tsutomu Sakamoto, Mamoru Saito, Michiyuki Sato, Hiroyuki Shimada, Carmen Dobrovie-Sorin, Patrick Sturt, Koji Sugisaki, Shoichi Takahashi, Katsuo Tamaoka, Jun-Ichi Tanaka, Satoshi Tomioka, Shigeo Tonoike, Mieko Ueno, Ayumi Ueyama, Naomi Yamada, Hiroko Yamakido, Kyoko Yamakoshi, Hiroko Yamashita, Henny Yeung and Tomoko Yoshino.

My thanks go also to people outside of the world of linguistics, for their friendship. Thanks go to Makoto Imaizumi, Yoshito Okuyama, Mika Kato, Miki Tsukada, and Kanako Yano. I would like to show my deep gratitude, especially to Miki and Kanako for their constant friendship, help and encouragement.

Finally and most of all, I would like to thank my parents, Isao and Fukiko Yoshida, my brother Tomoya Yoshida, and my aunt Setsuko Yokota, for their never-failing love, support and encouragement. I thank my parents for letting me do whatever I wanted to. This thesis is dedicated to them.

Table of Contents

Acknowledgements ii
Table of Contents xi
CHAPTER 1. INTRODUCTION 1
1. Incrementality 1
2. Predictive Mechanism 10
2.1. Evidence for Strong Incrementality 13
2.1.1. Evidence for Strong Incrementality 1: Prediction of Relative Clauses 14
2.1.2. Evidence for Strong Incrementality 2: Prediction of Conditional Clauses 19
3. Toward the Nature of Predictive Mechanism 24
CHAPTER 2. SYNTAX OF RELATIVE CLAUSES AND ADJUNCT CLAUSES 30
1. Introduction 30
2. The Syntax of Relative Clauses 31
2.1. Island Sensitivity 33
2.2. Connectivity/Reconstruction Effects 39
2.3. Summary 42
2.4. Issues of Verbal Morphology within Relative Clauses 43
2.4.1. Nominative Genitive Conversion 43
2.4.2. Predicate Adnominal Form and the C-system 47
2.4.3. Relative Clauses and the C-system 50
2.5. Conclusion 50
3. The Syntax of Conditional Clauses in Japanese 51
3.1. Non-islandhood of Japanese Conditionals 51
3.1.1. Adjuncthood of Conditional Clauses: English 54
3.1.2. Japanese Conditionals as Adjunct Clauses 58
3.1.3. Summary 63
3.2. Conditional Clauses as Weak Islands 64
3.2.1. Diagnosing Weak Islands 64
3.2.2. *Mosi* as a Weak Island Inducer 71
3.2.3. Summary 87
3.3. The Distribution of *Mosi* 87
3.3.1. Japanese Nominal Conditionals 88
3.3.2. Some Basics of Nominal Conditionals 89
3.3.3. The Conditional Focus Link 91
3.3.3.1. The Syntax of the Clausal Pronoun *sore* in Japanese 91
3.3.3.1.1. Nakao & Yoshida (2005) on Pronominal Sluicing 91
3.3.3.2. Summary 103
3.3.3.3. *Sore* in Nominal Conditionals 103
3.3.3.4. *Sore* and *Mosi* 106

3.3.3.5. A Problem Remains 108
3.3.3.6. Toward the Licensing Condition of *Mosi* 110
3.3.3.6.1. Rizzi (1997) 110
3.3.3.6.2. Focus Constructions in Japanese: Hiraiwa & Ishihara (2001) 117
3.3.3.7. The Analysis of Nominal Conditionals 136
3.3.3.7.1. Further Properties of Nominal Conditionals 136
3.3.3.7.2. The Derivation of Nominal Conditionals 145
3.3.3.7.3. The Licensing Condition on Mosi 148
3.3.4. Conclusion 148
3.4. The Question of *Mosi* and Weak Island Effects 149
3.4.1. Weak Island Inducers 149
3.4.2. Mosi as a Scope Marker 152
3.4.2.1. The Basic Observations 152
3.4.2.2. Disjunction Scope in English and Focus Particles in Japanese 158
3.4.3. Intervention Effects 161
4. Approaches to Adjunct (Non)Islands 161
4.1. Adjunct Islands and Left/Right Asymmetry 166
4.1.1. Spanish and English 166
4.1.2. Scrambling from Conditionals and the Left/Right Asymmetry 168
4.2. A Parametric Study on Adjunct Islands 188
4.2.1. Languages that Allow Extraction out of Adjuncts 188
4.2.2. Parametric Features 193
5. Conclusions 199
CHAPTER 3. RELATIVE CLAUSE PREDICTION IN JAPANESE 200
1. Introduction 200
2. Necessity of extra-lexical knowledge in sentence processing 201
3. On the Prediction of Japanese Relative Clauses 204
4. The Processing of Japanese Relative Clauses 210
5. Experiment 1 219
6. Experiment 2 227
7. Relative Clause Islands in Japanese 237
8. Processing Filler-Gap Dependencies in Japanese 240
9. Filler-gap Dependencies and Relative Clause Islands 244
10. Experiment 1b 245
11. Experiment 3 261
12. General Discussion 272
12.1. Processing of relative clauses 272
12.2. Incremental structure building with a powerful predictive mechanism 275
13. Conclusion 277
CHAPTER 4. CONDITIONALS AND LONGER DEPENDENCY FORMATION 278

1. Introduction 278
2. On the Motivations for the Longer Dependency Bias 279
3. The Source of the Bias for Longer Filler-Gap Dependencies in Japanese ... 282
3.1. Longer Dependency Bias 282
3.2. Manipulating the Trigger 287
3.3. The Japanese Filled-Gap Paradigm and the Motivation for the Longer Dependency Bias 289
3.3.1. Aoshima et al., (2004) and Aoshima (2003) 289
3.3.2. Manipulating the Features on the Fronted Phrase 295
3.4. Experiments 297
3.4.1. Experiment 1 297
3.4.2. Experiment 2 307
3.5. General Discussion 312
4. Conditionals and Long-Distance Dependency Formation 315
4.1. The Question 315
4.2. Japanese Conditionals from a Real Time Perspective 316
4.3. Experiments 324
4.3.1. Methodology and Predictions 324
4.3.2. Experiment 3: Sentence Fragment Completion Task 327
4.3.3. Experiment 4: Self-Paced Reading Task 330
4.3.4. Experiment 5: Self-Paced Reading Task: Japanese Filled Gap Paradigm 336
4.3.5. Experiment 6: Sentence Fragment Completion Task 347
5. General Discussion: The Silhouette of the Predictive Mechanism 359
5.1. The Nature of the Predictive Mechanism (or the Things that the Parser should be Able to do) 359
5.2. Reconsidering the Processing of Relative Clause Islands (or the Things that the Parser should be Able to Foresee) 363
CHAPTER 5. ON THE NATURE OF PREDICTIVE MECHANISMS 366
1. Introduction 366
2. The Requirements for the Predictive Mechanism 366
2.1. Subjacency Approaches 371
2.2. Morphology-Based Approaches 378
2.3. On the Requirements for the Predictive Mechanism: Summary 380
3. Perspectives on Online Sentence Processing: A Quest for Incrementality ... 381
3.1. Incrementality 381
3.2. Mechanisms for Sentence Processing 386
3.2.1. Purely Bottom-Up Strategy and Their Problems 387
3.2.2. Purely Top-Down Strategies and Their Problems 390
3.2.3. The Left-Corner Strategy 393
3.2.4. SPARSE (Schneider 1999): Incrementality and Flexibility .. 398
4. How to Block the Longer Dependency? 408
4.1. Dealing with the Indirect Cue 408
4.2. Enhancing the Predictive Component 416
4.3. Dealing with the Direct Cue 426
5. Conclusions 436

CHAPTER 6. GENERAL CONCLUSIONS 437
Bibliography 441

CHAPTER 1. INTRODUCTION

1. Incrementality

The aim of this thesis is to show that sentence processing is strongly incremental in the sense that the structural commitments are made and interpretations become available on a word-by-word basis and that strong incrementality is made possible by a powerful predictive mechanism. Throughout this dissertation I will show various lines of supporting evidence for strong incrementality and propose a mechanism that can derive it. In this Introduction I would like to discuss issues of incrementality and briefly summarize the themes and organization of this dissertation.

Over the history of sentence processing research it has been made clear that sentences are processed incrementally, in the sense that constituent structures are constructed and interpretations of constituents become available on a word-by-word basis (Aoshima et al. 2004; Bader and Lasser 1994; Frazier 1987; Inoue 1991; Marslen-Wilson 1973; Miyamoto and Takahashi 2002; Miyamoto 2002; Sturt and Crocker 1995; Sturt and Lombardo 2005).

Initial support for incrementality came from so-called garden path sentences (Bever 1970, among many others). It was suggested that garden path effects show that when the parser faces an ambiguous input, it does not delay its decision until crucial disambiguating information becomes available, and thus the parser makes a commitment to an initial analysis that may turn out to be an incorrect analysis. When the parser is garden-pathed, a consciously or experimentally detectable surprise effect occurs at the point where the initial misanalysis is resolved. Let us take an example from English.

(1) The boat floated down the river sank.

The string before the main verb *sank* in (1) can be parsed as a simple independent clause. When the main verb is encountered, a surprise effect takes place, and readers experience difficulty in reading. If the parser waits until a crucial element that can confirm the structure of the sentence, this garden path effect is not expected. The fact that readers experience the surprise effect suggests that the parser commits itself to the initial main clause analysis before the main verb is encountered, and the verb *sank* forced the parser to reanalyze the structure from the main clause structure to the reduced relative clause structure, resulting in the garden-path effect.

However, it has been pointed out that Garden Path sentences do not necessarily support the incrementality of sentence processing. To account for the garden path phenomena, many researchers have emphasized the importance of licensing heads. For example, the theory proposed by Abney (1987) or Pritchett (1992), the so-called head-driven models, explained the garden path phenomena in a way that the sentence processing is driven to satisfy the requirements of the licensing heads (typically verbs) at each step of processing. Under this approach, an example like (1) creates garden path effect because the materials before the verb *sank* can satisfy the grammatical requirements that the verb *floated* such as thematic requirements. As a result the string is understood as a simple sentence initially. Under strictly head-driven accounts of garden path phenomena, it is assumed that the structure is built when the information of the licensing head becomes available. Thus,

under this view, the sentence processing is not necessarily incremental, and garden path phenomena can be explained by head-driven mechanisms. This means that garden path effect is not necessarily a support for incremental model of sentence processing.

Head-driven strategies imply that in a head-final language such as Japanese the processor delays structure building until the final word of each constituent. In other words, head-driven theories predict that input materials are not integrated into a syntactic structure before the licensing head becomes available in the input (Mazuka and Lust 1988, 1990; Mazuka 1991; Pritchett 1991b, 1992a, 1992b among others). However, there are various pieces of evidence against this particular view either from head-final languages like Dutch (Frazier 1987) or Japanese (Aoshima et al. 2004; Inoue 1991; Inoue and Fodor 1995; Mazuka and Itoh 1995; Miyamoto 2002) or even from a head-final languages like English (Sturt and Lombardo 2005). Let us take a look at some examples from Japanese.

Inoue (1991) cites the following example from Japanese.

(2) Brown-ga White-ni ringo-o tabeta inu-o ageta.
B-nom W-dat apple-acc ate dog-acc gave
"Daniel gave Paul the dog which ate the apple."

Japanese native speakers experience some difficulty dealing with this sentence when they reach the verb *tabeta* "ate". One of the widely accepted interpretations of this garden path effect is the following. The garden path effect is not expected if the

parser delays the structure building until it encounter the verb. When the embedded verb "ate" becomes available, its argument structure information also becomes available. Thus, the parser recognizes that the verb "eat" does not take a dative argument. The difficulty indicates that the parser commits an analysis before it encounters the embedded verb. Based on the case-particles that each NP bears, the three NPs are initially postulated as coarguments of the same clause. When the embedded verb is encountered, it becomes clear that the initial analysis is failed, resulting in the garden path effect.

Inoue's argument is based on native speakers' intuitive judgments, but some experimental studies support the same point too. Miyamoto (2002) found that the upcoming relative clause structure could be facilitated when a sentence contains a sequence of NPs with the same Case Markers. He examined the following pair of sentences.

(3) a. Ofisu-de shokuin-ga kakaricho-o [$_{RC}$t$_1$ ocha-o
office-at employeee-nom manager-acc tea-acc
dasita] **josei-ni** teineini shookai-sita.
served woman-dat politely introduced
"At the office, the employee politely introduced the manager to the woman who served the tea."

b. Ofisu-de shokuin-ga kakaricho-ni [$_{RC}$t$_1$ ocha-o
office-at employeee-nom manager-dat tea-acc
dasita] **josei-o** teineini shookai-sita.
served woman-acc politely introduced
"At the office, the employee politely introduced the woman who served the tea to the manager."

Miyamoto makes two important observations. First, when there are two accusative NPs in a sentence as in (3a), the second accusative NP creates a disruption. Second, in (3a), the relative head noun was read more easily than in (3b). According to Miyamoto these two observations are not explained by head-driven approaches. Under the head-driven approaches, the disruptive effect of the second accusative NP is not predicted because these NPs come into the input before any of the verbs and thus they should not be processed. Furthermore, head-driven approaches do not predict the facilitation effect at the relative head position in (3a). In both of the conditions in (3), everything is the same except for the accusative NPs. Thus, the facilitation effect cannot be attributed to any other factors than the presence of the two accusative NPs in (3a).

Japanese has the so-called Double Accusative Constraint (Harada 1973a, 1974; Hiraiwa 2002; Kuroda 1965 among many others). Simply put, the constraint excludes a clause containing two accusative NPs. In (3a) the parser has to insert a clause boundary between the two accusative NPs in order to analyze the string grammatically because of the Double Accusative Constraint. Once a clause boundary is inserted, an embedded clause structure is created and it becomes clear that the

subject position in the embedded clause is empty. Based on this information, the parser can predict the upcoming relative clause structure, resulting in the facilitation of the relative head position.

According to this account, the accusative NPs before the embedded verb must be processed in advance of the embedded verb. Thus this finding is not compatible with head-driven approaches, and supports incremental structure building.

Finally, let us review one more study that supports the strong incrementality. Recently, Aoshima, Yoshida & Phillips (2006) show that a hierarchical structural relation such as c-command is established before the verb is encountered. First, they observe that in Japanese a pronoun corefers with a c-commanding antecedent, such as the nominative subject NP in (4a), but if the pronoun is not c-commanded by its antecedent as in (4b), coreference is less acceptable. When the dative NP containing the pronoun is scrambled and precedes the antecedent as in (4c), the coreference is acceptable, even though the pronoun precedes the antecedent. (4c) is acceptable because the pronoun is c-commanded by the antecedent at its original position.

(4) a. C-commanded

Daigakusei$_1$-ga kare$_1$-no tomodachi-ni
college-student-nom he-gen friend-dat
okane-o kasita.
money-acc lent
'The college student lent his friend some money.'

b. Not C-commanded

??Kare$_i$-no tomodachi-ga daigakusei$_i$-ni

he-gen friend-nom college-student-dat

okane-o kasita.

money-acc lent

'His friend lent the college student some money.'

c. Scrambled

[Kare$_1$-no tomodachi-ni]$_2$ daigakusei$_1$-ga t_2

he-gen-friend-dat college-student-nom

okane-o kasita.

money-acc lent

'The college student lent his friend some money.'

They took advantage of this c-command requirement on pronominal coreference to investigate whether the parser establish c-command relation before the verb is encountered. They tested the following types of sentences. In (5a) and (5b), the dative NP containing the pronoun *kare* 'he' is scrambled, and the gender of the first nominative subjects is manipulated. In (5a), it is female, *oba* 'aunt' and in (5b), it is male, *oji* 'uncle'. On the other hand, in (5c) and (5d), the NP containing the pronoun is the first nominative subject, and thus it is not scrambled. The gender of the second nominative subjects is manipulated in the same way.

(5) a. Scrambled Mismatch

… [[*Kare*-no dono kodomo-ni]$_1$ **oba**-ga t_1 obentoo-o
he-gen which child-dat **aunt**-nom lunch-box-acc
watasita-ka]$_2$ titioya-ga t_2 oboeteita…
handed-Q father-nom remembered.

"… the father remembered to which of his children the aunt passed a lunch box."

b. Scrambled Match

… [[*Kare*-no dono kodomo-ni]$_1$ **oji**-ga t_1
he-gen which child-dat **uncle**-nom
obentoo-o watasita-ka]$_2$ titioya-ga t_2
lunch-box-acc handed-Q father-nom
oboeteita…
remembered.

"… the father remembered to which of his children the uncle passed a lunch box."

c. Unscrambled Mismatch

… [[*Kare*-no dono kodomo-ga] **oba**-ni
he-gen which child-dat **aunt**-nom
obentoo-o watasita-ka]$_2$ titioya-ga t_2
lunch-box-acc handed-Q father-nom
oboeteita…
remembered.

"… the father remembered which of his children handed a lunch box to the aunt."

d. Unscrambled Match

… [[*Kare*-no dono kodomo-ga] **oji**-ni obentoo-o
he-gen which child-dat **aunt**-nom lunch-box-acc
watasita-ka]$_2$ titioya-ga t_2 oboeteita…
handed-Q father-nom remembered.

"… the father remembered which of his children handed a lunch box to the uncle."

Testing these four types of sentences using self-paced reading task, they found that the first nominative NP is read slower in scrambled mismatch condition than in scrambled match condition. However, the comparison between the unscrambled mismatch condition and the unscrambled match condition did not show significant difference in reading time.

They base their argument on the conclusions from the previous studies on the processing of backward anaphora in English that when an anaphor is encountered, the parser initiates an attempt finding its antecedent at the earliest grammatically sanctioned position in the sentence, and if the potential antecedent is mismatched in gender with the pronoun there is surprise effect so-called the Gender Mismatch Effect (van Gompel & Liversedge 2003). They argue that their finding parallels the Gender Mismatch Effect in English. When the pronoun is encountered, the parser starts searching for its antecedent. In the scrambled conditions, the first nominative NPs are potential antecedents for the pronouns in the scrambled NPs. Therefore there was a

surprise if the gender of the potential antecedent was mismatched. On the other hand, the first dative NPs in unscrambled conditions are not grammatically possible antecedents for the pronouns contained in the nominative subjects. Thus there was no surprise.

For our discussion, the most important point in their finding is that the parser calculated c-command relation before the verb becomes available. In other words, even before the verb comes into the input, the parser builds hierarchical syntactic structure. As we can see, the gender mismatch effect is detected at the first nominative NP position in the scrambled conditions that is the position before the verb becomes available. Furthermore, the coreference relation between pronouns and their antecedents in Japanese are sensitive to c-command as we have seen. Thus, the gender mismatch effect that they observed indicates that the parser calculated the coreference relation between the pronoun and its antecedent (c-command relation) before the verb becomes available.

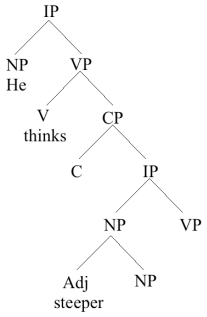
Results of these studies and many others suggest that human sentence processing is strongly incremental, in which input materials are assembled into a syntactic structure from left-to-right without a delay.

2. Predictive Mechanism

Even though various experimental studies have suggested that human sentence processing is strongly incremental, there are also various challenges to the incremental structure building. Lombardo and Sturt (2002) point out an example that creates a potential difficulty to the strong incremental structure building.

(6) a. He thinks [$_{CP}$[$_{IP}$[$_{NP}$[$_{Adj}$ steeper] prices] have come about because …]]

b.



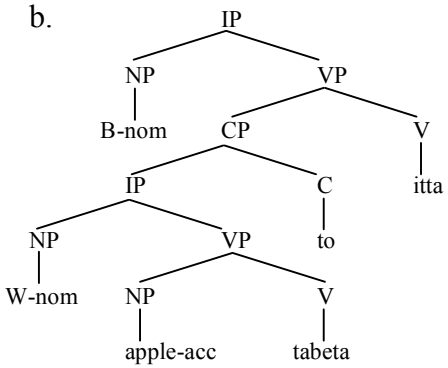
The example (6) contains an attributive adjective that modifies the subject NP of the embedded clause. When the adjective *steeper* is encountered during online sentence processing, a strongly incremental parser incorporates the adjective into the current representation without waiting for other lexical items to become available. In order to do so, however, the parser has to build the structure of the NP and IP, which do not have overt heads at the point where the word *steeper* is encountered. Upon encountering the adjective, the parser has to project an NP node to host it. The IP node is to be projected too to host the predicted NP. Thus a strongly incremental parser has to predict an IP node based on the predicted NP node in order to accommodate the adjective into the structure. A challenge for a strong incremental parser is whether such complicated prediction can be done without a guide from the lexical information.

If, on the other hand, the structure building is not strongly incremental and if it is allowed for a certain delay, this example does not create any problem. Facing an example like (6) a delay parser can wait until the NP *price* becomes available to host the adjective in to the structure.

In this respect, Japanese head-final sentences create this type of challenge almost everywhere. The following simple example can illustrate this point.

(7) a. Brown-ga [$_{CP}$[$_{IP}$ White-ga ringo-o tabeta-to]]
B-nom W-nom apple-acc ate-comp
itta.
said.
"Brown said that White ate an apple."

b.



(7) contains a simple multi-clausal structure in which a VP is taking a CP as its complement. When the second nominative subject NP, *White-ga* is encountered, a strongly incremental parser has to build an embedded IP node, a CP node a VP node and the matrix IP node in order to incorporate two nominative subjects into a structure. In doing so, the parser has to commit to a multiple syntactic predictions. It has to project two IP nodes in order to host *Brown-ga* and *White-ga* as subject NPs. It has to project a CP node, because the structure contains *White-ga* is an embedded clause. It has to project a VP node because there must be a V that takes a CP as a subordinate clause.

It is obvious that this type of approach is risky because the structure that the parser builds before encountering lexical verbs can turn out to be incorrect. For example, even if there are two nominative NPs the sentence can be a simple clause with a nominative object in Japanese such as *Brown-ga White-ga kiraida* "Brown hates White". Thus it is easy for a strongly incremental parser to make a mistake.

A delay parser, on the other hand, does not need to commit such a complicated multi-step calculation in order to process a sentence like (7). Furthermore, the delay parser does not need to commit to a particular analysis before it encounters a crucial licensing head. What the delay parser has to do is to wait until lexical heads, such as a verb to come into the input, and build the structure using the information conveyed by the lexical heads. For example the lexical verb with complementizer attached can help the parser to recognize that the there is a embedded clause and the second nominative NP is the subject of the embedded clause and so on.

With such an ambiguity problem, for example, is a strongly incremental parser a plausible model for human sentence processing? Is there any evidence that suggest strong incrementality?

2.1. Evidence for Strong Incrementality

In this dissertation, I will show evidence that strongly supports fully incremental models of sentence processing. Based on these pieces of evidence, I will investigate what mechanism can capture the strongly incremental nature of sentence processing. Evidence for strong incrementality comes from the studies on two constructions in Japanese, relative clauses and conditional clauses.

The dissertation is organized as follows. In Chapter 2, the basic syntactic properties of relative clauses and conditional clauses will be summarized. In Chapter 3, experimental studies on Japanese head-final relative clauses are summarized. Chapter 4 turns to experimental studies on Japanese conditional clauses. In these two chapters, I will present evidence for strong incrementality. Finally in Chapter 5, I will propose a possible mechanism for a strongly incremental structure building.

2.1.1. Evidence for Strong Incrementality 1: Prediction of Relative Clauses

The first evidence for strong incrementality comes from the processing of head-final relative clauses in Japanese. As is widely recognized, Japanese head-final relative clauses normally do not have any overt markers corresponding to the relative pronouns in English. Because of this property, it is very difficult to determine whether a string of phrases is part of a relative clause or not. For example, during the online processing of the example in (8), it is very difficult for the parser to recognize that this sentence contains a relative clause before the verb or the relative head is encountered. There are several reasons for this. One obvious reason is that the same string that constitutes the relative clause in (8) can be an independent simple clause as in (9) because of the lack of unambiguous relative clause markers or because of the availability of empty arguments (Hirose 1999; Inoue 1991; Mazuka 1991; Yamashita et al. 1993; Yamashita 1995 among many others).

(8) Brown-ga[$_{NP}$[$_{RC}$ White-ga Blue-ni gap ageta] hon]-o
B-nom W-nom B-dat gap gave book-acc
yonda.
read
"Brown read the book that White gave to Blue "

(9) Brown-ga White-ni gap ageta.
B-nom W-dat gave
"Brown gave something to White."

In Chapter 3 of this dissertation, I will show that there are some elements that can potentially mark the beginning of the relative clause unambiguously. (10) is one of such examples.

(10) Brown-ga [$_{NP}$ 3-satu-no [$_{NP}$[$_{RC}$ White-ga Blue-ni
B-nom 3-cl$_{(book)}$-gen W-nom B-dat
ageta] hon]]-o yonda.
gave book-acc read
"Brown read three books that White gave to Blue."

(10) contains a genitive-marked numeral classifier that is located before the relative clause. The important property of this type of construction is that the genitive-marked classifier can only be associated with the head of relative clause. Normally genitive classifiers can be associated with its adjacent NP. However, in the example above, it can only be associated with the relative head because it is not semantically

compatible with its adjacent NP, *White-ga*. This type of long-distance association of classifier and its host NP, and the semantic incompatibility between the classifier and its adjacent NP only take place in a complex NP environment and typically in relative clause environment. Thus, the semantic incompatibility between the genitive-marked classifier and its adjacent NP can unambiguously mark the beginning of the relative clause.

During online sentence processing, if the semantic incompatibility between the classifier and its adjacent NP is encountered, and if the parser can recognize that the semantic incompatibility is possible only in the relative clause environment, it is possible that the parser can predict the upcoming relative clause structure. Three experiments in Chapter 3 will show that this is indeed the case. The following two types of sentences were compared.

(11) a. Classifier Mismatch

Brown-ga 3-satu-no gakusee-ga katta

B-nom [$_{NP}$3-cl$_{(book)}$-gen [$_{NP}$[$_{RC}$ student-nom bought]

hon-o yonda

book]-acc read

"Brown read three books that the student bought."

b. Classifier Match

Brown-ga [$_{NP}$[$_{RC}$[$_{NP}$3-nin-no gakusee]-ga katta]

B-nom 3-cl$_{(person)}$-gen student-nom bought

hon]-o yonda.

book-acc read

"Brown read the book that the three students bought."

The first example contains a relative clause with semantic incompatibility between the classifier and its adjacent NP. It is called the classifier mismatch condition. In this case, as the brackets and the translation show that the classifier can only be associated with the relative head. The second example, on the other hand, contains a relative clause with the classifier that is semantically compatible with its adjacent NP. In this case, the classifier can be associated with its adjacent NP, and is not necessary to be associated with the relative head even if the relative head can host the classifier. This condition is called the classifier match condition.

A self-paced reading experiment is conducted and it is shown that the embedded relative clause verb is read more easily in the classifier mismatch condition than that of the classifier match condition. This result suggests that the parser can predict the upcoming relative clause structure by means of the classifier mismatch. Because the upcoming relative clause structure is predicted, the relative clause verb, which has a special morphology uniquely licensed in relative clauses and other complex NP environment (see Chapter 2 for details of this point), is read more easily. Furthermore, the result suggests that the relative clause structure is built before the embedded verb or the head of the relative clause is encountered. Thus this finding is fully compatible with strongly incremental parsing models that build the structure before encountering rich information bearing units like the verbs or the head of the relative clause.

In chapter 3, we will also see evidence that the predicted structure of relative clause is rich enough to block the long-distance dependency. This will be shown through a study on the island effect induced by relative clauses during online processing.

Japanese relative clauses like those in other languages are islands (Saito 1985). Thus, an overt movement like scrambling cannot escape the relative clause island as in the example (12a).

(12) a. *Dono-gakusee-ni$_1$ [Brown-wa [$_{NP}$[$_{RC}$ White-ga t_1
which-student-dat B-top W-nom
ageta] hon]-o yonda-no]?
gave book-acc read-Q
Lit. "Which student did Brown read the book that White gave the book to?"

b. ^{ok}Dono-gakusee-ni$_1$ [Brown-wa[$_{CP}$ White-ga t_1 hon-o
which-student-dat B-top W-nom book-acc
ageta-to] itta-no]?
gave-comp said-Q
"Which student did Brown say that White gave a book to?"

On the other hand, in the experimental studies by Aoshima (2004), it has been shown that when the sentence contains a scrambled phrase as in (12b), the scrambled phrase is preferentially associated with the embedded clause. Thus, during the online sentence processing, when the scrambled phrase *dono-gakusee-ni* "which student" is

encountered, the parser tries to interpret it in the embedded clause rather than in the matrix clause. Let us call this preference of the parser, the Longer Dependency Bias (LDB). Given this LDB and the islandhood of relative clauses as well as the possibility of predicting the upcoming relative clause structure, we can make following two predictions for the behavior of the parser. First, if the predicted relative clause structure does not represent islandhood, the fronted phrase is preferentially interpreted in the embedded relative clause, resulting in a violation of relative clause island constraints. On the other hand, if the islandhood of the relative clauses is represented, the LDB will be blocked. The results of the experiment will be shown in chapter 3 that indicate that LDB is blocked when the upcoming relative clause is predicted by means of the classifier mismatch.

An important observation in this study is the following. It will be shown that genitive-marked classifiers do not have any specific grammatical relation to the relative clause structure *per se*. Thus, the parser makes use of such "indirect" information to project the upcoming relative clause structure.

Taken together, these two pieces of evidence strongly suggest that the parser can build rich representation of the relative clause structure upon encountering the classifier mismatch before the information from the verb or the head of the relative clause becomes available.

2.1.2. Evidence for Strong Incrementality 2: Prediction of Conditional Clauses

In chapter 4, I will present another piece of evidence for strong incrementality. In this chapter we will look at Japanese conditional clauses.

The detailed descriptive syntactic studies on conditional clauses in chapter 2 will reveal three important properties of Japanese conditional clauses. First, Japanese conditional clauses can be introduced by so-called conditional adverbs. One such conditional adverb is *Mosi*. An important property of *Mosi* is that it can be licensed only by conditional clauses. Second, conditional clauses are distinguished from other types of clauses by specific verbal morphology. Typically the conditional suffix *-ra*, which is understood as a complementizer, is attached to the verb. Finally, conditional clauses are not islands even though they are adjunct clauses. An example of conditional clauses is in (13).

(13) Brown-wa [$_{\text{Cond}}$ *mosi* White-ga Blue-ni present-o
B-top mosi W-nom B-dat present-acc
ageta-**ra**] naki-dasu daroo
give-cond cry-start will
"Brown will cry if White gives a present to Blue."

These three properties of conditional clauses allow us to test strong incrementality too. In chapter 4, I will show the results of experiment that suggest that conditional verb is read more easily in the sentence containing *Mosi* compared to the one that does not contain *Mosi*. The following two types of sentences are compared in self-paced reading experiment.

(14) a. Mosi Condition

Brown-wa [$_{cond}$ mosi kawaii onnanoko-ga Blue-ni

B-top mosi pretty girl-nom B-dat

present-o ageta-ra] naki-dasu daroo.

present-acc give-cond cry-start will

"Brown will cry if the pretty girl gives a present to Blue."

b. Adverb Condition

Brown-wa [$_{cond}$totemo kawaii onnnanoko-ga Blue-ni

B-top very pretty girl-nom B-dat

present-o ageta-ra] naki-dasu daroo.

present-acc give-cond cry-start will

"Brown will cry if the very pretty girl gives a present to Blue."

The *Mosi* condition contains conditional adverbs. On the other hand Adverb condition contains degree adverbs that do not have any grammatical relation to conditional clause structures. Embedded clauses are conditional clauses in both conditions. In this comparison, the verb with the conditional morpheme, *-ra* is read more easily in the *Mosi* condition than in the Adverb condition. This result suggests that the parser sets a prediction of the upcoming conditional structure by means of *Mosi*. Because *Mosi* has a direct grammatical relation to conditional clauses, it will be argued that the parser can make use of such direct cue to predict the upcoming conditional clauses. This finding also is compatible with strong incrementality because the prediction of the upcoming conditional clause is set before the actual conditional verb is

encountered. Thus, the parser can project the structure of conditional clauses upon encountering the adverb *Mosi*.

Additionally, I will show that the predicted conditional clauses also block the LDB. As we have discussed, Japanese conditional clauses are not islands. Thus an overt movement like scrambling can escape conditional clauses as in (15).

(15) a. ^{ok}Dono-gakusee-ni$_1$ Brown-wa [$_{cond}$mosi White-ga t$_1$
which-student-dat B-top mosi W-nom
present-o ageta-ra] naki-dasu-no?
present-acc give-cond cry-start-Q
"Which student will Brown cry if White gives a present to?"

b. ^{ok}Dono-gakusee-ni$_1$ [Brown-wa [$_{CP}$ White-ga t$_1$
which-student-dat B-top W-nom
present-o ageta-to] itta-no]?
present-acc gave-comp said-Q
"Which student did Brown say that White gave a present to?"

Given the results of the studies on relative clauses, we expect that the fronted wh-phrase is preferentially associated with the embedded clause in examples like (15a) because of the LDB and non-islandhood of conditional clauses. However, contrary to our expectation, the LDB is blocked when the upcoming conditional clause is predicted by means of *Mosi*.

I will explain this blocking effect of the LDB in the following way. I will show two important observations. First, the results of two experiments will show that

the LDB is motivated by wh-feature on the fronted NP. The LDB is observed only when the fronted material is a wh-phrase. I will argue that the parser tries to associate wh-phrase with Question particle as soon as possible, resulting in the LDB. Second, the conditional verb, *V-ra*, cannot host the question particle, *-ka*. Based on these two observations, I will argue that the LDB is not seen in the predicted conditional clauses because the morphology of the conditional verb, *V-ra*, is predicted by *Mosi*. Because the predicted conditional verbal morphology is not compatible with the motivation for the LDB, it is blocked when the conditional clause is predicted.

The finding that the LDB is blocked when the conditional verb is predicted suggests that the parser projects a detailed structure of the conditional clause that can represent the morphology of conditional verbs. Based on the discussion in chapter 2, in which I will show that conditional verbal morphology is created by V-T-C complex, I will argue that the parser project the full structure of conditional clauses where V, T and C are all represented to support the conditional morphology.

Finally, I will argue that the finding that the predicted conditional morphology blocks the LDB, forces us to reconsider the online island effect induced by relative clauses. As we have discussed, conditional clauses are not islands in Japanese. However, they block the LDB. If so there can be some factor independent of islandhood, which is relevant to the blocking of the LDB. I will show that the morphology of the relative clause verbs is also not compatible with the question particle. On the basis of this observation, I will argue that the LDB is blocked when the relative clause is predicted not because relative clauses are islands but because the morphology of relative clause verbs is predicted.

3. Toward the Nature of Predictive Mechanism

The findings in chapter 3 and chapter 4 support the strong incrementality of sentence processing. Experimental results suggest that both relative clause structures and conditional clause structures are built before the verb or the head of the relative clause is encountered. These findings suggest that the parser is equipped with a powerful predictive mechanism that allows for the strong incrementality. Furthermore, this predictive mechanism should be able to project the detailed structure of upcoming relative clauses or conditional clauses by means of indirect cues such as the classifier mismatch or direct cues such as conditional adverbs like *Mosi*.

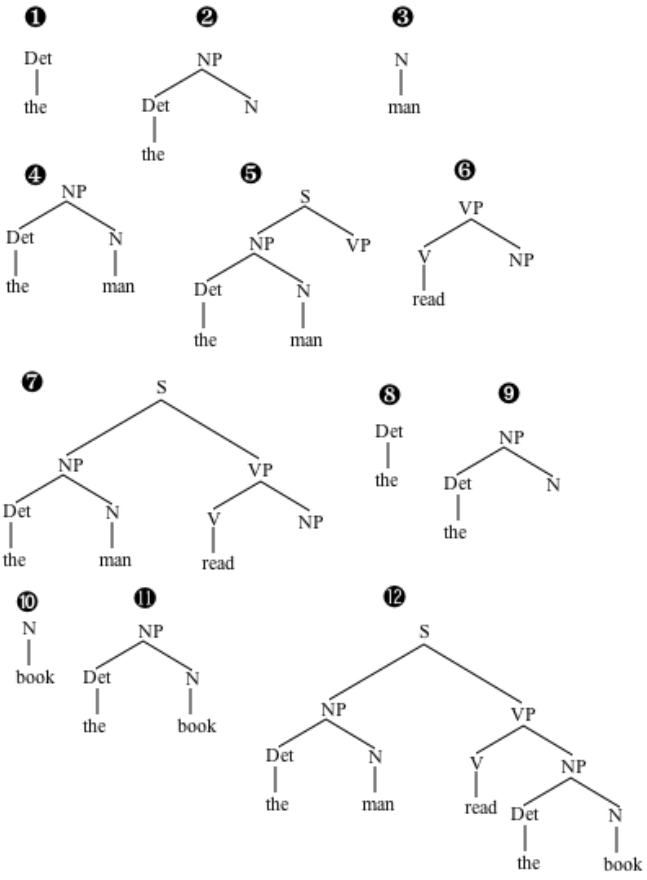
In Chapter 5 I will discuss the issue of the sentence processing mechanism. Specifically I will try to figure out what algorithm can derive the range of predictions that we have seen above. I will critically examine four types of parsing algorithm that have been proposed in the literature, and show that a variant of the so-called left-corner algorithm is a psychologically plausible algorithm for sentence processing.

First, I will argue that both purely bottom-up algorithms and purely top-down algorithms have problems. It will be shown that purely bottom-up parsing is not incremental. On the other hand, purely top-down algorithm has incrementality, but it has problem with left-branching structures. I will show that the so-called Left-Corner parser is psychologically more plausible than these two algorithms (Abney and Johnson 1991; Johnson-Laird 1983; Resnik 1992; Stabler 1994 among others).

Assuming some basic phrase structure grammar, a left-corner parser processes the sentence in the following way. It processes the leftmost daughter of the right-hand

side of a phrase structure rule in a bottom-up fashion, and the rest of the rule in a top-down fashion. Let us see how a left-corner parser processes the sentence “the man read the book.”

(16) The man read the book.



When the parser finds the word *the*, it allows a *Det* node to be built (the step 1: the bottom up structure projection). *Det* is the left-most category at the right-hand side of an NP rule like [NP -- > Det, N], the NP rule is applied and the NP node is built. The NP rule tells that it can be expanded as *Det* and *N* (NP -- > Det N). Thus the sibling of the NP, the N node is also projected (the step 2: the top down prediction). The word

man comes into the input, and it allows the N node to be built (the step 3), which can be attached as the right-hand member of the NP, and NP is completed by attaching the newly projected N into the predicted N node (the step 4). When the NP is completed, an *S* rule like [S -- > NP VP] tells that NP is the left-most category of the right-hand side of *S* rule. Thus, the NP allows the S node and its sibling VP-node to be built (the step 5). The word *read* comes into the input, which can project up to V and VP (the step 6). VP is the right most member of the currently constructed S node. So the newly projected VP is attached to the predicted VP (the step 7). The word *the* comes into the input and as we have seen it can project an NP node and an N node, the sibling of the NP can be predicted (the step 8 and 9). The word *book* comes into the input, projecting the N node. The newly projected N node is incorporated into the predicted N node (step 10 and 11). The newly projected NP is the right-most category of the current VP node, thus the newly projected NP can be incorporated into the predicted NP node (step 12) and the sentence structure is completed.

In this way, a left-corner parser holds incrementality by means of top-down prediction, and it can resolve the problems of bottom-up and top-down algorithms. However, Schneider (1999) points out that even this algorithm has some problems. He argues that because left-corner parsers refer to syntactic categories, they require frequent reanalysis in the processing of a sentence in head-final languages. For example, in German, a case-marked NP may be licensed by either a verb or a postposition. Thus it is risky to commit a particular analysis.

To minimize the amount of reanalysis needed in head-final languages, Schneider proposes a feature-based left-corner parser called SPARSE. In SPARSE,

features are the minimal building blocks. This parser allows for the prediction of features rather than categories. For example, for a German case-marked NP, SPARSE predicts a head with the case feature. Because it does not refer to a particular category, the predicted head is compatible with either a verb or a postposition. In this way, SPARSE has a flexibility in predictive component.

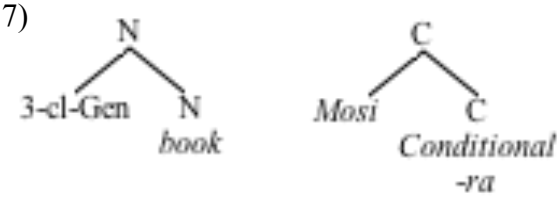
Based on this flexibility and incrementality, I will basically adopt SPARSE as the basic structure-building algorithm.

SPARASE allows flexible structure building. However, like other algorithms, left-corner parsers in general, including SPARSE, do not provide recursive prediction that is required for an example like (6) where a word to be incorporated in the structure, there must be multiple structural predictions. I will show that the recursive prediction problem arises in the prediction of relative clauses. To accommodate the recursive prediction, I will modify SPARSE by enhancing the predictive component. The modification of the parser raises a problem. Enhancing the predictive component to allow for the recursive prediction may make the parser too powerful in a way that the parser allows for an infinite recursive prediction. To resolve this problem, I will also discuss how we can restrict the power of predictive component.

In chapter 5, I will show how the prediction of relative clauses or conditional clauses can be handled by SPARSE. I will point out that the original SPARSE algorithm cannot predict the sufficiently rich structure of relative clauses or conditional clauses based on the cues such as the classifier mismatch or *Mosi*. To resolve this problem, I will modify SPARSE so that we can enhance the predictive component of the parser.

The key points of my modification rely on the grammatical properties of the classifiers and *Mosi*. Both of these items should be able to set a prediction of specific heads. For example, the classifier for books should be able to predict the upcoming N head with the semantics of *book*. On the other hand, *Mosi* should be able to predict the upcoming C which is specified for conditional clause type. The predicted C's specific morphological form *-ra* or its variants, which are understood as complementizers (Chapter 2), should also be predicted by means of *Mosi* because it is the licensor of *Mosi* (Chapter 2).[1]

(17)



Based on this assumption, I will propose that only when the prediction of such specific heads is possible can the parser access the features of the predicted heads, and only then can these features allow for the top-down prediction. For example, if the N head *book* is predicted, the information that *book* does not take a complement clause is also predicted. I will make this type of top-down prediction using the information from the predicted head possible only when a specific head with specific semantic information is predicted. With this restriction, the parser allows a powerful recursive prediction in a certain limited cases. In a normal situation, the parser can

[1] The relation between *mosi* and its licensing morphology is to be understood as a general property of dependent elements and their licensors such as wh-phrase and Q-particle.

only predict the upcoming underspecified head as the original SPARSE does, thus the flexible structure building is not sacrificed even though a powerful multi-step prediction is possible.

In chapter 5 I will show how this modification allows for the parser to predict sufficiently rich structure of relative clauses or conditional clauses in which the LDB can be blocked.

CHAPTER 2. SYNTAX OF RELATIVE CLAUSES AND ADJUNCT CLAUSES

1. Introduction

The aim of this chapter is to figure out the descriptive properties of Japanese Relative Clauses (RCs) and Conditional Clauses. The descriptive syntactic studies help us understand what the representations of these clauses should look like, that the parser construct during online sentence processing. The study on Conditional Clauses is specifically important for us because in Japanese Generative Grammar, the syntax of Conditional Clauses have not been studied much, and therefore representation of conditional clauses are not well understood.

Throughout this chapter, we will try to figure out the internal syntax and external syntax of both RCs and Conditional Clauses by applying various well-known syntactic tests.

In the course of the discussion, it will become clear that Japanese Adjunct Clauses are not islands. The non-islandhood of adjunct clauses is surprising given the fact that adjunct islands are observed in wide varieties of languages (Stepanov 2001). Thus, in this chapter we will try to capture the non-islandhood of Japanese adjunct clauses too. Comparing Japanese with various other languages, we will propose a parameter that regulates the islandhood of adjunct clauses.

The organization of the chapter is as follows. In the next section, we will figure out the descriptive properties of Japanese RCs. In the following section, we will turn to the syntax of Conditional Clauses. Finally, we will turn to the cross-linguistic studies on adjunct clauses.

2. The Syntax of Relative Clauses

The aim of this section is to briefly review the syntactic properties of relative clauses (RCs) in Japanese. Although there are various types of RCs in Japanese, I will concentrate on one type of RCs, the so-called head external restrictive RCs. I will report three prominent properties of RCs, namely: (i) island sensitivity; (ii) reconstruction effects; and (iii) verbal morphology. I will show that they are best captured by a CP-analysis of RCs.

One of the well-known properties of Japanese RCs is their apparent insensitivity to island constraints. It has been reported in various places that Japanese relativization can escape islands (Fukui and Takano 2000; Kuno 1973; Murasugi 1991 among others). Taking some examples, Kuno cites the following examples containing, complex NP islands, adjunct islands, and subject islands. They are summarized in (18) to (20). All of these examples are acceptable even though the relative head is extracted out of island domains. Recently, it has been argued that clausal adjuncts and clausal subjects do not seem to be islands in Japanese (Ishii 1997; Mihara 1994 among others). Therefore, the facts about these two domains are not surprising. However, it is still surprising that Japanese relativization can escape Complex NPs that normally show strong island effects (Saito 1985).

(18) Complex NP Islands

[$_{NP}$[$_{CP}$ ___$_2$ [$_{NP}$[$_{CP}$ ___$_1$ kiteiru] huku$_1$]-ga yogoreteiru]

wearing-is suit-nom dirty-is

sinsi$_2$]

gentleman

"the gentleman who [the suit that he is wearing] is dirty"

(19) Adjunct Islands

[$_{NP}$[$_{CP}$[$_{CP}$ ___$_1$ sinda no de] minna-ga kanasin-da]

died because everyone-nom was-distressed

hito$_1$]

person

Lit. "a person who, because (he) died, everyone was saddened"

(20) Subject Islands

[$_{NP}$[$_{CP}$[$_{CP}$ watakusi-ga ___$_1$ au koto/no]-ga

I-nom meet comp-nom

muzukasii]hito$_1$]

difficult person

Lit. "The person whom that I see/meet (him) is difficult."

The above examples that demonstrate the island insensitivity of relativization, especially the examples with Complex NP islands, have created a controversy regarding the treatment of the syntactic derivation and the structure of RCs in Japanese. The above cases of island insensitivity have been one of the important pieces of evidence for the claim that Japanese RCs are different from those in English

in the important sense that they do not involve movement of the relative head or the relative operator, and that RCs have an IP structure as in (21a) rather than the CP structure such as (21b) (Fukui and Takano 2000; Murasugi 1991).

(21) a. [$_{NP}$[$_{IP}$ Quinn-ga __$_1$ tabeta] ringo$_1$]
Q-nom ate apple

b. [$_{NP}$[$_{CP}$ *Op$_1$* [$_{IP}$ Quinn-ga __$_1$ tabeta]] ringo]
Op Q-nom ate apple
"The apple that Quinn ate."

In the following subsections, I will try to show three arguments against this view, and try to support the movement and CP-analysis of Japanese RCs.

2.1. Island Sensitivity

One of the strongest arguments against the movement analysis, and thus non-CP analysis of RCs comes from the apparent island insensitivity of relativization (Kuno 1973; Murasugi 1991). The claim is that if operator movement is involved in relativization, they should show island sensitivity. Because they do not show island sensitivity, operator movement is not motivated. If operator movement is not motivated, then the position that hosts the operator (CP-spec) is not also motivated. Thus, there is no strong reason to assume CP structures inside RCs. However, there are several arguments that this apparent lack of island effects does not necessarily indicate a lack of movement in Japanese relative clause formation. Sakai (1994), incorporating the so-called major subject analysis of topicalization in Japanese of

Kuroda (1986a), suggests that Japanese RCs can involve movement of a null operator (Op) from the major subject position. According to this analysis, the derivation of RCs is something like (22a), and that of topicalization is (22b)

(22) a. [$_{CP}$ *Op*$_1$ [t$_1$ [$_{NP}$[$_{CP}$ pro$_1$ e$_2$ kiteiru] yoohuku$_2$]-ga
Op pro wearing-is suit-nom
yogoreteiru] [sinsi$_1$]]
dirty-is gentleman
"The gentleman who [the suit that he is wearing] is dirty"

b. [$_{IP}$**(sono) sinsi$_1$-ga** [$_{NP}$[$_{CP}$ pro$_1$ e$_2$
that gentleman-nom pro
kiteiru][yoohuku$_2$]]-ga yogoreteiru]
wearing-is suit-nom dirty-is
"(that) gentleman is such that the suit that he is wearing is dirty"

The upshot of this analysis is the following. In Japanese, major subjects that can be generated outside of RCs are always available (Kuroda 1986a; Watanabe 2003 among others). Thus, if the relative operator is generated as a major subject, and moves from major subject position to the CP-Spec position, the operator movement can circumvent Complex NP Constraint violations because it is a movement that originates outside of the RC (Hoshi 1995, 2004; Sakai 1994). In the same vein, the fact that relativization can escape other types of islands is also explained by major subject analyses.

A specific prediction of this major subject analysis is that if the movement from major subject position is somehow not available, then island effects should emerge. Hoshi (2004) points out that the following example in (23a) is one such case. (23a) is derived from (23b). In (23a), the major subject position is occupied by an NP, *sono sinsi-ga* "that gentleman-nom", and the outermost relative head corresponds to the indirect object of the verb *okutta* "gave" in the RC. In this example, thus, the relative head is extracted out of an RC and the example creates severe unacceptability.

(23) a. *[$_{NP}$[$_{CP}$ sono sinsi$_1$-ga kinoo [$_{NP}$[$_{CP}$ pro$_1$
that gentleman-nom yesterday pro
itinen-mae-ni t_2 t_3 okutta]yubiwa$_2$]-ga
a-year-ago gave ring-nom
nusumareta] okusan$_3$]
was-stolen wife
"The wife that that gentleman is such that yesterday the ring which he gave her a year ago was stolen."

b. sono sinsi$_1$-ga kinoo [$_{NP}$[$_{CP}$ pro$_1$
that gentleman-nom yesterday pro
itinen-mae-ni t_2 okusan-ni okutta] yubiwa]-ga
a-year-ago wife-dat gave ring-nom
nusumareta.
was-stolen

"That gentleman is such that yesterday the ring which he gave his wife a year ago was stolen."

The same point can be made using another type of island constraint, such as the Coordinate Structure Constraint (Ross 1967) in Japanese (Kato 2005a, 2005b among many others). First of all, Japanese VP-coordination constructions (Takano 2004) do not allow a major subject in the configuration illustrated in (24).

(24) [$_{IP}$ Auster kyooju-ga [$_{VP}$[$_{VP}$ Quinn-o sikari]
A-prof.-nom Q-acc scold
[$_{VP}$ Stillman-o home]]-ta]
S-acc praise-past
"Prof. Auster scolded Quinn and praised Stillman."

a. *Sono gakusee$_1$-ga, [$_{IP}$ Auster kyooju-ga
that student-nom A-prof.-nom
[$_{VP}$[$_{VP}$ Quinn-o sikari][$_{VP}$ pro$_1$ home]]-ta].
Q-Acc scold pro praise -past
"That student is such that Prof. Auster scolded Quinn and praised him."

b. *Sono gakusee$_1$-ga, [$_{IP}$ Auster kyooju-ga
that student-nom A-prof.-nom
[$_{VP}$[$_{VP}$ pro$_1$sikari][$_{VP}$ Stillman-o home]]-ta]
pro scold S-acc praise-past
"That student is such that Prof. Auster scolded him and praised Stillman."

Furthermore, Japanese VP-coordination constructions are islands as in (25). The Examples in (25) are cases of scrambling (Kato 2005a). Here, I just cite examples of scrambling, but basically any type of extraction is sensitive to the Coordinate Structure Constraint.

(25) a. *Stillman$_1$-o [$_{IP}$ Auster kyooju-ga [$_{VP}$[$_{VP}$ Quinn-o
S-acc A-prof.-nom Q-acc
sikari][$_{VP}$ t_1 home]]-ta]
scold praise-past
"Prof. Auster scolded Quinn and praised Stillman."

b. *Quinn$_1$-o [$_{IP}$ Auster kyooju-ga [$_{VP}$[$_{VP}$ t_1 sikari]
Q-acc A-prof.-nom scold
[$_{VP}$ Stillman-o home]]-ta]
S-acc praise-past
"Prof. Auster scolded Quinn and praised Stillman."

Whatever the explanation of major subject formation might be[2] (Kuno 1973; Kuroda 1986a, 1986b; Mikami 1960; Saito 1982, 1983; Takezawa 1987 among many others), given the unavailability of major subjects and the islandhood of VP-coordination, the straightforward prediction of the major subject analysis of relativization is that Japanese relativization is constrained by the Coordinate Structure

[2] For our purpose, just showing the island sensitivity of relativization suffices, so I leave the explanation of major subject formation open. For details, see the references cited.

Constraint. This prediction seems to be correct. The following examples show that Japanese relativization is constrained by the Coordinate Structure Constraint.

(26) a. *[$_{NP}$[$_{CP}$ Auster kyooju-ga [$_{VP}$[$_{VP}$ Quinn-o sikari]
A-prof.-nom Q-acc scold
[$_{VP}$ t_1 home]]-ta] gakusee$_1$]
praise-past student
Lit. "The student that Prof. Auster scolded Quinn and praised."

b. *[$_{NP}$[$_{CP}$ Auster kyooju-ga [$_{VP}$[$_{VP}$ t_1 sikari]
A-prof.-nom scold
[$_{VP}$ Stillman-o home]]-ta] gakusee$_1$]
S-acc praise-past student
Lit. "The student that Prof. Auster scolded and praised Stillman."

These examples suggest that when the major subject becomes unavailable, the relative operator has to move from its thematic position, and thus it creates island effects if the original position is embedded in an island. This, in turn, suggests that there are always two possible derivations for Japanese RCs. One involves the movement of relative operator from the major subject position, and the other involves the movement from its thematic position (Hoshi 2004).

Taken together, the data we have seen so far strongly suggest the following two points. First, it seems that there are always two types of possible derivations for Japanese RCs: relative head/operator movement from major subject position; and that from its original position. If one of these derivations becomes unavailable for some

reason (e.g., island violation) the other derivation becomes the only possible derivation. Second, there is indeed movement of some sort, and thus, apparent island insensitivity of Japanese RCs is not a strong argument against movement analyses.

2.2. Connectivity/Reconstruction Effects

In the previous subsection, we have established that Japanese relativization involves some type of movement (it can either be head raising (Bianchi 2000; Hoshi 2004; Kayne 1994) or null operator movement (Browning 1987; Chomsky 1977; Ishii 1991; Sakai 1994; Watanabe 1992)). In this subsection, I would like to point out an additional argument for the movement analysis of Japanese RCs, namely connectivity/reconstruction effects.

Ishii (1991) demonstrates that Japanese RCs show connectivity effects when the relative head contains a local anaphor such as *kare-zisin* "himself" or *kanojo-zisin* "herself". A relevant example is shown in (27). In this example, the local anaphor that is embedded in the relative head is bound by the subject in the RC. As long as these anaphors in Japanese are to be c-commanded by their antecedents (Katada 1991), this example suggests that the relative head is originated in the RC where the subject can c-command the anaphor.[3]

[3] Examples of binding reconstruction effects have been used to support the head-raising analysis of RCs (Kayne 1994). It is problematic to the operator movement analysis because it is normally assumed that empty operator does not show reconstruction effects (Epstein 1989).

(27) [$_{NP}$[$_{CP}$ Quinn$_1$-ga t_2 taipu-sita][kare-zisin$_1$-no
Q-nom type-do himself-gen
ronbun]$_2$]
paper
"the paper of himself that Quinn typed"

Ishii (1991) also observes an apparent counter example to the possibility of connectivity effects within Japanese RCs. He cites the example in (28). In this example, the relative head NP is extracted out of a complex NP, and the coreferential reading of *Stillman* and the local anaphor is not available.

(28) ?*[$_{NP}$[$_{CP}$ Virginia-ga [$_{NP}$[$_{CP}$ Stillman$_1$-ga e_3 e_2 miseta
V-nom S-nom showed
koto-ga aru] hito$_3$]-o sitteiru] [kare-zisin$_1$-no
fact-nom exist person-acc know himself-gen
syasin]$_2$]
picture
Lit. "The picture of himself which Virgnina knows the person to whom Stillman has once showed."

Given the fact that the relative head can escape complex NP islands, this example is not expected to be bad. However, Hoshi (2004) points out that this example is not truly problematic. Remember that there are two derivations available for Japanese RCs. One of them involves movement from the major subject position and the

starting point of the movement of relative head or relative operator is outside of the complex NP island. According to Hoshi, a major subject construction with a local anaphor in the same configuration is indeed unacceptable, as the following example shows. Thus, he concludes that the apparent problematic example in (28) is derived from an underlying structure that involves a major subject.

(29) *[kare-zisin$_1$-no syasin]-ga [$_{IP}$ Virginia-ga
himself-gen picture-nom H-nom
[$_{NP}$[$_{CP}$ Stillman$_1$-ga t$_3$t$_2$ miseta koto-ga aru]
S-nom show fact-nom exist
hito]-o sitteiru.]
person-acc know
Lit. "The picture of himself is such that Virginia knows the person to which Stillman showed it."

What I will argue in the section regarding adjunct clauses in Japanese is that they are not islands for movement. Together with Hoshi's theory of RCs, the non-islandhood of Japanese adjunct clauses makes a specific prediction, namely that relativization out of adjunct clauses should show binding connectivity effects. This is because, if the relative head or relative operator moves from a non-island domain, movement from the non-major subject position becomes possible. Looking at an example like (30), the prediction seems to be correct. In this example, the coreferential interpretation between *Stillman* and *kare-zisin* "himself" in the relative head position is possible without any trouble.

(30) [$_{NP}$[$_{CP}$ minna-ga [$_{Cond}$ Stillman$_1$-ga t_2 sutete-simatta-ra]
everyone-nom S-nom trash-end-up-with-cond
komatte-simau] [kare-jisin$_1$-no ronbun]$_2$]
have-trouble-end-up-with himself-gen paper
Lit. "The paper of himself that will make a trouble to everyone if Stillman trash it "

2.3. Summary

The discussion so far established two important claims about Japanese relativization. First, it involves some type of movement. Specifically, there are two starting points for the relative head or relative operator: one is from the major subject position; and the other is from the thematic position inside the relative clause. Second, if the derivation involving a major subject is not available, the relative operator has to move out of its original, thematic position inside the relative clause. Thus, in this case, the relativization exhibits island effects if the underlying position of the relative head is inside an island. These two observations basically reject the claim that Japanese relativization does not involve movement of any kind. Thus this, in turn, supports the claim that Japanese Relative Clauses and those found in English are basically the same type of construction. Only the difference, besides their basic word orders, between those two is the availability of a major subject.

2.4. Issues of Verbal Morphology within Relative Clauses

The third property of Japanese RCs that I would like to discuss concerns the morphology of embedded verbs. Following Hiraiwa's series of studies on the so-called Nominative-Genitive Conversion, I will show that embedded verbs in Japanese RCs have special verbal morphology, the so-called predicate adnominal form. I will further show that detailed examination of this predicate adnominal form will provides us with an important clue for the internal phrase structure of RCs in Japanese, namely the existence of CP-layers.

2.4.1. Nominative Genitive Conversion

It is well-known that Japanese nominative Case is optionally converted to Genitive Case in specific environments, the phenomenon called Nominative-Genitive Conversion (Fukui 1995; Harada 1971, 1976; Hiraiwa 2000, 2001; Miyagawa 1989, 1993; Ochi 2001; Saito 1982; Sakai 1994; Shibatani 1978; Watanabe 1994, 1996). Nominative-Genitive Conversion is typically observed in RCs. In (31), for example, the nominative case on the subject in an RC is converted to genitive case.

(31) a. Kinoo Quinn-**ga** katta hon
yesterday Q-**nom** bought book
"The book that Quinn bought yesterday"

b. Kinoo Quinn-**no** katta hon
yesterday Q-**gen** bought book

Traditionally, it has been argued that genitive case on the subject is licensed in a clause that is headed by a nominal element such as RCs, complement clauses of nouns, or clauses headed by nominalizing complementizers. Backed up by the fact that the genitive case is typically licensed by nominals, researchers have argued that the genitive case in Nominative-Genitive Conversion is also licensed by these nominal heads (Fukui 1995; Harada 1971, 1976; Miyagawa 1989, 1993; Ochi 2001; Saito 1982; Sakai 1994; Shibatani 1978). However, recently, problems with this view have been pointed out. Watanabe (1994, 1996) and Hiraiwa (2000, 2001) show that genitive subjects can be licensed in environments with no nominal heads. They cite examples like (32). (32) was first pointed out by Watanabe. (32) contains examples of comparative constructions. Genitive subjects are licensed in comparative constructions without problems, even though they do not utilize any nominal heads in order to mark embedded clauses. This point can be shown by the fact that the marker of the embedded clause *yori* "than" does not take the genitive form of the pronoun *sono* but they select the full DP form *sore* as summarized in (34). This suggests that *yori* is not a nominal element. Hiraiwa cites the example in (33) containing an *until*-clause. Until-clauses also do not contain nominal heads, but license genitive subjects. Hiraiwa cites six more types of examples and confirms that genitive subjects can indeed be licensed in non-nominal environments.

(32) a. Quinn-wa [Virginia-**ga** yonda yori] takusan-no
Q-top V-**nom** read than many-gen
hon-o yonda.
book-acc read
"Quinn read more books than Virginia did."

b. Quinn-wa [Virginia-**no** yonda yori] takusan-no
Q-top V-**gen** read than many-gen
hon-o yonda.
book-acc read

(33) a. Quinn-wa [ame-**ga** yamu made] office-ni ita.
Q-top rain-**nom** stop until office-at be-past
"Quinn was at his office until the rain stopped."

b. Quinn-wa [ame-**no** yamu made] office-ni ita.
Q-top rain-gen stop until office-at be-past

(34) a. *sono yori/ *sono made
it-gen than it-gen until

b. sore yori/ sore made
it-gen than it-gen until

Hiraiwa concludes that the crucial factor in the licensing of genitive subjects is not a nominal head, rather it is the special verbal morphology that licenses genitive subjects. According to Hiraiwa all of the environments in which genitive subjects are legitimate, including RCs, involve the same morphology on the embedded verb, i.e.,

the so-called predicate adnominal form. Thus, he concludes that genitive subjects are licensed by a predicate that employs the predicate adnominal form.

A note is in order here. In the verbal morphological paradigm in modern Japanese, it is difficult to distinguish the predicate adnominal form and the matrix sentence final form, the so-called end form, because both of them have the same form, i.e., verbs without overt complementizers. However, the so-called verbal adjectives and copula in Japanese preserve an explicit distinction between the adnominal form and the end form. Thus, we can test whether the environments that license genitive subjects contain the predicate adnominal form or not by using verbal adjectives or copulas. The end form of verbal adjectives is *-da* and the adnominal form is *-na* respectively. Thus, we expect that in genitive subject environments, these predicates should bear the *-na* form rather than the *-da* form. As we can see in (35), in all the environments where the genitive subject is licensed, these elements inflect with the adnominal form.

(35) a. Relative Clauses

Quinn-ga suki-**na/*-da** ongaku

Q-nom like-AdNom/*End music

"The music that Quinn likes"

b. Comparatives

Quinn-no koto-ga simpai-**na/*-da** yori mo

Q-gen thing-nom worried-AdNom/*End than

Virginia-ga simpai-da.

V-nom worried-be-End

"I am worried about Quinn than about Virginia"

c. *Made* Clause

Quinn-wa ijou-**na**/-***da** mad-ni
Q-top extraordinary-AdNom/*-End extent-to
sinkeisitu da.
nervous be-End
"Quinn was extraordinarily nervous"

Because the adnominal form is the legitimate form in these environments, we can conclude that the verbs in these environments also bear the adnominal form, and thus, that the genitive subject is licensed by the adnominal morphology of the embedded verb.

2.4.2. Predicate Adnominal Form and the C-system

One of the well-known constraints on Nominative-Genitive Conversion is the so-called complementizer blocking effect. Simply put, if the embedded clause is headed by an overt complementizer, the genitive subject is not licensed. Let us see some examples from Hiraiwa (2000).

(36) a. [$_{NP}$[$_{RC}$ syoorai daijisin-ga okiru]
in-the-future great-earthquate-nom occur
kanousei]
possibility
"The possibility that a great earthquake will occur in the future."

b. [$_{NP}$[$_{RC}$ syoorai daijisin-no okiru]
in-the-future great-earthquate-gen occur
kanousei]
possibility
"The possibility that a great earthquake will occur in the future."

c. [$_{NP}$[$_{CP}$ syoorai daijisin-ga okiru
in-the-future great-earthquate-nom occur
toiu]kanousei]
comp possibility
"The possibility that a great earthquake will occur in the future."

d. *[$_{NP}$[$_{CP}$ syoorai daijisin-no okiru
in-the-future great-earthquate-gen occur
toiu] kanousei]
comp possibility
"The possibility that a great earthquake will occur in the future."

The important point of these examples is that, although all of the four sentences have basically the same meaning, only the example hosting the overt complementizer disallows the genitive subject. This paradigm suggests that adnominal predicate formation is blocked by the presence of an overt complementizer, and thus adnominal formation has some relation to the complementizer in the embedded clause. As a support for this position, we can cite examples from verbal adjectives again. In the example (37), if the complementizer is present, the verbal adjectives cannot bear the adnominal form, i.e., the *-na* ending.

(37) a. [Quinn-ga gengogaku-ga kirai-na] kanousei
Q-nom linguistics-nom hate-AdNom possibility
"The possibility that Quinn does not like linguistics"

b. *[Quinn-ga gengogaku-ga kirai-na **toiu**]
Q-nom linguistics-nom hate-AdNom comp
kanousei
possibility
"The possibility that Quinn does not like linguistics"

c. [Quinn-ga gengogaku-ga kirai-da **toiu**]
T-nom linguistics-nom hate-end comp
kanousei
possibility
"The reason that Quinn does not like linguistics"

The discussion so far suggests that there is a tight connection between the complementizer and the licensing of the adnominal forms of the embedded predicates. The simplest way to capture the above paradigm regarding the adnominal predicate formation is to assume that there is an empty complementizer position that stands in an agreement-like relation with the embedded verb. Otherwise, it is difficult to explain why the presence of the overt complementizer blocks the adnominal predicate formation. Hiraiwa indeed provides this line of argument. Hiraiwa's theory of verbal morphology assumes that an adnominal predicate is formed by the amalgamation of v, T and C though the operation Agree (Chomsky 2001 among others), and his theory can give a straightforward explanation for the above paradigm.

For our purposes, however, it is more important to show that predicate adnominal form is licensed by a specific type of C, because if it is true, it suggests that Japanese RCs contain a C-layer.

2.4.3. Relative Clauses and the C-system

Remembering that RCs are one of the typical environments where nominative-genitive conversion is allowed, it shall now be clear that RCs be the environment where the predicate adnominal form is also allowed. The conclusion we reached in the previous subsection is that the adnominal form is licensed by an empty complementizer. Thus, we can draw the same conclusion for the phrase structure of RCs, i.e., the embedded clause in an RC is headed by the empty complementizer.

2.5. Conclusion

In this section, I have reviewed three arguments for the movement and CP analysis of Japanese RCs. Based on the observations on island sensitivity and connectivity effects, we have concluded that Japanese relativization involves movement of a relative head or relative operator. We have also seen that embedded verbs in Japanese RCs have a predicate adnominal form. Based on the facts about nominative-genitive conversion and complementizer blocking effects on adnominal predicate formation, we concluded that embedded clauses in Japanese RCs are headed by an empty complementizer. Taken together, Japanese RCs, like English RCs, involve movement of a relative operator and a CP structure. The structure of Japanese RCs is thus something like the following.

(38) $[_{NP}[_{CP}\ Op_1 \ldots t_1 \ldots\ V\text{-}_{AdNom}]\ NP_1]$

3. The Syntax of Conditional Clauses in Japanese

3.1. Non-islandhood of Japanese Conditionals

We will start our discussion from the curious fact that Japanese conditional clauses do not show strong island effects. The examples in (39) are the relevant cases. Unlike English overt movement, scrambling out of conditional clauses is not degraded. This point can be made clear by comparing examples like (39b) and examples of long-scrambling out of relative clauses or complement clauses.

(39) a. Virginia-wa [mosi Quinn-ga sono cake-o tabe-
V-top mosi Q-nom that cake-acc eat-
ta-ra] naki-dasu daroo.
past-cond cry-start will
"Virginia will start crying if Quinn eats that cake."

b. *Sono cake-o*$_1$ Virginia-wa [mosi Quinn-ga t_1
that cake-acc H-top mosi T-nom
tabe-ta-ra] naki-dasu daroo.
eat-past-cond cry-start will
"Virginia will start crying if Quinn eats that cake."

As it will be discussed in later sections, Japanese conditional clauses show the signature properties of so-called Weak Islands. However, as it will become clearer,

the weak island effects are induced by the Conditional Adverb (CA) *mosi*. This suggests that conditional clauses per se do not induce island effects.

The aim of this subsection is to evaluate one of the possible approaches to the problem of non-islandhood of conditional clauses. The claim that we will evaluate is that conditional clauses in Japanese are more like arguments than adjuncts. The intuition behind this claim is the following. Overt extraction out of conditional clauses is allowed because conditional clauses are complement clauses, which do not induce island effects under normal circumstances (Chomsky 1981, 1986b; Huang 1982; Ross 1967).

The type of analysis that certain adjunct clauses are more like complement clauses has been suggested in previous literature. To the best of my knowledge, it was Mihara (1994) who first acknowledged that adjunct clauses in Japanese allow overt extraction, and he actually hints at the possibility that adjunct clauses that allow extraction are more like complement clauses. His examples come from scrambling. He shows that long scrambling out of *because*-clauses in Japanese is allowed.

(40) a. Quinn-wa [Stillman-ga bungakubu-ni
Q-nom S-nom literature-department-to
nyuugaku-sita-node] odoroi-ta
enter-did-because get-surprised-past
"Quinn got surprised because Stillman entered the department of literature"

b. bungakubu-ni Quinn-wa [Stillman-ga
literature-department-to Q-nom S-nom
nyuugaku-sita-node] odoroi-ta
enter-did-because get-surprised-past
"Quinn got surprised because Stillman entered the department of literature (rather than some other department)"

Mihara suggests that the notion of "degree of subordination" in traditional Japanese grammar (Kuno 1973; Masuoka and Takubo 1992 among others) might be relevant to the adjunct clauses' generosity with regard to extraction. His conclusion is basically that adjunct clauses in Japanese are not islands. However, he suggests a correlation between extractability and the degree of subordination, i.e., if the degree of subordination is lower, the extraction would be worse. Because in traditional Japanese grammar, however, the term "degree of subordination" has not been well defined, we cannot argue anything definitely. Informally, we can understand degree of subordination as corresponding to the extent to which a clause behaves like a matrix/independent clause or a complement clause. Thus, under this understanding of the term, we can interpret what Mihara suggests as that adjunct clauses that allow extraction are more like complement clauses.

In this section, however, I will show that conditional clauses should be analyzed as adjunct clauses. Thus, contrary to the widely accepted view that clausal adjuncts do not allow overt extraction cross-linguistically (Huang 1982; Saito 1985; Stepanov 2001), I will argue that Japanese conditionals allow the extraction even though they are adjuncts.

3.1.1. Adjuncthood of Conditional Clauses: English

Basically the argument that conditional clauses in English are adjuncts is backed up by four observations. First, there seem to be no verbs that select conditional clauses as their complements. This point is related to the optionality of conditional clauses. Second, unlike arguments, the positioning of conditionals in a sentence is relatively free. Third, movement of conditional clauses is severely constrained by wh-islands. Fourth, conditionals behave like other adjuncts with respect to constituency tests. I will review these points one by one.

The first indication of the adjuncthood of conditional clauses comes from the fact that they are optional elements and there seems to be no predicate that takes conditionals as its complement. The fact that in the following examples the presence or absence of conditional clauses does not have an impact on acceptability supports this point. These examples also show the adjuncthood of conditional clauses because verbs like *leave* that do not take complement clause can cooccur with conditional clauses.

(41) a. Quinn will leave [if you get angry].

b. Quinn will leave.

Another major argument for the adjuncthood of conditional clauses is based on the fact that their positioning in a sentence is relatively free. A conditional clause can appear both sentence-initially and sentence-finally without changing the basic interpretation of the sentence. The circumstance is different if an argument NP is

moved from its original position. For example, if a direct object NP is fronted to the beginning of a sentence, it obligatorily receives a topicalized interpretation. The sentence initial conditionals, however, does not necessarily receive a topicalized interpretation. Thus, this freedom of the positioning in a sentence distinguishes conditional clauses from arguments.

(42) a. Quinn will leave [if you get angry].

b. [If you get angry] Quinn will leave.

(43) a. Quinn will eat that cake.

b. That cake, Quinn will eat.

One of the most powerful arguments for the adjuncthood of conditionals in English comes from extraction from WIs such as wh-islands. It is well known that argument extraction from wh-islands is more or less allowed but that adjunct extraction creates severe degradation. Collins (1998) shows that English conditionals cannot be extracted out of wh-islands, taking advantage of the fact that they can be clefted. In the examples in (44) an if-conditional is clefted. (44a) shows that the clefting of if-clauses is unbounded, i.e., the if-clause can be extracted out of a complement clause. On the other hand, in (44b) the if-clause is extracted out of a wh-island, and the example is not acceptable. According to Collins, the severity of the violation in (44b) can be analyzed as an ECP violation. Given that conditional clauses are obviously not subjects, this extraction pattern strongly suggests that they are adjunct clauses.

(44) It is if the student fails that the teacher will fire the TA.

a. ?It is if the student fails that the Stillman said that the teacher would fire the TA.

b. *It is if the student fails that Stillman wonders why the teacher will fire the TA.

Finally, let us take a look at an argument based on classical constituency tests. Constituency tests such as VP-deletion (Jayaseelan 1990; Sag 1976) and *do so* substitution (Lakoff and Ross 1976) show clearly that conditional clauses in sentence final position behave like adjuncts (Bhatt and Pancheva 2001). Bhatt and Pancheva cite the following examples. In both of the examples in (45), the conditional clauses are interpreted in the ellipsis site or *do so* anaphora. These examples suggest that they are attached to the VP, and thus can be inside the scope of deletion or *do so* substitution. Together with these examples, the ones in (46) suggest that conditionals are adjuncts rather than arguments, where if-clauses are stranded outside of VP-deletion or *do so* anaphora.

(45) a. I will leave if you do and Quinn will $_{\text{leave}}$ <u>if you do</u>, too.

b. I will leave if you do and Quinn will *do so* too.

(46) a. I will leave if you do and Quinn will $_{\text{leave}}$ if Virginia does.

b. I will leave if you do and Quinn will *do so* if Virginia does.

So far, I have reviewed a set of argument for the adjuncthood of conditionals in English. One might argue that this issue is too obvious and trivial to spend a lot of space. However, laying out these arguments is crucially necessary especially for the analysis of conditionals in Japanese where a systematic argument for their constituency has not been offered in the literature. In the following section, I will apply some of the constituency tests reviewed above in order to show that conditional clauses in Japanese show adjuncthood in the same way as English conditionals.

3.1.2. Japanese Conditionals as Adjunct Clauses

Even in Japanese, there seems to be no verb that selects conditional clauses as its complement. Thus, we can basically conclude that they are adjuncts rather than arguments. However, syntactically showing their adjuncthood is not so easy in Japanese. When we try to test the constituency of sentences in Japanese, we always face the problem of whether Japanese has the same type of constituency tests that we can see in English. For example, we realize that the first two tests in the previous section, optionality on the one hand, and freedom of word order on the other, are not so informative, given that Japanese allows empty elements that can refer to both arguments and adjuncts (Kuno 1973; Kuroda 1965) and given that Japanese allows free word order. Furthermore, movement of non-wh-phrases is not constrained by wh-islands in Japanese (Boskovic and Takahashi 1998; Saito 1985; Takahashi 1993; Watanabe 1992 among others). Therefore, making use of wh-islands is also not so straightforward. Facing these problems, I can pick only one test from the four that we have seen above for the adjuncthood of conditional clauses, namely VP-constituency tests.

Like English, Japanese has a variety of constituency test that can tap on the structure of VP. There are two of them that have been frequently mentioned in the literature. One is VP-fronting (Hoji et al. 1989; Tateishi 1994), and the other is *soo su* "do so" substitution (Hinds 1973a, 1973b; Inoue 1976; Nakau 1973; Tateishi 1994). The utility of these two syntactic operations is that both of them are sensitive to the differences between arguments and adjuncts in the same way as English constituency

tests. For this reason, by making use of these tests, we can examine the adjuncthood of conditional clauses in Japanese.

Let us start our discussion from VP-fronting. Japanese VP-fronting has some prominent grammatical properties. It is allowed if the topic marker -*wa*, the emphatic marker -*mo*, or any of the contrastive markers are attached to the fronted VP, and if an auxiliary verb is inserted to support the stranded tense morpheme (Hoji et al. 1989). Even though Japanese makes use of some explicit morphological markers, these properties also hold true in English. The first property is related to the fact that VP-fronting is a phenomena associated with Topic or Focus (Rochemont 1986) and the second property can be understood as the Japanese counterpart of the constraint on stranded affixes (Lasnik 1981). For example, we can observe the following contrast (Saito 1985; Tateishi 1994).

(47) a. Quinn-ga sono hon-o kat-ta koto.
Q-nom that book-acc buy-past fact
"The fact that Quinn bought that book."

b. *[$_{VP}$ Sono hon-o kau]$_1$ Quinn-ga t$_1$ ta koto.

c. [$_{VP}$ Sono hon-o kai]$_1$-***wa*** Quinn-ga t$_1$ ***si***-ta koto.

(47b) is excluded because there is no contrastive marker on the fronted VP and the tense morpheme is stranded alone. If these two conditions are satisfied, the example becomes acceptable as in (47c). These properties, therefore, allow us to conclude that Japanese VP-fronting is basically the same type of operation as in English.

VP-fronting in Japanese has various other properties, but for our purposes, the following is the most relevant: if a VP is fronted, the complement must be fronted. Compared to (47c), the example in (48) is not at all acceptable.

(48) *[$_{VP}$ kai]$_1$-***wa** Quinn-ga sono hon-o t$_1$ ***si***-ta koto.

Given this property, it follows that if conditional clauses are arguments, they must be fronted together with the VP. However, as the pair of examples in (49) suggests that it is not the case. The stranding of conditionals does not make the sentence unacceptable.

(49) Quinn-wa [(mosi) Stillman-ga kita-ra] nigedasu
Q-top mosi S-nom come-cond run-away
daroo.
will
"Quinn will run away if Stillman comes."

a. [$_{VP}$[(mosi) Stillman-ga kita-ra] nigedasi]$_1$-sae
mosi S-nom come-cond run-away -even
Quinn-wa t$_1$ suru daroo
Q-nom do will
"Quinn would even run away if Stillman comes."

b. ?[$_{VP}$ nigedasi]$_1$-sae Quinn-wa [(mosi) Stillman-ga
run-away-even T-top mosi H-nom
kita-ra] t_1 suru daroo.
come-cond do will

For some speakers, (49b) does not sound good. However, this seems to be because of the length and complexity of the conditional clause itself. As an indication of this, if we replace the arguments inside conditionals to empty arguments and omit *mosi*, the sentence becomes much more acceptable even to these speakers. In the example in (50), a context is inserted before the sentence to support the empty arguments, and this sentence is acceptable to those speakers.

(50) Quinn and Virginia are discussing what will happen to Stillman if he eats a really spicy hot pepper.

Quinn: Naki-sura Stillman$_1$-wa [pro$_1$ pro$_{(hot\ pepper)}$
cry-even J-top
tabeta-ra] suru daroo.
eat-cond do will
"Stillman will even cry if he eats the hot pepper."

Another constituency test, the *soo su* substitution test (Shibatani 1973; Terada 1990 among others), provides a support for the argument above. Just like *do so* in English, Japanese *soo su* has been used for testing VP-constituency. Although the status of *soo su*, whether it is a deep anaphor or a surface anaphor (Hoji 1987), is

controversial, the important generalization that has been accepted in the previous studies is that *soo su* replaces a VP. In other words, when *soo su* is used, the complement must be replaced together with the verb. The examples in (51) illustrate this point clearly.

(51) Quinn-wa ringo-o tabe-ta.
Q-top apple-acc eat-past
"Quinn ate an apple."

a. Stillman-mo soo-si-ta.
S-also so-do-past.
"Stillman did so too."

b. *Stillman-mo ***remon-o*** soo-si-ta.
S-also lemon-acc so-did-past
"Stillman did so a lemon."

Following the same logic as above, we expect that if conditionals are arguments, they must be replaced by *soo su* together with the other elements in the VP. However, they can be stranded out of the VP.

(52) Quinn-wa [mosi Virginia-ga kaetta-ra] kaeru
Q-top mosi V-nom go-home-cond go-home
deshoo.
will
"Quinn will go home if Virginia goes home."

a. Stillman-mo soo-suru deshoo.
S-also so-do will
"Stillman will do so too."

b. Stillman-mo [mosi Auster-ga kaetta-ra]
J-also mosi A-nom go-home-cond
soo-suru deshoo.
so-do will
"Stillman will do so if Auster goes home too."

In exactly the same way as the English examples, they strongly indicate that conditional clauses are adjuncts rather than arguments.

The above observations tell us some important properties of conditionals in Japanese. One is that Japanese conditional clauses are indeed adjuncts rather than arguments. Otherwise, we cannot capture the different behaviors of conditional clauses and arguments with respect to VP-constituency tests. The other is that sentence medial conditional clauses, at least, are generated in a VP-adjoined position. The fact that the conditional can move together with VP, and replaced by *soo su* together with VP suggests that they are a part of VP. Thus, we can conclude that conditional clauses in Japanese are base generated in a VP-adjoined position.

3.1.3. Summary

In this section, we have reviewed one potential approach to the non-islandhood of conditional clauses, namely that conditionals are complement clauses. Through VP-constituency tests in which conditionals show systematic differences

from complement clauses, we have reached the conclusion that they should be analyzed as adjunct clauses rather than complement clauses.

3.2. Conditional Clauses as Weak Islands

The aim of this section is to show that Conditional Clauses in Japanese exhibit signature properties of so-called Weak Islands (henceforth WIs) (Cinque 1990 among many others). Specifically, my claim is the following. The Conditional Adverb (CA) *mosi* is the inducer of the weak-islandhood of conditional clauses and conditional clauses per se are not islands. This means that if *mosi* is present in a conditional clause, it becomes a WI. On the other hand, if *mosi* is not present, the conditional clause does not show WI effects. To establish this claim, I will apply various tests of WIs to conditional clauses.

3.2.1. Diagnosing Weak Islands

Referentiality

Although there are many properties of WIs that have been reported in the literature, there is one commonly acknowledged feature, namely that extraction out of WIs is dependent on the referentiality of wh-phrases (Cinque 1990; Rizzi 1990 among many others).

The basic pattern in the data is that if a wh-phrase is referential, it is easier to move it out of WIs. On the other hand, if a wh-phrase is not referential, extraction out of WIs becomes worse. Let us see this pattern by looking at some examples from English.

Rizzi, following Cinque (1984), proposes that only elements assigned referential theta roles can be extracted from a WI, and everything else cannot. A referential NP is understood as an NP that refers to specific members of a preestablished set. Based on this notion of referentiality, Rizzi draws a distinction between arguments with referential theta-roles such as agent, theme, patient, experiencer and so on, and the other elements. The necessity of this distinction can be clearly seen in the following pair of examples. In both of the examples the verb *weigh* takes a complement NP, *apples* and *200 lbs* respectively. According to Rizzi, because they are complement NPs, they should not be structurally different, i.e., they are both sisters of the verb *weigh*, and thus presumably assigned theta roles.

(53) a. Quinn weighed *apples*.

a. Quinn weighed *200 lbs*.

However, as far as movement is concerned, these complements show a clear difference. The question in (54a) is ambiguous but if a wh-phrase is extracted from a wh-island as in (54b), only the agentive reading becomes available. Therefore, the question in (54b) can be properly answered *apples*, but not *200 lbs*.

(54) a. What did Quinn weigh t?

b. ?What did Quinn wonder how to weigh t?

Rizzi's point is that this difference cannot be captured by a theory that simply incorporates a complement/non-complement asymmetry (Chomsky 1986b; Huang 1982) because they are both complements. On the other hand, given the referential/non-referential distinction, this difference can be easily captured. NPs such as *apples* receive referential theta role according to Rizzi. This is so because it is an element that refers to participants in the event described by the verb. *200 lbs*, on the other hand, is an expression that qualifies something, not something that participates in the event.

I will basically follow Rizzi's proposal here, and test the weak-island status of Japanese Conditional clauses referring to the notion of referentiality.

Constructions Sensitive to Weak Islands

We have briefly seen that referentiality is a crucial property of deciding whether an element can be extracted out of WIs or not. In this subsection, let us see more concretely which constructions exhibit WI effects and which do not.

First, as we have already seen, arguments and adjuncts show differences. Normally arguments can be extracted from WIs but adjuncts cannot (Chomsky 1986; Huang 1982; Lasnik and Saito 1992).

(55) a. ?Which man are you wondering whether to invite t?

b. *How are you wondering whether to behave t?

(Szabolsci & den Dikken 1999)

It has been observed however that among the types of adjuncts, *why* and *how* show strong WI sensitivity, *when* has an intermediate status, but adjuncts such as *where* do not (Chomsky 1986b; Huang 1982; Lasnik and Saito 1992; Szabolcsi and den Dikken 1999 among others). Szabolsci and den Dikken cite the following examples.

(56) a. *Why did Quinn ask whether to do this t?

b. *How did Quinn ask whether to do this t?

c. ??When did Quinn ask whether to do this t?

d. Where did Quinn ask whether to read this book t?

These differences among adjuncts suggest that WI-sensitivity does not depend just on whether the extracted elements are subcategorized for or not. Rather, the patterns illustrated in (55) and (56) are compatible with Rizzi's dichotomy of extracted elements based on referentiality. According to Rizzi, although *manner* or *reason* phrases may be arguments, they do not receive referential theta roles. Thus, in this sense they are similar to the amount phrases that we have seen in (53). An amount phrase may also be an argument but it does not receive a referential theta role.

Whether the extracted elements are D-linked or not is another case that affects extractability. If the wh-phrase is D-linked, it can be extracted from WIs in contrast to Non D-linked wh-phrases (Cinque 1990; Lasnik and Saito 1992; Pesetsky 1987 among others). Szabolcsi and den Dikken cite the following minimal pair in (57). In (57b), even though the wh-phrase is an amount phrase, it is D-linked in the sense that it involves a contextually established set of books. On the other hand, (57a) is not D-

linked in this sense. As the minimal pair suggests, if a non D-linked wh-phrase is extracted from a weak island, the sentence becomes unacceptable.

(57) a. *How many books are you wondering whether to write t next year?

b. How many books on the list are they wondering whether to publish t next year?

Another support for the claim that D-linking is a significant factor for extraction comes from examples of so-called aggressively non D-linked wh-phrases. Pesetsky points out that *wh-the-hell* phrases are not compatible with overt markers of D-linking such as *which*, and they indeed cannot be D-linked, i.e., it cannot refer to the member of a preestablished set of entities.

(58) a. What the hell book did you read that in?

b. *Which the hell book did you read that in?

Because they cannot be D-linked, the straightforward expectation is that these wh-phrases cannot be extracted from WIs, and this indeed seems to be correct.

(59) a. ??Who the hell are you wondering whether to invite t?

b. Which man are you wondering whether to invite t?

According to Cinque (1990), D-linking and referentiality are the same notion. Both of them crucially refer to the notion of preestablished set, a certain presupposition. Thus, examples of the-hell question also show the referentiality asymmetry.

Finally, let us see some examples from scope reconstruction cases. It is well known that a wh-phrase like *how many books* has wide or narrow scope with respect to a verb like *want* (Cinque 1990; Kroch 1989; Rullmann 1995). This is illustrated by the following example cited from Rullmann (1995). The wide scope reading of (60) is paraphrased in (60a), where it is assumed that there are certain number of books which Chris wants to buy and the speaker asks how many such books there are. On the other hand, the narrow scope reading is paraphrased as (60b). Under this reading, it is not assumed that there is any specific set of books that Chris wants to buy, but it is assumed that Chris wants to buy a certain number of books. Whether the wh-phrase refers to the preestablished set of entities or not is the crucial notion to distinguish these two reading.

(60) How many books does Chris want to buy?

a. What is the number *n* such that there are *n* books that Chris wants to buy.

b. What is the number *n* such that Chris wants it to be the case that there are *n* books that he buys?

In the literature, it is observed that this *how-many* question is sensitive to WIs (Cinque 1990; Kroch 1989; Rullmann 1995). If the *how-many* phrase moves across a

WI, the narrow scope reading becomes unavailable, i.e., the non-referential reading becomes unavailable. To illustrate this point, Rullmann (1995) cites the following example.

(61) How many books did no student want to buy?

a. What is the number *n* such that there are *n* books that no student wants to buy?

b. *What is the number *n* such that no student wants it to be the case that there are *n* books that s/he buys?

The negation *no* is known to create WIs (Beck 1996; Ross 1984; Rullmann 1995 among others), and if *how-many* is extracted from under negation, the narrow scope reading is no longer available anymore. Only the presuppositional, wide scope reading is available. This is understood as that the WI induced by the negation blocks the reconstruction of the *how-many* phrase, and as a result narrow scope non-referential interpretation becomes unavailable.

Let us summarize the discussion so far. We have seen that referentiality is the crucial notion for distinguishing the elements that can be extracted from WIs and those that cannot. Following this line of argument, we have seen at least four cases of extraction that are sensitive to WIs.

Cinque argues that the major difference between WIs and Strong Islands is the following. WIs allow extraction of referential elements but disallow the extraction of non-referential elements. On the other hand Strong Islands do not allow any

extraction. Given this property, we can test whether a domain is a WI or not by making use of the referentiality asymmetry. In other words, if movement of a referential element is allowed but non referential element is not allowed from a certain domain, we can call this domain a weak island. Following this way of thinking, we can test whether Japanese Conditionals are WI or not.

3.2.2. *Mosi* as a Weak Island Inducer

In the discussion so far, I have shown cases of extraction that are sensitive to WIs. From now, I will take advantage of the WI sensitivity of these elements and show that Japanese Conditionals are WIs. In the course of the discussion, I will further show that the WI effects shown by Japanese Conditionals are induced by the CA, *mosi*.

First let us see the simple argument-adjunct asymmetry. In (62), two sentences are compared where (62a) contains the in-situ argument wh-phrase *dono-ringo* "which apple" inside the conditional clause, and (62b) contains the in-situ adjunct wh-phrase *naze* "why." In this comparison, we can see a clear argument/adjunct asymmetry.[4] The in-situ *naze* in a *mosi* conditional is totally unacceptable, but in-situ *dono-ringo* is quite acceptable.

[4] Note that recently, it becomes clearer that the interpretation of *naze* and its interaction with islands are heavily influenced by prosodic pattern (Kitagawa 2006 and references therein). In this study, I keep presenting the data based on the judgments from native speakers I interviewed without much consideration in the influence of prosody, just for sake of the simplicity of the argument. However, a careful examination of the examples containing wh-phrases is, indeed, required.

(62) a. Quinn-wa mosi Stillman-ga dono-ringo-o
Q-top mosi S-nom which-apple-acc
tabeta-ra okorimasu-ka?
eat-cond get-angry-Q
Lit. "Which apple will Quinn get angry if Stillman eats?"

b. *Quinn-wa mosi Stillman-ga naze ringo-o
Q-top mosi S-nom why apple-acc
tabeta-ra okorimasu-ka?
eat-cond get-angry-Q
"Why$_1$ will Quinn get angry [if Stillman eats an apple t_1]?"

As we have seen in the earlier sections, *mosi* in Japanese conditionals is an optional adverb. Interestingly, if *mosi* is absent from conditional sentences, the extraction of both arguments and adjuncts becomes acceptable. The Examples in (63) are exactly the same as the examples in (62) except that mosi is not present in both examples.

(63) a. Quinn-wa Stillman-ga dono-ringo-o tabeta-ra
Q-top S-nom which-apple-acc eat-cond
okorimasu-ka?
get-angry-Q
Lit. "Which apple will Quinn get angry if Stillman eats?"

b. ?Quinn-wa Stillman-ga naze ringo-o tabeta-ra
Q-top S-nom why apple-acc eat-cond
okorimasu-ka?
get-angry-Q
"Why$_1$ will Quinn get angry [if Stillman eats an apple t_1]?"

For some speakers, *naze* in conditional clauses is still not perfect. However, the remarkable fact is that comparing examples like (62b) and (63b), (63b) is much better than (62b). This fact indicates that the presence or absence of CA *mosi* is crucial for the weak islandhood of Japanese conditional clauses. Differently put, we can plausibly consider that *mosi* is an inducer of WI effects.

Argument wh-phrases can also be overtly moved out of conditional clauses as the examples in (64) suggest. On the other hand, because long-scrambling of adjunct wh-phrases in Japanese is prohibited even out of complement clauses as in (65) (Nemoto 1993; Saito 1985) we cannot test whether the overt movement of *naze* creates the same results or not. However, the cases of long-scrambling suggest that as long as the extracted elements are referential arguments, conditional clauses are not islands.

(64) a. dono-ringo-o$_1$ Quinn-wa [mosi Stillman-ga t_1
which-apple-acc Q-top mosi S-nom
tabeta-ra] okorimasu-ka?
eat-cond get-angry-Q
Lit. "Which apple will Quinn get angry if Stillman eats?"

b. dono-ringo-o$_1$ Quinn-wa [Stillman-ga t$_1$
which-apple-acc Q-top S-nom
tabeta-ra] okorimasu-ka?
eat-cond get-angry-Q
Lit. "Which apple will Quinn get angry if Stillman eats?"

(65) *Naze$_1$ Quinn-wa Stillman-ni [Virginia-ga t$_1$
why Q-top S-dat V-nom
ringo-o tabe-ta to] it-ta no?
apple-acc eat-past comp say-past Q
"Why did Quinn say to Stillman that Virginia ate the apple?"

Let us turn our attention to referentiality. Miyagawa (1998) points out that two types of presumably non-referential elements are sensitive to weak islands in Japanese. One of his examples involves the so-called Floated Numeral Quantifier (FNQ). As quantity arguments in English are not referential, Numeral Quantifiers (NQ), modifiers denoting quantity, are also not referential. According to Miyagawa, NQ can be moved to the beginning of a sentence as in (66a). However, if the movement of NQ crosses a WI inducer, in his example it is the focus marker *sika* "only" (Hagstrom 1998; Hoji 1985; Pesetsky 2000; Takahashi 1990; Tanaka 2003a among many others) the sentence becomes unacceptable. In example (66b), the NQ *futa-tu* moves over the WI inducer *sika* and the example becomes unacceptable.

(66) a. *Futa-tu*$_1$ Quinn-ga ringo-o t_1 tabe-ta (koto).
2-cl$_{(thing)}$ Q-nom apple-acc eat-past (fact)
"Two, Quinn ate apples"

b. **Futa-tu*$_1$ Quinn-sika ringo-o t_1 tabe-nakat-ta
2-cl$_{(thing)}$ Q-Foc apple-acc eat-not-past
(koto).
(fact)
"Two, only Quinn ate apples"

In exactly the same way, overt movement of an indefinite quantifier shows WI sensitivity. Thus, if the indefinite quantifier *dareka* "some" is moved across *sika* the sentence becomes unacceptable.

(67) a. Virginia-ga gakusee-o *dareka* yon-da
V-nom student-acc some invite-past
(koto).
(fact).
"Hanako invited some student."

b. *Dareka*$_1$ Virginia-ga gakusee-o t_1 yon-da
some V-nom student-acc invite-past
(koto).
(fact).
"Virginia invited some student."

c. *Dareka$_1$ Virginia-sika gakusee-o t_1
some V-foc student-acc
yoba-nakat-ta (koto).
invite-not-past(fact).
"Only Virginia invited some student."

Given the WI sensitivity of these extractions, our expectation is the following. If *mosi* conditionals are WIs, extraction of these elements should be prohibited. First let us see cases of FNQ. There is a remarkable difference in acceptability between a case where a quantifier moves across *mosi* and one that does cross *mosi* or one without *mosi* i.e., among the examples in (68), only (68b) is unacceptable, where the quantifier moves across *mosi*. This contrast is exactly what we expect.

(68) a. Quinn-wa [(mosi) Virginia-ga ringo-o *futa-tu*
Q-top (mosi) V-nom apple-acc 2-cl$_{(thing)}$
tabe-ta-ra] okoru-daroo.
eat-past-cond get-angry-will
"Quinn will get angry if Virginia eats two apples."

b. *Quinn-wa [*futa-tu*$_1$ mosi Virginia-ga ringo-o t_1
Q-nom 2-cl$_{(thing)}$ mosi V-nom apple-acc
tabe-ta-ra] okoru-daroo.
eat-past-cond get-angry-will
"Quinn will get angry if Virginia eats two apples."

c. Quinn-wa [mosi *futa-tu*$_1$ Virginia-ga ringo-o t_1
Q-nom mosi 2-cl$_{(thing)}$ V-nom apple-acc
tabe-ta-ra] okoru-daroo.
eat-past-cond get-angry-will
"Quinn will get angry if Virginia eats two apples."

d. Quinn-wa [*futa-tu*$_1$ Virginia-ga ringo-o t_1
Q-nom 2-cl$_{(thing)}$ V-nom apple-acc
tabe-ta-ra] okoru-daroo.
eat-past-cond get-angry-will
"Quinn will get angry if Virginia eats two apples."

The same contrast emerges in examples involving indefinite quantifiers. Like Miyagawa's examples with *sika*, if the indefinite quantifier *dareka* moves over *mosi*, the example becomes unacceptable.

(69) a. Quinn-wa [(mosi) Virginia-ga tomodati-o *dareka*
T-top mosi H-nom friend-acc some
party-ni yon-da-ra] naki-dasu daroo.
party-dat invite-past-cond cry-start will.
"Quinn will start crying if Virginia invites some friend to the party."

b. *Quinn-wa [*dareka* mosi Virginia-ga tomodati-o t_1
Q-top some mosi V-nom friend-acc
party-ni yon-da-ra] naki-dasu daroo.
party-dat invite-past-cond cry-start will.
"Quinn will start crying if Virginia invites some friend to the party."

c. Quinn-wa [mosi *dareka* Virginia-ga tomodati-o t_1
Q-top mosi some V-nom friend-acc
party-ni yon-da-ra] naki-dasu daroo.
party-dat invite-past-cond cry-start will.
"Quinn will start crying if Virginia invites some friend to the party."

d. Quinn-wa [*dareka* Virginia-ga tomodati-o t_1
T-top some H-nom friend-acc
party-ni yon-da-ra] naki-dasu daroo.
party-dat invite-past-cond cry-start will.
"Quinn will start crying if Virginia invites some friend to the party."

Taken together, the extraction of non-referential elements further supports the view that *mosi* conditionals in Japanese are WIs. Specifically the contrast between extraction of the referential argument and NQs is the same contrast that we have seen in the English examples, i.e., referentiality matters for extraction out of *mosi* conditionals.

Let us turn to the D-linking test. As we have seen, examples from English suggest that D-linked wh-phrases are referential and thus they can move out of WIs but non-D-linked wh-phrases are not referential and not extractable from WIs. In our

earlier discussion, we have seen that this difference becomes clear if we see the contrast between the D-linked wh-phrases and aggressively non-D-linked wh-phrases. Pesetsky (1987) suggests that the so-called *ittai* wh-phrases in Japanese show the same behavior as aggressively non-D-linked wh-phrases in English (see also Lasnik and Saito 1992 for related discussion). Details aside, if *ittai* wh-phrases behave in the same way as aggressively non-D-linked wh-phrases, we can simply expect that *ittai* wh-phrases cannot be extracted out of WIs but D-linked wh-phases can. This expectation can be tested using WI effects induced by *sika*. The following examples show that under the scope of *sika*, a D-linked wh-phrase is acceptable but both a non-D-linked wh-phrase and an *ittai* wh-phrase are marginal at best.

(70) a. Quinn-sika dono-hon-o yoma-nai-no?
Q-only which-book-acc read-neg-Q
"Which book will only Quinn read?"

b. ??Quinn-sika nani-o yoma-nai-no?
Q-only what-acc read-neg-Q
"What will only Quinn read?"

c. ??/*Quinn-sika ittai nani-o yoma-nai-no?
Q-only ittai what-acc read-neg-Q
"What the hell will only Quinn read?"

These examples suggest that, as in the English examples, there is a D-linked/non-D-linked asymmetry with respect to extraction from WIs. Now, let us see if this D-linked/non-D-linked asymmetry holds in the case of conditionals. If *mosi* induces a

WI effect, we expect that the extraction of D-linked wh-phrases should be allowed from *mosi* conditionals but that extraction of non-D-linked wh-phrases or *ittai* wh-phrases should create degradation. Although a simple wh-phrase does not create severe degradation, this prediction seems to be confirmed. Even though the effects do not seem to me to be so strong, still there is a difference in acceptability between the case of the D-linked and Non-D-linked wh-phrases in (71).[5] (72) summarizes the examples without *mosi*, and all the examples are acceptable.

(71) a. Quinn-wa [mosi Virginia-ga dono-hon-o
Q-top mosi V-nom which-book-acc
yon-da-ra] bikkuri-simasu-ka?
read-past-cond be-surprised-Q
"Which book will Quinn be surprised if Virginia reads? "

b. ?Quinn-wa [mosi Virginia-ga nani-o
Q-top mosi V-nom what-acc
yon-da-ra] bikkuri-simasu-ka?
read-past-cond be-surprised-Q
"What will Quinn be surprised if Virginia reads?"

[5] The strength of the effects is different among speakers. Some detected clear differences but some do not. The author conducted acceptability judgment with six native speakers, including myself. Two of them detected clear differences but four speakers recognized the differences in the same direction, but for them the effect is rather mild. Further examination of the data is required.

c. ??Quinn-wa mosi Virginia-ga ittai nani-o
Q-top mosi V-nom ittai what-acc
yon-da-ra bikkuri-simasu-ka?
read-past-cond be-surprised-Q
"What the hell will Quinn be surprised if Virginia reads?"

(72) a. Quinn-wa Virginia-ga dono-hon-o yon-da-ra
bikkuri-simasu-ka?

b. Quinn-wa Virginia-ga nani-o yon-da-ra
bikkuri-simasu-ka?

c. Quinn-wa Virginia-ga ittai nani-o yon-da-ra
bikkuri-simasu-ka?

A note is in order on the acceptability of the case of simple wh-phrases. It is true that for some speakers, a simple wh-phrase in a *mosi* conditional sounds bad. However, at least to my ear, the example does not sound as bad as (71c), the example containing the aggressively-non-D-linked wh-phrase. Why does this pattern emerge? Here, I would like to present a potential account. Pesetsky (1987) offers one possible answer for this question. He argues that an argument wh-phrase such as *what* in English can be D-linked even though it does not have the form of *which*. My conjecture here is that what Pesetsky argues holds true for Japanese wh-phrases also. In other words, the example (71b) sounds good because the simple wh-phrase *nani* can potentially be D-linked. This conjecture leads us to the following prediction: if we can make simple wh-phrase non-D-linked, the example becomes worse. This prediction seems to be true. As we have seen, if we attach *ittai* to the wh-phrases, the examples become bad

regardless of whether the wh-phrase is D-linked or non-D-linked. Thus, the contrast between (71b) and (71c), can be understood that a simple wh-phrase like *nani* can be D-linked. Eexample (71c) is bad because the wh-phrase is forced to be non-D-linked because of the presence of *ittai*. To support this claim, we can consider cases of D-linked wh-phrases. Even wh-phrases with a D-linked form like *dono*-x "which-x" become unacceptable if *ittai* is attached to it under *mosi* conditionals. Recall also that a wh-phrse like *naze* "why", which cannot be D-linked/referential regardless of whether *ittai* is attached or not, is bad under *mosi* conditionals. These contrasts can be easily accounted for under our assumption that only referential wh-phrases can be extracted out of WIs, i.e., simple wh-phrases can be D-linked and referential, but *ittai* wh-phrases and adjunct wh-phrases cannot be D-linked and thus cannot be referential. Thus, the simple wh-phrases are extractable out of WIs, but the *ittai* wh-phrases are not. Although there may be several possible alternatives, if we adopt Pesetsky's idea outlined here, we can have a coherent account for the data we have.

(73) ??Quinn-wa mosi Virginia-ga ittai dono-hon-o
T-top mosi H-nom ittai which-book-acc
yon-da-ra bikkuri-simasu-ka?
read-past-cond be-surprised-Q

Finally, let us examine *how-many* questions in Japanese. As in English counterparts, Japanese *how many* phrases show an ambiguity. Consider (74) which is ambiguous. In the reading (74a), the existential quantifier 'n-many articles' scopes above 'should'. This reading involves a presupposition: for instance, the committee

has 10 topics on which 10 articles are supposed to be written. (74a) asks about the number of articles that *Quinn* is required to write by the committee. By contrast, the other reading in (74b), this sentence asks about the number of the articles that Quinn should write, say, by the end of this month. If we scramble the *how many* phrase to the beginning of the sentence, the same ambiguity is observed. (75) is an example with scrambling of the *how many* phrase. With this much in mind, let us turn to an example that contains a WI inducer, such as (76).

(74) Sono iinkai-wa [Quinn-ga **[nan-bon-no kizi-o]**
that committee-top[Q-nom [how many article-acc]
kakubekida to] hookokusita nodesuka?
should-write comp] reported Q
'How many articles did that committee report that Quinn should write?'

a. For what number n: there are n-many articles x such that that committee reported that Quinn should write x (wide reading)

b. (?)For what number n: that committee reported that it is necessary for there to be n many articles x such that Quinn writes x (reconstructed reading)

(75) **[nan-bon-no kizi-o]**$_1$ sono iinkai-wa
[how many article-ACC] that committee-NOM
[Quinn-ga **t**$_1$ kakubekida to] hookokusita nodesuka?
[Q-NOM should write COMP] reported Q
'How many articles did that committee report that Taro should write?'

a. For what number n: there are n-many articles x such that that committee reported that Quinn should write x (wide reading)

b. (?)For what number n: that committee reported that it is necessary for there to be n many articles x such that Quinn writes x (reconstructed reading)

(76) is minimally different from (75). In (76), the matrix subject is associated with the quantificational expression *hotondo dono* "almost every." On the other hand, in the(76), the matrix subject is not associated with an overt quantifier. This type of quantifier is another instance that induces WI effects in Japanese (Hoji 1985; Miyagawa 1998 among others). If this type of WI inducer is inserted, the example does not show the ambiguity anymore. Thus in (76), only the presuppositional reading is available, and the example does not have the reading (76b). This effect can be understood in the same way as we have seen in the English examples, i.e., because of the WI inducer, the amount reading, a non-referential reading, becomes unavailable.

(76) [nan-bon-no kizi-o]$_1$ hotondo dono iinkai-mo
[how many article-acc] almost every committee
[Quinn-ga t_1 kakubekida to] hookokusita nodesuka
[Q-nom should write COMP] reported Q
'How many articles did almost every committee report that Quinn should write?'

a. For what number n: there are n-many articles x such that almost every committee reported that Quinn should write x. (wide)

b. *For what number n: almost every committee reported that it is necessary for there to be n-many articles x such that Quinn writes x. (reconstructed)

Now our question is whether the same type of disambiguation takes place if we locate the *how many* question in a *mosi* conditional. The answer seems to be positive. Let us examine the following example. A subtle judgment is required. In (77) the presuppositional reading (77a) is available. However, the quantity reading (77b) is not available. On the other hand, if *mosi* is not present , both of the readings are available as illustrated in (78).

(77) Nan-satu-no-hon-ni$_1$ gakubuchoo-wa [mosi
how-many-books-dat department-chair-top mosi
Virginia-ga t$_1$ shohyoo-wo kakeba] shoogakukin-ga
V-nom review-acc write scholarship-nom
das-**eru**-no?
offer-CAN-Q?

Lit. "How many books can the department chair offer the scholarship to Virginia if she wrote a review to?"

a. For what number n: there are n-many books x such that the department chair can offer the scholarship to Virginia if she writes a review to x?

b. *For what number n: the department chair can offer the scholarship to Virginia if there are n-many books x such that she writes a review to x?

(78) Nan-satu-no-hon-ni$_1$ gakubuchoo-wa
how-many-books-dat department-chair-top
[Virginia-ga t$_1$ shohyoo-wo kakeba] shoogakukin-ga
V-nom review-acc write scholarship-nom
das-**eru**-no?
offer-CAN-Q?
Lit. "How many books can the department chair offer the scholarship to Virginia if she wrote a review to?"

a. For what number n: there are n-many books x such that the department chair can offer the scholarship to Virginia if she writes a review to x?

b. For what number n: the department chair can offer the scholarship to Virginia if there are n-many books x such that she writes a review to x?

The pattern in (77) parallels the English examples that we have seen, where the presuppositional reading is not available because of the wh-island. Therefore, this suggests that *mosi* conditionals indeed behave like WIs. On the other hand, the contrast between (77) and (78) confirms that the presence of *mosi* is a crucial factor for the WI effects of conditionals.

3.2.3. Summary

Our main observations so far are the following. *Mosi* conditionals show the signature properties of WIs, i.e., *mosi* conditionals show a referential/non-referential asymmetry. Thus, referential elements are allowed to move out of *mosi* conditionals but non-referential elements are not. To establish this claim we have examined four diagnostic cases: argument/adjunct asymmetry; NQ movement; D-linked/non-D-linked asymmetry; and *how many* reconstruction. Furthermore, we have seen that these asymmetries go away if *mosi* is absent from conditional clauses.

Based on these observations, we can draw the following conclusions. Japanese conditional clauses show various WI effects. However, these WI effects are not induced by conditional clauses per se, rather the Conditional Adverb *mosi* creates these effects. Thus, there are asymmetries between *mosi* conditionals and conditionals without *mosi*.

Given this conclusion, a question arises: Why does an item like *mosi* create WI effects? In the following sections we will try to answer this question.

3.3. The Distribution of *Mosi*

In this section, I will investigate the distributional properties of *mosi*. I will establish the following claim: the conditional adverbial *mosi* is licensed by a finite clause. Furthermore, I will show that an investigation of the distribution of *mosi* gives us a window into the internal structure of Japanese conditional clauses.

3.3.1. Japanese Nominal Conditionals

Japanese conditional marking is typically realized as an inflection on the verb. For example, in (79), the morpheme that marks the conditional (*nara*) appears after the tense marker of the embedded verb. There are, however, cases in which the same conditional marker is attached to NPs (1b).

(79) a. Quinn-ga ringo-o tabe-ta-nara
Q-nom apple-acc eat-tense-cond
"If Quinn eats an apple"

b. Quinn-nara gakkoo-ni iru-hazu-da.
Q-cond school-at be-must-copula
"If you are talking about Quinn, he is at school."

In this section, I will concentrate on examples like (79b), which I call Nominal Conditionals. In doing so, I have two goals in mind; first, to reveal the distributional properties of *mosi*, and second, to figure out the internal syntax of conditional clauses. I will show that nominal conditionals give us important clues towards achieving these two goals.

The basic line of argumentation that I will pursue is that the two constructions, nominal conditionals and standard conditionals, are derived from the same base structure. I will show that Rizzi's (1997) Split CP hypothesis gives us an important insight in this regard. What I will specifically claim is that a nominal conditional is a residue of clausal ellipsis (specifically the ellipsis of FinP) applied to a standard conditional clause.

3.3.2. Some Basics of Nominal Conditionals

First let us review some basic properties of nominal conditionals. In Japanese traditional grammar, the marker *nara* as used in nominal conditionals has been called *Teidai Joshi* (Topic Particle) (Takubo & Masuoka 1992). As the name suggests, this particle can be used basically in the same context as a typical topic marker such as *wa* (Kuno 1973). For example, in (80), both *nara* and *wa* are used in a typical context for a topic. In this context, the topicalized NP *Tanaka-san* must have some topic marker, and if a Nominative case marker such as *ga* is attached the sentence sounds bad. Both *wa* and *nara* can be used naturally in this context as opposed to *ga*.

(80) A: Quinn mi-na-katta-kai?
T see-neg-past-Q
"Have you seen Quinn?"

B: Quinn-nara/wa/*ga toshokan-de
Q-nara/wa library-at
benkyoo-site-ta-yo.
study-do-prog-part
"Quinn was studying at the library"

Besides its use as a topic marker, the nominal conditional has another interesting property. As can be seen in (79), nominal conditionals use the same morphology as standard conditional clauses. Japanese conditional clauses can host conditional adverbs (CAs) such as *mosi*. Interestingly, nominal conditionals also can host *mosi*, but in a limited environment, namely when there is a pronoun *sore*

referring to a clausal element. The intuition behind the presence of *mosi* is something like the following: if *mosi* is inserted in a nominal conditional the focus of the sentence is on the proposition in the context. On the other hand, if *mosi* is absent, the focus is on the noun or the entity that is marked by *nara*. An interesting fact is that *mosi* requires the presence of the pronoun *sore* but this pronoun does not require the presence of *mosi*. This fact suggests that *mosi* is dependent upon the presence of *sore* but not vice versa.

(81) a. *mosi* Quinn-ga ringo-o tabe-ta-nara
mosi Q-nom apple-acc eat-past-cond
"if Quinn eats an apple"

b. *(sore-ga) *mosi* Quinn-nara
it-nom mosi Q-cond
"if it is Quinn"

c. sore-ga Quinn-nara
it-nom T-cond
"if it is Quinn"

Given examples like (81b), an obvious question arises. Why does *mosi* in nominal conditional requires the pronoun *sore*? This brings us back to the bigger questions about what properties compose the differences and similarities between standard conditional clauses and nominal conditionals. In the following section, I will answer these questions. Furthermore, I will show that answering these questions gives us important clues to the internal syntax of Japanese conditional clauses.

3.3.3. The Conditional Focus Link

To clarify the licensing condition on *mosi* in nominal conditionals, in this section I will examine the syntax of the clausal pronominal *sore* in Japanese. Example (81b) suggests that the presence of the pronoun *sore* is crucial for the licensing of *mosi* in nominal conditionals. Therefore, revealing the nature of *sore* naturally leads us to the nature of *mosi* licensing. Once I reach the licensing of *mosi*, I will reveal the internal structure of Japanese conditionals using the distribution of *mosi* as a clue. In so doing, I will show that there is a tight connection between conditional clauses and various focus constructions in Japanese.

3.3.3.1. The Syntax of the Clausal Pronoun *sore* in Japanese

3.3.3.1.1. Nakao & Yoshida (2005) on Pronominal Sluicing

To understand the nature of *sore*, I will first review a previous study by Nakao & Yoshida (2005). Their analysis of *sore* is the basis for my analysis of nominal conditionals.

Nakao and Yoshida (henceforth, N&Y) observe an interesting property of *sore*. According to N&Y, the clausal pronoun *sore* has the properties of both pronouns (or nominals) and clauses. The intuition behind this view is that clausal structure is hiding behind the pronominal expression. To capture this intuition, N&Y analyze *sore* as the residue of ellipsis of CP. More specifically, they argue that the sluicing construction that contains *sore* is derived from the so-called specificational cleft construction through deletion of the presuppositional clause, in which *sore* occupies the place where the presuppositional clause is originally generated. N&Y

base their argument on the properties of *sore* in Japanese sluicing constructions. They call the sluicing construction containing *sore* Japanese Pronominal Sluicing (JPS).

N&Y's basic observations on JPS can be summarized in (82). Let us consider each of the observations.

(82) a. The wh-remnant in JPS shows connectivity effects.

b. The copular construction in JPS induces honorification agreement with *sore* but not with the wh-remnant.

c. JPS does not allow the inversion of presuppositional clause and focused elements.

Connectivity Effects

The wh-remnant in JPS shows variety of connectivity effects. (83a) is an example of binding and Case connectivity effects, and (83b) is an example of postposition connectivity effects. There are two important points in example (83a). Accusative Case in Japanese can be analyzed as a structural Case (Saito, 1982; Takezawa, 1985, among others). I assume, following Saito (1982), or Takezawa (1985), that it is assigned to an NP that is the sister of the verb, the direct object. Therefore, it is most plausible to think that the accusative case on the remnant phrase, *zibun-zisin* (self) in (83a) is assigned by the verb in the sister relation, not by any other means. Second, the verb *siru* (know) is not an ECM type verb (Hoji, Takano, Hiraiwa, Tanaka Takezawa among others). Thus it is not likely that *siru* or its negative form *sira-nai* assigns accusative case to the remnant NP. These two points

strongly suggest that the remnant NP is assigned accusative case by the verb that is in the elided site, i.e., *seme-ta* "blame." Turning to the binding connectivity, the local anaphor *zibun-zisin* (self) is bound by the subject in the first conjunct. *Zibun-zisin*, however, requires a local c-commanding antecedent (ref) as example (84b) indicates. Thus, it is not plausible to think that the subject in the first conjunct—which does not c-command the local anaphor—directly binds it. Rather, this example strongly suggests that the anaphor is bound by an antecedent in the elided structure. The combination of the two connectivity effects, Case and local anaphor licensing, strongly suggest that there is hidden clausal structure in JPS. The example in (83b), the connectivity effect involving postpositions, points to exactly the same conclusion. The postpositional phrase in (83b) is selected by a particular class of verb such as *okuru* (send), and it is not compatible with a verb like *wakar* (know) in Japanese. Thus, as in the case of Case connectivity, the remnant postpositional phrase should be selected by the verb that is in the elided structure.

(83) a. Quinn$_1$-ga dareka-**o** seme-ta ga, watasi-wa
Q-nom someone-acc blame-past but, I-top
sore-ga zibun-zisin$_1$-**o** kadooka sira-nai.
it-nom self-acc whether know-not
"Quinn$_1$ blamed someone, but I don't know whether it was himself$_1$."

b. Quinn-ga dokoka-**kara** nimotu-o
Q-nom somewhere-from baggage- acc
okuttarasii ga, watasi-wa sore-ga Tokyo-**kara**
sent-seem but I-top it-nom Tokyo-from
kadooka wakara-nai.
whether know-not
"It seems that Quinn sent his baggage to us from somewhere, but I don't know from where"

(84) a. Quinnm-wa Virginia-o*/ga[ok] tensai-dearu kadooka
Q-nom V-acc/nom genius-be whether
sira-nai.
know-not
"Quinn does not know whether Virginia is a genius or not."

b. Quinn$_1$-wa Stillman$_2$-no hahaoya$_3$-ga
Q-top S-gen mother-nom
zibun-zisin$_{1*/2*/3}$[ok]-o semeta-to itta.
self-acc blame-comp said
"Quinn said that Stillman's mother blamed herself."

Honorification Agreement

The second property of JPS is its agreement property. In Japanese generative grammar, honorification has been treated as a grammatical agreement phenomenon (Harada 1976, Boeckx and Niinuma 2003 among others). Although this analysis contradicts the general view that Japanese is a language without agreement (Fukui

1995, Kuroda 1988), the reason that honorification is treated as an agreement phenomenon is mainly because it is sensitive to a certain grammatical function, and induces a certain type of locality effect. Before examining honorification in JPS, let us briefly review this point.

Harada pointed out that what he calls propositional honorifics are sensitive to grammatical functions such as subject or object. Basically, both subject honorification and object honorification induce special morphology on the predicate. Subject honorification is marked by the form of *o-predicate-ni-naru* if the subject is a person socially superior to the speaker and object honorification is marked by the form of *o-predicate suru* if the object is a person socially superior to the speaker, respectively. They are illustrated in the following examples.

(85) a. Subject Honorification

Sasaki-sensei-wa watasi-ni koo

S-teacher-top I-dat this-way

o-hanasi ni nat-ta.

speak-past

"Mr. Sasaki told me this way"

b. Object Honorification

Watasi-wa Sasaki-sensei-ni koo

I-top S-teacher-dat this-way

o-hanasi si-ta.

speak-past

"I told Mr. Sasaki this way."

An important point of the paradigm above is that subject honorification cannot be used in the context of object honorification or vice versa. If honorification were not a grammatical phenomenon, this sensitivity to the grammatical functions would not be expected. This point can be easily shown by switching the subject and the object of each example in (85). In (86a), verb has the subject honorification marker despite the fact that the subject is not a person socially superior to the speaker, and in (86b) object honorification is induced but the object is not the person socially superior to the speaker. Both examples are indeed unacceptable.

(86) a. *Watasi-wa Sasaki-sensei-ni koo
I-top S-teacher-dat this-way
o-hanasi-ni nat-ta.
speak-past
"I told Mr. Sasaki this way."

b. *Sasaki-sensei-wa watasi-ni koo
S-teacher-top I-dat this-way
o-hanasi si-ta.
speak-past
"Mr. Sasaki told me this way."

Another supporting argument for the position that honorification is a grammatical phenomenon comes from the fact that it is subject to a certain type of locality condition (Boeckx and Niinuma 2003). According to Boeckx and Niinuma, object honorification is an agreement relation between the small V and an object.

They further argue that this agreement relation is disrupted if there is an offending intervener between the small V and the object NP. To illustrate this point, they cite the examples in (87). The pattern illustrated by these examples is that the verb can have object honorification when the indirect object is a person socially superior to the speaker, but it cannot if the direct object is the honorification inducer. The comparison between (87b) and (87c) further supports this point. If the indirect object is not present, the sentence sounds much better.

(87) a. Virginia-ga Tanaka-sensei-ni Quinn-o
V-nom T-teacher-dat Q-acc
go-syookai si-ta.
introduce-past
"Virginia introduced Quinn to Mr. Tanaka."

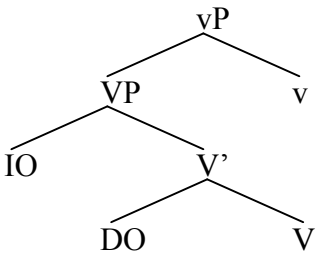
b. *Virginia-ga Quinn-ni Tanaka-sensei-o
V-nom Q-dat T-teacher-acc
go-syookai si-ta.
introduce-past
"Virginia introduced Mr. Tanaka to Quinn."

c. Virginia-ga Tanak-sensei-o go-syookai si-ta.
V-nom T-teacher-acc introduce -past
"Virginia introduced Mr. Tanaka (to someone)."

Boeckx and Niinuma argue that this pattern indicates that object honorification is subject to locality constraints. Simply put, if there is an indirect object, the direct

object cannot control the object honorification. They argue that honorification agreement is established by the operation Agree, and Agree is subject to locality conditions, the so-called Defective Intervention Constraint (Chomsky 2000). Assuming that the indirect object is base generated at the higher position than the direct object, i.e., the indirect object c-commands the direct object, the indirect object intervenes v and the direct object. Schematically, (88) illustrates this point, where the indirect object is intervening between v and the direct object.

(88)



Boeckx and Niinuma cite various types of supporting evidence for their position, and argue that honorification is best analyzed as being established by the operation Agree rather than by movement, thus supporting the validity of a syntactic theory which incorporates Agree.

Putting the technical details aside, for us the most crucial point in both Harada's and Boeckx and Niinuma's studies is that honorification is a grammatical phenomenon, and best treated as an agreement relation between a verb and its object.

With the discussion so far in mind, let us turn to honorification agreement in JPS. The crucial example for us is in (89). In (89), the copula in the second conjunct shows subject honorification. The wh-word *dare* is the object position and does not

have the form of a person socially superior to the speaker. If the wh-word is used for a socially superior person, it has the form of *dotira* or *dono-kata* but not *dare*. Thus, it is clear that the honorification is induced by *sore* but not by the wh-remnant. This fact, that if *sore* is not present, the example is not acceptable, shows that the agreement is established between *sore* and the copula.

(89) ?Quinn-ga aru kata-ni **o-ai si-ta**
Q-nom certain person-dat meet(Obj.Hon) -past
rasii ga watasi-wa sore-ga dare-ni
seem but I-top it-nom who-dat
de-irassyat-ta ka sira-nai.
be(Subj.Hon)-past Q know-not
"It seems that Quinn met a certain person, but I don't know who it was."

This agreement pattern can be captured if we assume there is a hidden clausal structure in the guise of the pronoun *sore* as illustrated in (90). (90) is a cleft construction that Nakao and Yoshida assume to underlie JPS. In (90), the presuppositional clause occupies the subject position. Basically the presuppositional clause can be analyzed as a nominalized clause that refers to the NP in the focus position. If this is correct, the presuppositional clause can be honorificational, and this is plausible because the gap corresponds to the focalized wh-phrase, *dare-ni* "to whom" is honorified by the object honorificational verb in the presuppositional clause. The copular in the second conjunct in (89) agrees with this hidden

presuppositional clause, and because it is in the subject position, it induces subject honorification.

(90) …watasi-wa [$_{CP}$[$_{IP}$ Quinn-ga t_1 **o-ai si-ta**] no]-ga
I-top Q-nom meet(Obj.Hon) comp-nom
dare$_1$-ni de-irassyat-ta ka sira-nai.
who-dat be(Subj.Hon.)-past Q know-not
"I don't know who it was who Quinn met."

If we do not assume that the presuppositional clause is underlying *sore*, it is not clear why subject honorification can take place in JPS. Putting the use as a clausal pronoun aside, the most basic use of *sore* is to refer to an inanimate object. Thus *sore* itself does not have a feature that can induce honorification agreement. If we assume that *sore* is a simple pronoun without any clausal structure, and that it can be coreferential to its antecedent or interpreted in the way that deep anaphora is interpreted (Hoji 1995), then we have to assume that somehow the honorificational feature as well as an animacy feature is assigned to the pronoun via a coreference relation or a deep anaphoric relation because, as we have seen, honorification is an agreement phenomenon. However, such feature assignment does not seem to be observed in the other environments. On the other hand, under Nakao and Yoshida's analysis, the honorification phenomenon in JPS is given a straightforward explanation, i.e., there is a hidden presuppositional clause that has a feature for honorification.

Finally, Nakao and Yoshida point out that the same kind of agreement pattern can be seen in the English specificational pseudocleft construction. (91) shows that if a pseudocleft construction is interpreted specificationally, the copular agrees in number with the pre-copular presuppositional clause, rather than with the post-copular focus phrase.

(91) [What you have bought] **is** fake **jewels**.

This parallelism between JPS and English specificational pseudoclefts support the validity of Nakao and Yoshida's analysis.

Ban on Inversion

The fourth property of JPS is the ban on inversion of the pre- and post-copular elements.

For some unknown reasons, Japanese specificational constructions do not allow inversion of the presuppositional clause and the focused phrase.

(92) a. [Quinn-ga katta no]-wa hon-o
Q-nom bought comp-top book-acc
san-satu da.
three-classifier be
"What Quinn bought was three books "

b. *Hon-o san-satu-wa [Quinn-ga
book-acc three-cassifier-top Q-nom
katta no] da.
bought comp be
"What Quinn bought is three books "

In the same way, JPS does not allow the inversion of the pronoun *sore* and focused phrase. (93a) is an example of JPS with the regular order of the presuppositional clause and the focused phrase. In (93b), on the other hand, they are inverted, and the example is not acceptable.

(93) Quinn-ga nanika-o katta-rasii ga
Q-nom something-acc bought-seem but
a. watasi-wa sore-ga nani-o (da) ka sira-nai.
I-top it-nom what-acc be Q know-not
b. *watasi-wa nani-ga sore (da) ka sira-nai.
I-top what-nom it be Q know-not
"It seems that Quinn bought something, but I don't know what it is."

The ban on inversion also indicates a strong parallelism between the specificational construction and JPS. Under the assumption that the unseen presuppositional clause is represented, the ban on inversion is given a straightforward explanation, i.e., because the presuppositional clause and the focused phrase cannot be inverted, the pronoun *sore* and the focused phrase cannot be inverted either.

3.3.3.2.Summary

So far I have shown various reasons to believe that some clausal structure, the presuppositional clause of the pseudocleft construction, underlies *sore* in JPS. The connectivity effects indicate that there is a clausal structure in *sore*, and the parallelisms between JPS and pseudocleft suggest that these two are actually the same elements, and thus support the claim that *sore* corresponds to the presuppositional clause.

3.3.3.3. *Sore* in Nominal Conditionals

Based on the discussion so far, I will argue that *sore* in nominal conditionals is exactly the same element as the one found in JPS. To this end, I will show that all the properties of JPS and *sore* in JPS that we have reviewed hold true for nominal conditionals with *sore*.

The properties of *sore* in JPS that we have seen so far are summarized in (94). As I summarize in the chart in 0), all of the properties of *sore* in JPS are seen in the *sore* in nominal conditionals. Let us examine these properties one by one.

(94) a. The wh-remnant in JPS shows connectivity effects.

b. The copular in JPS induces honorification agreement with *sore* but not with the wh-remnant.

c. JPS does not allow the inversion of pre- and post-copular elements.

(95)

	Properties		
	Connectivity	*Honorification*	*Inversion*
JPS	√	√	√
Nominal Conditionals	√	√	√

The examples in (96) show that *sore* in nominal conditionals exhibits connectivity effects of Case, local anaphor licensing and postpositions. Exactly like the examples of JPS, *sore* in nominal conditionals shows connectivity effects of local anaphor licensing and Case in (96a) where the anaphor *zibun-zisin* is bound by the subject in the first conjunct, and accusative case is assigned to the remnant, and the connectivity effect of the postposition *kara* (from) in (96b).

(96) a. Quinn$_1$-ga dareka-**o** semete-iru rasii ga,
Q-nom someone-acc blame-prog seem but
mosi *(sore-ga) **zibunzisin$_1$-o** nara
mosi it-nom self-acc cond
yameta hoo-ga yoi.
stop it-is-better
"It seems that Quinn is blaming someone but if it is himself, he had better stop doing that."

b. Quinn-ga dokoka-kara nimotu-o
Q-nom somewhere-from baggage-acc
okutta-rasii ga mosi *(sore-ga) Tokyo-kara nara
sent-seem but mosi it-nom Tokyo-from cond
sugu-todoku daroo.
soon-arrive will
"It seems that Quinn sent the baggage from somewhere, but if it is from Tokyo, the baggage will arrive soon."

The pronominal nominal conditional shows a similar honorification pattern to JPS. In (97) the copular shows subject honorification even though the honorifiable person, *Yamada-sensee* (Mr. Yamada) is in the object position. Thus, as in JPS, we can conclude that the copular agrees with *sore*, and so the same explanation of honorification agreement holds for pronominal nominal conditionals.

(97) Quinn-wa Yamada-sensee ka Sasaki-sensee-ni
Q-top Y-teacher or S-teacher-dat
oai-sita rasii ga, mosi sore-ga Yamada-sensee-ni
meet(Obj.Hon.) seem but mosi it-nom Y-teacher-dat
de-irasshatta-nara kitto kinchoo-sita daroo.
cop(Subj.Hon)-cond certainly nervous-did may
"It seems that Quinn met Mr. Yamada or Mr. Sasaki, but if it was Mr. Yamada, Quinn was certainly nervous."

Finally inversion of *sore* and the remnant phrase is not allowed in pronominal nominal conditionals, as (98) shows.

(98) Quinn-wa ringo ka mikan-o tabe-ta-rasii ga
Q-top apple or orange-acc eat-past-seem but

a. mosi sore-ga ringo-nara
mosi it-nom apple-cond

b. *mosi ringo-ga sore-nara
mosi apple-nom it-cond
oisi-katta hazu-da.
tasty-past must-copular

"Quinn ate an apple or an orange, but if it was an apple that he ate, it must have been tasty."

These parallelisms between JPS and pronominal nominal conditionals strongly suggest that *sore* in these constructions is the same element. If we analyze these elements in different ways we cannot account for the parallelism of these two *sore*s.

3.3.3.4. *Sore* and *Mosi*

The striking parallelism between JPS and pronominal nominal conditionals strongly suggests that there is a clausal structure in the guise of a pronominal in these constructions. The examples of connectivity effects strongly suggest the existence of clausal hierarchical structure in *sore* in these examples. Relatedly, JPs and nominal

conditionals show yet another similarity, i.e., the connectivity effects are not legitimate if *sore* is not present. We can see this point in the examples in (99). In both JPS and nominal conditionals, the examples without *sore* are degraded with the local anaphor, i.e., the local anaphor is not bound in the examples without *sore*.

(99) Connectivity effects

a. Quinn$_1$-ga dareka-**o** seme-ta ga, watasi-wa
Q-nom someone-acc blame-past but, I-top
*(sore-ga) zibun-zisin$_1$-**o** kadooka sira-nai.
it-nom self-acc whether know-not
"Quinn$_1$ blamed someone, but I don't know whether it was himself$_1$."

b. Quinn$_1$-ga dareka-**o** semete-iru rasii ga,
Q-nom someone-acc blame-prog seem but
*(sore-ga)**zibunzisin$_1$-o** nara yameta hoo-ga yoi.
it-nom self-acc cond stop it-is-better
"It seems that Quin is blaming someone but if it is himself, he better stop doing that."

This correlation between the presence of *sore* and connectivity effects can be easily understood if there is indeed a clausal structure in the second conjunct when *sore* is present, but there is no such hidden clause when *sore* is absent. Interestingly, the distribution of *mosi* also observes an analogous effect of connectivity. In other words, if *sore* is not present, *mosi* is not licensed, in much the same way that the local anaphor is not licensed.

(100) Quinn$_1$-ga dareka-**o** semete-iru rasii ga,
Q-nom someone-acc blame-prog seem but
mosi *(sore-ga) zibunzisin$_1$-o nara yameta
mosi it-nom self-acc cond stop
hoo-ga yoi.
it-is-better
"It seems that Quinn is blaming someone but if it is himself, he better stop doing that."

It should be clear that when *mosi* is licensed, *mosi* clearly requires the clausal element that is replaced by the pronoun *sore*. If not, the unacceptability of the above example is not easily understood. Thus, I conclude that *mosi* is licensed by a clause that can be replaced by the clausal pronoun *sore*. Now the question is the identity of the "clausal element" that is replaced with *sore*. If we get a better idea of the structure of this clausal element, we can get a better understanding of the licensing condition on *mosi*.

3.3.3.5.A Problem Remains

Before turning to the investigation of the identity of the clausal element, let us note one problem with the current analysis that *sore* replaces the presuppositional clause of pseudocleft.

As it is well known, NPs in Japanese can be phonologically empty. The inanimate pronoun *sore*, for example, can be empty. In (101b), the object that corresponds to the pronoun *sore* in A's utterance is omitted.[6]

(101) a. A: Quinn-wa **sore**-o tabe-ta-no?
Q-top it-acc eat-past-Q?
"Did you eat that, Quinn?"

b. B: Un, ø tabe-ta-yo.
yes, eat-past
"Yes, I ate that."

Given this property of empty argument licensing in Japanese, the most straightforward expectation with respect to the clausal pronoun *sore* is that it can also be empty. However, the fact that connectivity effects cannot be observed in the sluicing construction or nominal conditionals in the absence of *sore* suggests to us the opposite. If empty clausal pronoun *sore* exists, we expect no differences between Japanese standard sluicing constructions or nominal conditionals without *sore* and JPS and pronominal nominal conditionals. Even if we assume some empty element is

[6] Various analyses of this phenomenon have been proposed. These proposed analyses can be classified into two types. One is known as the *pro* analysis (Hoji 1999), and the others are VP-ellipsis analyses (Otani and Whitman 1991). The treatment of this type of construction is still controversial. I will leave this problem open. For us, it is important that the object corresponding to the inanimate pronoun *sore*, or the larger constituent that contains *sore*, can be omitted.

there, we are forced to conclude that this empty element is different from the clausal pronoun *sore*.

3.3.3.6. Toward the Licensing Condition of *Mosi*

So far, I have shown that the clausal pronoun *sore* in JPS and the pronominal nominal conditional are the same element by showing the parallelisms between these two constructions. The conclusion we have reached is that *mosi* requires a clausal structure that can license it. In this subsection, I will show that by adopting the analysis of focus construction in Japanese proposed by Hiraiwa and Ishihara, we can successfully derive these parallelisms, and go closer to the nature of the Conditional Adverb *mosi*.

In this subsection, I will establish the following two claims. First, Japanese conditional clauses have an articulated CP structure that can be best analyzed by Rizzi's (1997) Split-CP analysis. Second, *mosi* is licensed by FinP under the Split-CP analysis. In the course of the discussion, I will further propose a general licensing condition on a certain type of clausal ellipsis in Japanese.

3.3.3.6.1. Rizzi (1997)

The heart of Rizzi's (1997) Split-CP hypothesis is that just as there is evidence that IP is better analyzed by splitting it up into multiple projections, there is also evidence that CP should be split up into several projections. Specifically, he claims that CP has the internal makeup illustrated in (102).

(102) $[_{ForceP}$ Force $[_{TopP*}$ Top $[_{FocP}$ Foc $[_{TopP*}$ Top $[_{FinP}$ Fin $[_{IP}$ …]]]]]]

Rizzi shows numerous pieces of empirical evidence for the split-CP system, but before reviewing his empirical arguments, let us look at his conceptual motivation for splitting CP up into multiple functional categories.

It has been traditionally argued that CP has selectional relations with the verb that takes it as a complement and with the clause that it takes as a complement. Viewed from outside, different types of clauses are selected by different verbs. For example, *believe* selects declarative clauses, *wonder* selects interrogative clauses, and *know* selects either declarative clauses or interrogative clauses. In this way, CP marks different *clause types* in accord with the semantics of the verb that takes the CP as its complement (Bresnan, Chen, Chomsky, etc). Rizzi calls this function of CP marking clause types the *force system*.

(103) a. Quinn believes that Virginia went to school.

b. *Quinn believes who went to school.

(104) a. *Quinn wonders that Virginia went to school.

b. Quinn wonders who went to school

(105) a. Quinn knows that Virginia went to school.

b. Quinn knows who went to school.

On the other hand, C itself selects a certain type of clause. Just as a specific type of verbs takes a specific type of CP, a specific complementizer takes a specific

type of clauses. Taking an example from English, the complementizer *that* takes a finite IP as its complement, but *for* takes an infinitival IP as its complement. He calls this a *finiteness system*.

(106) a. Quinn thinks that Virginia went to school.

b. *Quinn thinks that Virginia to go to school.

(107) a. Quinn wants very much for Virginia to go to school.

b. *Quinn wants very much for Virginia went to school.

Thus, it seems that CP is responsible for two different functions. Rizzi's claim is that these two different functions can be best expressed as properties of different heads in "CP."

Besides his conceptual argument, Rizzi offers many empirical arguments for the existence of multiple layers in CP. Below, I pick several of his arguments to discuss so that we can see how things differ under the split-CP framework.

Rizzi bases his arguments for the existence of FocP and TopP on the relative order of focalized and topicalized elements with respect to each other and to "complementizers." Let us start from the relative orderings of topicalized phrases and complementizers. In Italian, the complementizer *di* that introduces non-finite embedded clauses follows topicalized phrases. On the other hand, another complementizer, *che*, that introduces finite embedded clauses, precedes them.

(108) a. Credo **che** loro apprezzerebbero molto il tuo libro.

"I believe **that** they would appreciate your book very much."

b. Credo **che** il tuo libro, loro *lo* apprezzerebbero molto.

that your book

"I believe **that** your book, they would appreciate *it* a lot."

c. *Credo, il tuo libro, **che** loro *lo* apprezzerebbero molto.

your book **that**

"I believe, your book, **that** they would appreciate *it* a lot. "

(109) a. Credo **di** apprezzare molto il tuo libro.

"I believe **of** to appreciate your book very much."

b. *Credo **di** il tuo libro, apprezzar*lo* molto.

of your book

"I believe **of** your book, to appreciate it a lot. "

c. Credo il tuo libro, **di** apprezzar*lo* molto.

your book **of**

"I believe your book **of** to appreciate *it* a lot."

The relative ordering of complementizers and topicalized phrases cannot be easily accommodated by a system with unique C position. However, the split-CP system that Rizzi is adopting can easily capture their distribution by assuming that *che* is the Force^0 element and *di* is Fin^0 element respectively as is illustrated in (110).

(110) $[_{\text{ForceP}}$ *che* $[_{\text{TopicP}} \ldots [_{\text{FinP}}$ *di* $[_{\text{IP}} \ldots]]]]$

Rizzi also examines the relative order of wh-words and topicalized phrases, to make the same point. What he specifically shows is that wh-words in relative clauses must precede topicalized phrases, but interrogative wh-phrases must follow them. Schematically the ordering can be illustrated as in (111). Relevant examples are (112) and (113), both from Italian.

(111) … Relative-Wh … Topic Phrases … Interrogative-Wh … IP

(112) Wh-words in Relative Clauses

a. Un uomo **a cui**, il premio Nobel, *lo* daranno senz'altro.

to whom the Nobel Prize

"A man **to whom**, the Nobel Prize, they will give *it* undoubtedly."

b. *Un uomo, il premio Nobel, **a cui** *lo* daranno senz'altro.

the Nobel Prize **to whom**

"A man, the Nobel Prize, **to whom** they will give *it* undoubtedly"

(113) Interrogative Wh-words

a. ***A chi**, il premio Nobel, lo daranno?

to whom the Nobel Prize

"**To whom**, the Nobel Prize, will they give it?"

b. Il premio Nobel, **a chi** *lo* daranno?

the Nobel Prize **to whom**

"The Nobel Prize, **to whom** will they give it?"

These examples suggest that wh-phrases in relative clauses occupy the specifier of the highest phrase, and that interrogative wh-phrases occupy the specifier of a lower phrase than TopP, and this in turn suggests that interrogative wh-phrases occupy the specifier of FocP. To support this claim, Rizzi further shows that interrogative wh-phrases cannot cooccur with focalized phrases, but relative wh-phrases can and must precede them. Rizzi cites the examples in (114) and (115). The examples in (114) suggest that interrogative wh-phrases and focus phrases compete for the same spot. Assuming that there is just one position for focus, at least in Italian, the contrast in (114) can be easily captured under the split-CP system, i.e., interrogative wh-phrases occupy specifier of FocP that is lower than TopP. Turning to (115), we see that relative wh-phrases are compatible with focused phrases and, additionally, that they must precede focus phrases. This is predicted by the split-CP system that Rizzi assumes, if we hypothesize that relative wh-phrases occupy the specifier of ForceP and that FocP is located lower than ForceP.

(114) Focus is not compatible with interrogative wh-phrases

a. *<u>A GIANNI</u> **che** cosa hai detto (, non a Piero)?
to Gianni what
Focus *Wh*
"TO GIANNI what did you tell (not to Piero)?"

b. ***Che** cosa A GIANNI hai detto (, non a Piero)?

what to Gianni

Wh *Focus*

"What TO GIANNI did you tell (not to Piero)?"

(115) Focus is compatible with relative wh-phrases

a. Ecco un uomo **a cui** IL PREMIO NOBEL dovrebbero dare (non il

to whom THE NOBEL PRIZE

Rel.Wh *Focus*

premio X).

"Here is a man to whom THE NOBEL PRIZE they should give (not prize X)."

b. *Ecco un uomo IL PREMIO NOBEL **a cui** dovrebbero dare…

THE NOBEL PRIZE to whom

Focus *Rel.Wh*

Taken together, the distribution of these elements can be accurately captured by the split-CP system, but not by traditional analyses assuming that there is single CP projection.

Rizzi shows many other types of supporting evidence for his split-CP system. For us, however, the above should be sufficient to see the motivations for splitting CP up into multiple layers.

3.3.3.6.2. Focus Constructions in Japanese: Hiraiwa & Ishihara (2001)

Based on Rizzi's split-CP system, Hiraiwa and Ishihara offer an analysis of several Japanese focus constructions. Specifically, they pick three focus constructions that have been known to have a lot of similarities, but for which no analyses were previously offered which could capture these similarities, and they show that the properties of these constructions can be readily derived by Rizzi's articulated CP-system. In this subsection, I will briefly review their study, and build a bridge to the analysis of conditional clauses.

Properties of Focus Constructions

In Japanese there are several focus constructions that share significant similarities. Hiraiwa and Ishihara call them the cleft construction, the *no-da* in-situ focus construction, and the sluicing construction, respectively. Examples of these constructions are summarized in (116).

(116) a. Cleft

[Quinn-ga tabe-ta no]-wa kono ringo-o
Q-nom eat-past comp-top this apple-acc
(mit-tsu) da.
three-cl cop
"It is this apple that Quinn ate."

b. No-da

[Quinn-ga kono ringo-o tabe-ta no] da.
Q-nom this apple-acc eat-past comp cop

"It is that Quinn ate this apple."

c. Sluicing

Quinn-ga Virginia-ni nanika-o
Q-nom V-dat something-acc
age-ta arasii ga boku-wa [nani-o (da) ka]
give-past seem but I-top what-acc cop Q
wakara-nai.
know-not.

"Quinn gave something to Virginia but I don't know what."

Let us first summarize the important properties of the cleft construction, and see to what extent the properties of clefts are shared by the other two constructions.

Properties of Clefts

Hiraiwa and Ishihara report five main properties of clefts, which I list in (117).

(117) a. The possibility of multiple foci

b. Island effects

c. The complementizer *no* cannot be substituted with a pronoun/NP

d. The unavailability of Nominative/Genitive Conversion (NGC)

e. Clause mate condition on multiple foci

As is well known, Japanese focus constructions (including cleft constructions), allow multiple foci (Koizumi 1995). In (118a), two NPs are in the focus position. Interestingly, if any of these focused elements are not case-marked, multiple foci are not allowed. This is, however, not true when there is just one focus element. If there is just one focus element, the case marker can be dropped. The cleft construction containing the focus element without case is sometimes called a Pseudocleft (Kuroda 1999) or a Non-Case Marked-Cleft (Hoji 1990). Here following Hiraiwa and Ishihara's notation, I will refer to it as a pseudocleft.

(118) a. [Quinn-ga age-ta no]-wa Virginia-ni
Q-nom give-past comp-top V-dat
ringo-o (mit-tsu) da.
apple-acc three-cl cop
"(Lit.) It is (three) apples to Virginia that Quinn gave."

b. *[Quinn-ga age-ta no]-wa Virginia*(-ni)
Q-nom give-past comp-top V-dat
ringo*(-acc) (mit-tsu) da.
apple-acc three-cl cop

(119) [Quinn-ga Virginia-ni age-ta no]-wa ringo da.
Q-nom V-dat give-past comp-top apple-ø cop.
"It is an apple that Quinn gave to Virginia."

The second property is island sensitivity. As we can see in the contrast in (120), clefting but not pseudoclefting is sensitive to islands. In the cleft sentence, (120a), the focus element is extracted out of a complex NP and the sentence is not acceptable. On the contrary, the pseudocleft sentence in (42b) is acceptable even though the focus element is extracted out of a complex NP. This contrast is taken to indicate that clefting involves movement whereas pseudoclefting does not.

(120) a. *[Quinn-ga [[e_1 e_2 kai-ta] hito$_1$]-o
Q-nom write-past person-acc
hihansi-ta no]-wa kono-ronbon-o$_2$ da.
criticize-past comp-top this-paper-acc cop
"(Lit) It is this paper that Quinn criticized the person who wrote."

b. [Quinn-ga [[e_1 e_2 kai-ta] hito$_1$]-o
T-nom write-past person-acc
hihansi-ta no]-wa kono-ronbon$_2$ da.
criticize-past comp-top this-paper-ø cop
"(Lit) It is this paper that Quinn criticized the person who wrote."

The third property is relevant to the properties of the complementizer involved in clefting. In a cleft construction, the complementizer *-no* cannot be substituted with an NP. In a pseudocleft, however, such alternation is possible.

(121) a. [Quinn-ga tabe-ta no/*mono/*kudamono]-wa
Q-nom eat-past comp/thing/fruit -top
ringo-o (mit-tu) da.
apple-acc three-cl be
"It was (three) apples that Quinn ate"

b. [Quinn-ga tabe-ta no/mono/kudamono]-wa ringo
Q-nom eat-past comp/thing/fruit-top apple-ø
(mit-tu) da.
three-cl be
"It is (three of) apples that Quinn ate"

The fourth property is concerned with the status of the complementizer *no* that is used in cleft and pseudocleft constructions. As we have reviewed in section 2 in this chapter (relative clause section), a nominative NP in Japanese optionally converts to a genitive NP in a structure in which a verb bears adnominal form (see Hiraiwa 2002; Watanabe 1996). This is called Nominative-Genitive Conversion. Even though both in cleft and pseudocleft constructions the predicate bears adnominal form, cleft resists nominative-genitive conversion whereas pseudocleft allows it. The following examples illustrate this contrast.

(122) a. Cleft
[Quinn-ga/??no tabe-ta no]-wa kono ringo-**o**
Q-nom/??gen eat-past comp-top this apple-acc
(mit-tu) da.
three-cl cop
"It is (three of) these apples that Quinn ate."

b. Pseudocleft

[Quinn-ga/no tabe-ta no]-wa kono ringo

Q-nom/??gen eat-past comp-top this apple-ø

(mit-tu) da.

three-cl cop

"It is (three of) these apples that Quinn ate."

Finally, the fifth property of clefts that Hiraiwa and Ishihara point out is the *clausemate* condition. As we have seen, unlike pseudoclefting, clefting allows multiple foci. The cleft with multiple foci, which is called a multiple cleft, is known to be subject to the so-called Clause-Mate Condition: The elements undergoing Multiple Clefting must be clausemates. (123b) and (123c) are derived from the same base sentence (123a). (123b) violates the clausemate condition and the sentence is unacceptable, but (123c) satisfies the condition and it is acceptable.

(123) a. Base sentence

Virginia-ga sensei-ni [Quinn-ga kono ringo-o

V-nom teacher-dat Q-nom this apple-acc

tabe-ta to] iituketa.

eat-past comp told

"Virginia told the teacher that Quinn ate this apple."

b. Clausemate Condition violation

*[Virginia-ga e_1 [Quinn-ga e_2 tabe-ta to]

V-nom Q-nom eat-past comp told

iituketa no]-wa **sensee-ni$_1$** **kono ringo-o$_2$** da.

comp-top teacher-dat this apple-acc cop

"(Lit.) it is the teacher, this apple that Virginia told that Quinn ate."

c. Satisfying the clausemate condition

[Virginia-ga sensee-ni [e_1 e_2 tabe-ta to]

V-nom teacher-dat eat-past comp

iituketa no]-wa **Quinn-ga$_1$** **kono ringo-o$_2$** da.

told comp-top Q-nom this apple-acc cop

"(Lit.) It is Quinn, this apple that Virginia told the teacher that ate."

Properties of No-da Constructions and Sluicing Constructions

Now, let us turn to the other two focus constructions and see the similarities and differences among these three constructions.

Hiraiwa and Ishihara point out that *no-da* constructions show significant syntactic parallelism with cleft constructions. Other than the island sensitivity and the clausemate condition, the *no-da* construction shows all the properties of cleft constructions: It allows multiple foci, it resists substitution of the complementizer with an NP, and nominative-genitive conversion is not allowed. However, the focused phrase can be inside an island, and if there are two focused phrases, they are not necessarily clausemates. These two properties presumably derive from constraints on overt movement, and focused phrases in the *no-da* construction do not seem to move

overtly. Given that the *no-da* construction does not contain any word order permutations, it is thus rather natural that it should differ from the cleft constructions on exactly these two properties.

(124) a. Multiple Foci

[QUINN-ga KONO RINGO-o tabe-ta no] da

Q-nom this apple-acc eat-past comp cop

focus *focus*

"(Lig.) It is Quinn, this apple that (he) ate."

b. NP-substitution

[QUINN-ga kono ringo-o tabe-ta

Q-nom this appke-acc eat-past

no/*mono/*kudamono] da.

comp/thing/fruit cop

"It is Quinn who ate this apple."

c. Nominative-Genitive Conversion

[QUINN-ga/*no kono ringo-o tabe-ta no] da

Q-nom/*gen this apple-acc eat-past comp cop

"(Lig.) It is Quinn, this apple that (he) ate."

d. Island

[Quinn-ga [[e_1 KONO RONBUN-o kai-ta]

Q-nom this paper-acc write-past

$hito_1$]-o hihansi-ta no] da.

person-acc criticize-past comp cop

e. Clausemate Condition

[Virginia-ga SENSEE-ni [Quinn-ga KONO RINGO-o
V-nom teacher-dat Q-nom this apple-ac
tabe-ta to] iituke-ta no] da.
eat-past comp told comp cop
"Virginia told the teacher that Quinn ate this apple."

Next let us look at sluicing. Hiraiwa and Ishihara show that sluicing and pseudosluicing in which the remnant NP does not bear any case markers or post-positions[7] (Fukaya and Hoji 1999), show exactly the same type of contrast that can be seen between clefting and pseudoclefting, except for the lack of a clausemate condition. Although I will not go into any details about the clausemate condition, I would like to note that they suggest that the clausemate condition does not hold in sentences with wh-phrases, and that sluicing is crucially a focus construction with wh-phrases. Thus there may be independent reasons for the lack of the clausemate condition, (for more details, see Hiraiwa and Ishihara).

Putting the matter of the clausemate condition aside, according to Hiraiwa and Ishihara the only difference between cleft and sluicing is the application of ellipsis and ellipsis-specific properties (if any). The relevant cases are the following: Sluicing

[7] Fukaya and Hoji (1999) use different terms for these constructions. They call the former case-marked sluicing and the latter non-case-marked sluicing. Here again, I follow Hiraiwa and Ishihara's (2002) terminology. Fukaya and Hoji show varieties of differences between these two constructions. For more details, see Fukaya and Hoji and also Fukaya (2002).

allows multiple foci but pseudosluicing does not; sluicing shows island effects but pseudosluicing does not. The examples below (125)-(126) show these points.

(125) Multiple foci

Quinn-ga dareka-ni nanika-o age-ta rasii ga…
Q-nom someone-dat something-acc give-past seem but…
"It seems that Quinn gave something to someone but"

a. Sluicing

boku-wa [dare-ni nani-o] (da) ka wakara-nai.
I-top who-dat what-acc cop Q know-not.
"I don't know what he gave to whom."

b. Pseudosluicing

*boku-wa [dare-ø nani-ø] (da) ka wakara-nai.
I-top who-dat what-acc cop Q know-not.
"I don't know who, what."

(126) Island sensitivity[8]

Quinn-wa [[otooto-ni nanika-o okuttekita]hito]-o

Q-top brother-dat something-acc sent person-acc

syootai-sita rasii ga…

invite-past seem but

"It seems that Quinn invited a person who had sent something to his brother, but…"

a. Sluicing

*boku-wa [nani-o (da) ka] sira-nai

I-top what-acc cop Q know-not

"I don't know what"

b. Pseudosluicing

boku-wa [nani-ø (da) ka] sira-nai.

I-top what cop Q know-not

"I don't know what."

[8] Extraction out of a complement clause does not show the degradation.

(i) Taroo-wa Hanako-ga otooto-ni nanika-o okuttekita-to

T-top H-nom brother-dat something-acc sent-comp

itta-ga, boku-wa nani-o da-ka sira-nai.

said-but, I-top what-acc cop-Q know-not.

"Taroo said that Hanako sent something to his brother, but he does not know what."

The example relevant to the clausemate condition is the following: In (127) the remnants of sluicing are clearly not clausemates, and still the sentence is perfectly acceptable.

(127) Virginia-ga dareka-ni [Quinn-ga nanika-o
H-nom someone-dat T-nom something-acc
tabe-ta to] iituke-ta rasii ga, boku-wa **dare-ni**
eat-past comp tell-past seem but I-top who-dat
nani-o (da) ka
what-acc cop Q
wakara-nai.
know-now.
"It seems that Virginia told someone that Quinn ate something, but I don't know to whom what."

Finally, I would like to note that Hiraiwa and Ishihara report the example in (128a) as an example that shows a ban on NP-substitution, and they argue that this also is a similarity between cleft and sluicing. According to their judgment, the example is not good. However, as we have discussed earlier, this type of sentence is an example of JPS, and I have used the same type of example to show Case connectivity in JPS. To my ear, this example does not sound bad, and according to the native speakers of Japanese who I have interviewed, this example is as good as its pseudocleft counterpart. Given my earlier argument, and the judgment of this sentence, this suggests that the pronoun in this example corresponds to the elided

clause itself rather than the complementizer. Note though that, although I disagree with their treatment of this particular pair of examples, this does not diminish the validity of their broader point. The many other similarities and differences of sluicing and cleft must still be explained.

(128) Virginia-ga nanika-o kat-ta rasii ga…
V-nom something-acc buy-past seem but
"It seems that Virinia bought something, but…"

a. Sluicing
boku-wa [sore-ga nani-o (da) ka] wakara-nai.
I-top it-nom what-acc cop Q know-not
"I don't know what it is."

b. Pseudosluicing
boku-wa [sore-ga nani (da) ka] wakara-nai.
I-top it-nom what-ø cop Q know-not
"I don't know what it was."

Japanese Focus Constructions Under a Split-CP analysis

So far we have seen that three types of focus constructions in Japanese show significant similarities. To capture these similarities Hiraiwa and Ishihara offer an analysis incorporating Rizzi's split-CP hypothesis. Their argument can be summarized as follows: Clefting and sluicing are derived from the underlying *no-da* construction through movement of focalized phrases followed by topicalization or

ellipsis of the clause headed by *no*. Thus, the three constructions show similarity because they share the same base structure. On the other hand, their differences are derived from the operation responsible for deriving each surface form. If ellipsis takes place, properties related to ellipsis show up; if movement takes place, the properties related to movement show up. Let's see their analysis.

First, they assume Split CP system. According to their analysis, sentence-final particles are hierarchically distributed to the heads of three CP-projections: FinP/CP, FocP and TopP. (129) is the schematic representation of the split-CP system that they adopt.

(129) $[_{TopP}$ XP-wa …$[_{FocP}$ … $[_{FinP}$ … $[_{IP}$ … $[_{VP}$ …V]-tense]-no]-da]]

Under this split-CP system, the derivation of each construction can be illustrated as in (130) for clefts and (131) for sluicing.

(130) The derivation of clefts

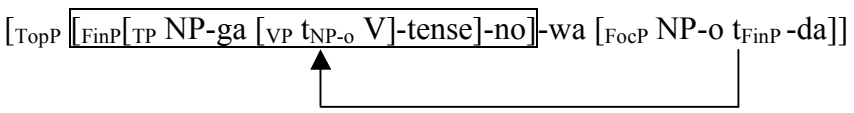
a. Step 0

$[_{TopP}[_{FocP}[_{FinP}[_{TP}$ NP-ga $[_{VP}$ NP-o V]-tense]-no]-da]]

b. Step 1: Focus movement

$[_{TopP}[_{FocP}$ NP-o $[_{FinP}[_{TP}$ NP-ga $[_{VP}$ $t_{NP\text{-}o}$ V]-tense]-no]-da]]

c. Step 2: Topic movement of FinP

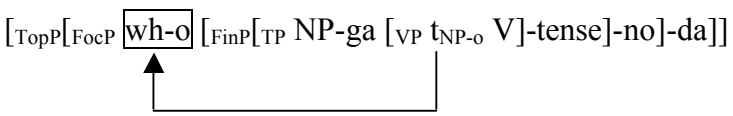


(131) The derivation of sluicing

a. Step 0

[TopP[FocP[FinP[TP NP-ga [VP wh-o V]-tense]-no]-da]]

b. Step 1: Focus movement



c. Step 2: Ellipsis of FinP

[TopP[FocP NP-o ~~[FinP[TP NP-ga [VP tNP-o V]-tense]-no]~~-da]]] -ka

As can be seen clearly in each derivation, the crucial step for driving clefting or sluicing is focus movement. From these derivations, it is obvious why clefting and sluicing are sensitive to islands but not the *no-da* construction: There is overt focus movement in the two former constructions but not in the *no-da* construction. An indication of the topicalization of FocP in cleft construction is the resulting word order, as well as the topic marker on FocP. Note that if we do not adopt this analysis, it is very difficult to capture the similarity and differences among these constructions.

A Modification: The Treatment of Pronominal Sluicing

Here I would like to argue that Hiraiwa and Ishihara's analysis of sluicing requires some modification.

In the earlier section, I showed that JPS is best analyzed as the ellipsis of the presuppositional clause of cleft construction. Based on the observation that JPS and the cleft construction show similarities, it is natural to extend the analysis of clefting to JPS. Under the split-CP analysis that we have just outlined, the derivation of JPS is something like (132).

(132) The Derivation of JPS

a. Step 0

$[_{TopP}[_{FocP}[_{FinP}[_{TP}$ NP-ga $[_{VP}$ wh-o V]-tense]-no]-da]]

b. Step 1: Focus movement

$[_{TopP}[_{FocP}$ wh-o $[_{FinP}[_{TP}$ NP-ga $[_{VP}$ $t_{NP\text{-}o}$ V]-tense]-no]-da]]

c. Step 2: Topic movement of FinP

$[_{TopP}$ $[_{FinP}[_{TP}$ NP-ga $[_{VP}$ $t_{NP\text{-}o}$ V]-tense]-no]-wa $[_{FocP}$ wh-o t_{FinP} -da]]

d. Step 3: Ellipsis and Pronominalization[9] of FinP

$[_{TopP}$ ~~$[_{FinP}[_{IP}$ NP-ga $[_{VP}$ $t_{NP\ o}$ V] tense] no] wa~~ $[_{FocP}$ wh-o t_{FinP} -da]]

sore

Unlike the derivation of sluicing that we have seen before, in this derivation, FinP undergoes topic movement before the ellipsis takes place. The necessity of this step is suggested by the observation that the *sore* or the presuppositional clause of the cleft construction cannot stay in-situ.

[9] Nakao and Yoshida assume a different derivation from the one I review here. For them a cleft is formed by a base generated focus element and null operator movement inside the presuppositional clause. They argue that *sore* is derived from the "specific presuppositional clause" with the genitive demonstrative *so-no*, with the structure of $[_{DP}$ so-no $[_{CP}$ …]]. Adopting the null operator analysis, they can sidestep the problem of specificity island violations by the focused phrase. Furthermore, they have a principled account of pronominalization. According to them, if a CP deletion takes place, *so-no* is stranded. However, in Japanese, since the genitive demonstrative cannot stand alone, it must be converted to the form of *sore* for language specific morphophonological reasons.

If we assume that JPS is derived from the cleft construction, and if we adopt the split-CP analysis, Nakao and Yoshida's approach raises a problem: Focus movement is not constrained by the specificity island. If, on the other hand, we discard the analysis that *sore* is derived from the specific presuppositional clause, then we lose a principled account of pronominalization. In this study, however, I adopt the split-CP analysis, which has the potential to account for not only the nature of focus constructions, but also that of nominal conditionals, and leave this problem open. I simply stipulate that FinP can be converted to *sore*. For more details of this theory of pronominalization, see Nakao and Yoshida (2006).

(133) Quinn-ga nanika-o tabe-ta rasii ga…

Q-nom something-acc eat-past seem but

"It seems that Quinn ate something, but…"

a. Wh-Cleft

Boku-wa [Quinn-ga tabe-ta no]-ga/wa

I-top Q-nom eat-past comp-nom/top

nani-o (da) ka

what-acc cop Q

wakara-nai.

know-not

"I don't know what it is that Quinn ate."

b. Wh-Cleft with scrambling

*Boku-wa nani-o [Quinn-ga tabe-ta no]-ga/wa

I-top what-acc Q-nom eat-past comp-nom/top

(da) ka wakara-nai.

cop Q know-not

"I don't know what it is that Quinn ate."

c. JPS

Boku-wa sore-ga/wa nani-o (da) ka wakara-nai.

I-top it-nom/top what-acc cop Q know-not

"I don't know what it is."

d. JPS with scrambling

*Boku-wa nani-o sore(-ga/wa) (da)ka wakara-nai.
I-top what-acc it(-nom/wa) cop Q know-not
"I don't know what it is."

Now the question is whether the step of topic movement before ellipsis is specifically necessary for JPS or is a more general requirement. There is no obvious reason to assume that JPS and sluicing are derived differently, except for the necessity of pronominalization for JPS. Thus, I would like to propose that the ellipsis of the FinP is only possible if it undergoes topic movement. I therefore assume that even in the standard sluicing construction, FinP is topic-moved before being elided.

3.3.3.7. The Analysis of Nominal Conditionals

The stage is now set for our analysis of nominal conditionals in Japanese, after our necessarily long and complicated review of Rizzi and Hiraiwa and Ishihara. What I will show here is that exactly the same properties of clefts that Hiraiwa and Ishihara try to derive hold true for nominal conditionals, and that the split-CP system is the best tool to capture the properties of nominal conditionals.

3.3.3.7.1. Further Properties of Nominal Conditionals

There are five properties of clefts that Hiraiwa and Ishihara try to derive from their analysis. They are summarized in (134).

(134) a. The possibility of multiple foci

b. Island effects

c. The complementizer *no* cannot be substituted with a pronoun/NP

d. The unavailability of Nominative/Genitive Conversion (NGC)

e. Clausemate condition on multiple foci

In this subsection, I show that the same contrast seen between clefts and pseudoclefts can be seen in the underlying form of nominal conditionals, i.e., there is a bifurcation between case-marked and non-case-marked conditionals.

In conditional constructions, the standard conditional in (135a) corresponds to the *no-da* construction, in which no word order permutation occurs. From this sentence, we can make a conditional sentence containing a cleft-like focus construction. Let us call it the cleft conditional. We can understand this construction as the conditional version of the cleft construction. As the example (135b) indicates, a focused phrase without a Case marker is also a possible option. Let us call this the pseudocleft conditional.

(135) a. mosi [[[[Quinn-ga ringo-o Virginia-ni age-ta
mosi Q-nom apple-acc V-dat give-past
no] de-are/nara] ba]…
comp cop cond
"If Quinn gave apple to Virginia…"

b. mosi [[[[Quinn-ga Virginia-ni age-ta
mosi Q-nom V-dat give-past
no]-ga ringo(-o) (mit-tu) de-are/nara] ba]…
comp-nom apple-acc three-cl cop cond
"If it is (three apples) that Quinn gave to Virginia…"

Now, I will show that this cleft conditional shows all the signature properties of clefts in (134). First, the cleft conditional allows multiple foci. On the other hand, the pseudocleft conditional does not allow multiple foci.

(136) a. The Cleft Conditional
mosi [Quinn-ga e_1 e_2 age-ta no]-ga
mosi Q-nom give-past comp-nom
Virginia-ni_1 ringo-o_2 (mit-tu) nara-ba.
V-dat apple-acc three-cl cop cond
"(Lit.) If it is to Virginia, (three) apples that Quinn gave…"

b. The Pseudocleft Conditional
mosi [Quinn-ga e_1 e_2 age-ta no]-ga
mosi T-nom give-past comp-nom
Virginia*(-$ni)_1$ ringo*(-$o)_2$ (mit-tu) nara-ba.
V-dat apple*-acc three-cl cop cond
"(Lit.) If it is to Virginia, (three) apples that Quinn gave…"

Second, the cleft conditional is sensitive to island constraints. Thus, in (137) the focused phrase is extracted from a complex NP island, and the sentence is severely degraded. In the pseudocleft conditional, on the other hand, extraction out of a complex NP is possible.

(137) a. mosi [Quinn-ga [$_{RC}$[Virginia-ga e_1 ringo-o
mosi Q-nom V-nom apple-acc
age-ta] hito$_1$]-ni at-ta nara ba]
give-past person-dat meet-past cop cond
"If Quinn meets the man who Virginia gave the apple to…"

b. *mosi [[Quinn-ga [$_{RC}$[Virginia-ga e_1 e_2 age-ta]
mosi Q-nom V-nom give-past
hito$_1$]-ni at-ta] -no] -ga ringo-o$_2$
person-dat meet-past-comp-nom apple-acc
(mit-tu) nara-ba…
three-cl cop-cond
"(Lit.) If it is (three) apples that Quinn met the man who Virginia gave to…"

c. mosi [[Quinn-ga [$_{RC}$[Virginia-ga e_1 e_2 age-ta]
mosi Q-nom V-nom give-past
$hito_1$]-ni at-ta] -no] -ga $ringo_2$ (mit-tu)
person-dat meet-past-comp-nom apple-ø three-cl
cop-cond
nara-ba…

"(Lit.) If it is (three) apples that Quinn met the man who Virginia gave to…"

Third, the complementizer *no* cannot be substituted with an NP in the cleft conditional, but it is possible in the pseudocleft conditional.

(138) a. Cleft Conditional
mosi [Quinn-ga tabe-ta no/*mono/*kudamono]-ga
mosi Q-nom eat-past comp/thing/fruit-nom
ringo-o (mit-tu) nara-ba…
apple-acc three-cl cop-cond…
"If it was (three) apples that Quinn ate…"

b. Pseudocleft Conditional
mosi [Quinn-ga tabe-ta no/mono/kudamono]-ga
mosi Q-nom eat-past comp/thing/fruit-nom
ringo (mit-tu) nara ba…
apple-ø three-cl cop cond…
"If it was (three) apples that Quinn ate…"

The fourth property is the possibility of Nominative/Genitive Conversion. The cleft conditional, like its cleft counterparts, resists nominative/genitive conversion, in contrast to the pseudocleft conditional.

(139) a. Cleft Conditional

mosi [Quinn-ga/*no tabe-ta no]-ga ringo-o
mosi Q-nom/gen eat-past comp-nom apple-acc
(mit-tu) nara-ba…
three-cl cop-cond
"If it was (three) apples that Quinn ate…"

b. Pseudoleft Conditional

mosi [Quinn-ga/no tabe-ta no]-ga ringo
mosi Q-nom/gen eat-past comp-nom apple-ø
(mit-tu) nara-ba…
three-cl cop-cond
"If it was (three) apples that Quinn ate…"

Finally, the cleft conditional with multiple foci is sensitive to the clausemate condition.

(140) a. mosi [Virginia-ga sensee-ni [Quinn-ga ringo-o
mosi V-nom teacher-dat Q-nom apple-acc
tabe-ta-to] iituke-ta nara-ba]
eat-past-comp tell-past cop-cond
"If Vrirginia tells the teacher that Quinn ate the apple…"

b. *mosi [Viriginia-ga e_1 [Quinn-ga e_2 tabe-ta-t
mosi V-nom Q-nom eat-past-comp
iituke-ta no]-ga sensee-ni_1 rigo-o_2
tell-past comp-nom teacher-dat apple-acc
(mit-tu) nara-ba
three-cl cop-cond
"(Lit.) If it was to the teacher, three apple that Virginia told that Quinn ate… "

Looking at these five properties, it is clear that cleft and the cleft conditional are quite similar creatures.

Now let us turn to the nominal conditional. Our expectation here is that the nominal conditional corresponds to sluicing in Hiraiwa and Ishihara's paradigm. To test this point, what I have to show is whether the contrast between sluicing and pseudosluicing can be seen between the nominal conditional with a remnant bearing case-markers and the one with a remnant without case-markers. What Hiraiwa and Ishihara show was the following two properties: sluicing allows multiple foci but pseudosluicing does not; sluicing shows island effects but pseudosluicing does not. Nominal conditionals, indeed exhibit exactly the same contrast as sluicing and

pseudosluicing, as the following examples show. Example (141) shows that the case-marked nominal conditional allows multiple foci but pseudocleft conditionals do not.

(141) Multiple Foci

Quinn-ga dareka-ni nanika-o age-ta rasii ga…
Q-nom someone-dat something-acc give-past seem but
"It seems that Quinn gave something to someone, but"

a. Case Marked Nominal Conditional

(mosi sore-ga) Virginia-ni ringo-o (mit-tu)
mosi it-nom V-dat apple-acc three-cl
nara-ba…
cop-cond
"(Lit.) if it was (three) apples to Virginia …"

b. Non Case Marked Nominal Conditional

(mosi sore-ga) Virginia ringo (mit-tu)
mosi it-nom V-ø apple-ø three-cl
nara-ba…
cop-cond
"(Lit.) if it was (three) apples to Virginia …"

Example (142) shows that the case-marked nominal conditional is sensitive to the complex NP island, but the non-case-marked nominal conditional is not. This, in turn, suggests that movement is not involved in the non-case-marked nominal conditional.

(142) Island Sensitivity

Quinn-wa [[otooto-ni nanika-o okuttakita]hito]-o
Q-nom brother-dat something-acc send person-acc
syootai-sita rasii ga…
invite-past seem but…

"It seems that Quinn invited a person who had sent something to his brother, but…"

a. Case Marked Nominal Conditionals

*(mosi sore-ga) ringo-o (mit-tu) nara-ba
mosi it-nom apple-acc three-cl cop-cond

"If it was (three) apples…"

b. Non Case Marked Nominal Conditionals

(mosi sore-ga) ringo-o (mit-tu) nara-ba
mosi it-nom apple-ø three-cl cop-cond

"If it was (three) apples…"

Summarizing the discussion so far, the conditional clauses reviewed in this subsection show significant parallelism with the *no-da*, cleft, and sluicing constructions that Hiraiwa and Ishihara examined. Furthermore, we have seen that JPS and the pronominal nominal conditional also show significant similarities. From these observations, it is quite plausible to conclude that they are basically the same type of construction, and thus should be analyzed in the same fashion.

3.3.3.7.2. The Derivation of Nominal Conditionals

The basic line of analysis of the nominal conditional that I will outline from this point is the same as the analysis of sluicing that is adopted by Hiraiwa and Ishihara. However, I have not yet described the internal makeup of conditional clauses under the split-CP system. Let us start from this point.

From the discussion so far, it is easy to see that many parts of the complex verbal morphology of conditional clauses are shared by the *no-da* construction, even though some of them have different surface forms. Simply put, the right edge of conditional clauses is the same as that of cleft constructions, except for the following two points: first, the top nodes of these constructions are occupied by different elements; second, the morphemes that compose conditional morphology can be omitted relatively freely. Compared with the *no-da* construction, the structure of conditionals can be illustrated as in (143b). In the embedded context, the *no-da* construction must be followed by complementizers like *-to* if it is declarative or *-ka* if it is interrogative. On the other hand, in conditionals, the rightmost element is *-ba* and it seems that this *-ba* marks the clause type of a conditional, even though it can be omitted.

(143) a. *No-da* Constructions

$[_{ForceP}[_{TopP}[_{FocP}[_{FinP}[_{IP}[_{VP}\ldots V]\text{-tense}]\text{-no}]\text{-da}]]\text{-to/ka}]$

b. Conditional

$[_{ForceP}[_{TopP}[_{FocP}[_{FinP}[_{IP}[_{VP}\ldots V]\text{-tense}](\text{-no})]\text{-(na)-ra/de-are}]](\text{-ba})]$

Based on the optionality of some of the morphemes, conditional verbs can have seven possible surface forms. (144) shows all seven possibilities with the verb *taberu* (eat). These examples show how the omission of conditional morphemes can take place. The most radical pattern is (144g) where all the morphemes are omitted except the force marker -*ba*.

(144) a. *tabe-ta*-no-de-are-ba

b. *tabe-ta*-ra-ba

c. *tabe-ta*-ra

d. *tabe-ta*-no-nara-ba

e. *tabe-ta*-nara-ba

f. *tabe-ta*-nara

g. *taber-e*-ba

A note on the traditional treatment of conditional clauses is in order. In the paradigm in (144) we can see three of the four patterns of conditional forms that are traditionally treated as different expressions: *re-ba* conditionals, *ta-ra* conditionals, *na-ra* conditionals respectively. Although there are some subtle semantic or pragmatic differences among these three forms of conditionals, in Japanese traditional grammar these are arbitrary distinctions. Under the traditional treatment, we cannot capture the intuition that they can all have the same basic meaning and, above all, a pattern like (144a) where all the morphemes are expressed cannot be captured.

With the structure in (143b) in mind, let us see the derivation of nominal conditionals. As in the derivation of sluicing, the nominal conditional results from the

base sentence through the cleft conditional structure and ellipsis of FinP. (145) illustrates the stepwise derivation.

(145) a. Step 0

$[_{ForceP}[_{TopP}[_{FocP}[_{FinP}[_{IP}$ NP-ga $[_{VP}$ NP-o V]-tense]-no]-nara]]-ba]

b. Step 1: Focus Movement

$[_{ForceP}[_{TopP}[_{FocP}$ NP-o $[_{FinP}[_{IP}$ NP-ga $[_{VP}$ $t_{NP\text{-}o}$ V]-tense]-no]-nara]]-ba]

c. Step 2: Topic Movement of FinP

$[_{ForceP}[_{TopP}$ $[_{FinP}[_{IP}$ NP-ga $[_{VP}$ $t_{NP\text{-}o}$ V]-tense]-no]-ga $[_{FocP}$ NP-o t_{FinP} -nara]]-ba]

d. Step 3: Ellipsis/Pronominalization of FinP

$[_{ForceP}[_{TopP}$ ~~$[_{FinP}[_{IP}$ NP-ga $[_{VP}$ $t_{NP\text{-}o}$ V]-tense]-no]-ga~~ $[_{FocP}$ NP-o t_{FinP} -nara]]-ba]

sore-ga

Under this analysis, we can successfully capture the similarities between conditional constructions and focus constructions. As in the case of focus constructions, the standard conditional, the cleft conditional and the nominal conditional show similarities because they are derived from the same underlying structure. Furthermore, we can derive the parallelisms between JPS and nominal conditionals with *sore*. Both of them are residues of FinP ellipsis. The crucial

differences between focus constructions and conditional constructions are their surface verbal morphology, and Force^0 elements.

3.3.3.7.3. The Licensing Condition on Mosi

The discussion up to this point has made clear the nature of *sore* in the nominal condition. In our analysis *sore* replaces FinP. In the earlier discussion, I have shown that the CA *mosi* in the nominal conditional is legitimate only when *sore* is present. Now we have a better understanding of the licensing condition of *mosi*. Under the proposed analysis *mosi* is licensed by FinP. If we assume the notion of *Criterion* adopted by Rizzi, *mosi* is licensed by Fin^0 through a Spec-Head relation.

3.3.4. Conclusion

In this section, I have revealed two important aspects of conditional clauses in Japanese. One is the internal syntax of the clause, and the other is the distributional properties of *mosi*. What we have specifically seen is that there is a tight connection between Japanese focus constructions and conditional constructions, and this connection gives us a clue to understanding the distributional properties of *mosi*, i.e., *mosi* is licensed by FinP.

In the following sections, I will try to show why in the Japanese conditional clause an adjunct clause is not a strong island but rather is a weak island. The discussion here will give us an important insight into understanding the weak-islandhood of conditional clauses. Below, I will show that *mosi* acts like a focus marker. Based on this observation, I will argue that *mosi* creates a so-called

intervention effect. To capture this intervention effect, I will argue that *mosi* is generated in the specifier of FinP and induces a Relativized Minimality effect (Rizzi 1990).

3.4. The Question of *Mosi* and Weak Island Effects

In the earlier discussion, we have seen that the Conditional Adverb (CA) *mosi* is an inducer of Weak Island (WI) effects. I now turn to the question of why *mosi* induces WI effects. Answering this question will lead us to a better understanding of *mosi* and CAs in general.

The claim that I will establish in the course of this discussion is that *mosi* is a conditional scope marker.

3.4.1. Weak Island Inducers

To answer the question, it is worth thinking about what kind of elements induce WIs, and what properties they share so that we can test whether *mosi* is one of these elements or it is a totally different element. In previous studies on WIs, there is consensus among researchers that certain quantificational elements tend to induce WIs (Beck 1996; Hagstrom 1998; Harada 1972; Harada 1973b; Honcoop 1998; Hornstein and Uriagereka 2002; Kim 2002; Miyagawa 2002; Pesetsky 2000; Rizzi 1990; Ross 1984; Rullmann 1995; Tanaka 2003a among many others). The WI effects induced by these quantificational elements are sometimes called intervention effects (Beck 1996 among many others). As the name suggests, the basic pattern of the WIs induced by quantificational elements can be illustrated as that a harmful

quantifier intervenes between the links of an operator-variable chain. There are various analyses proposed in the literature, but basically they all try to capture this basic pattern. As an example of the intervention effects, it has been widely observed that negation induces an island effect (Beck 1996; Harada 1972; Linebarger 1987; Rizzi 1990; Ross 1984; Rullmann 1995). The examples in (146) are from English, showing that negation blocks the overt movement of non-referential wh-elements. (147) summarizes examples from German, showing that negation blocks covert movement. (147b) suggests that just containing a negation does not make a sentence bad. Rather, if the in-situ wh-phrase is c-commanded by the negation, the sentence becomes bad. Importantly the two sentences in (147) have basically the same meaning.

(146) a. Who do**n't** you think Quinn saw?

b. *How do**n't** you think Quinn behaved?

(147) a. Wer hat wo **niemanden** agetroffen?
who has where **nobody** met
"Who didn't meet anybody where?"

b. *Wer hat **niemanden** wo angetroffen?
Who has **nobody** where met
"Who didn't meet anybody where?"

As we have seen in the earlier discussion, intervention effects have also been observed in Japanese (Harada 1972; Hoji 1985; Miyagawa 2002; Takahashi 1990;

Tanaka 2003b). Miyagawa, for example, observes that, like in other languages, scope-bearing elements block movement. Aside from the examples that we have seen already, universal quantifiers and existential quantifiers, as well as a certain class of focus particles, are known to induce intervention effects. Let us see some of these examples. The examples in (148) show that universal quantifiers in Japanese block covert wh-movement.

(148) a. ??Dono hito-mo nani-o kat-ta-no?
every person-foc what-acc buy-past-Q
"What did every person buy?"

b. ?*Hotondo dono hito-mo nani-o kat-ta-no?
almost every person-foc what-acc buy-past-Q
"What did almost every person buy?"

c. Quinn-ga nani-o kat-ta-no?
Q-nom what-acc buy-past-Q
"What did Quinn buy?"

Having seen these examples, we can conclude that Quantificational Elements are an inducer of intervention effects. In addition, based upon observations of a wide range of intervention/WI effects in Japanese, Miyagawa (2002) concludes that all WIs should be analyzed as being induced by Quantifiers. In other words, WIs in Japanese are Quantifier Induced Barriers (QIBs). This position is, however, controversial given the current debates on WIs among researchers (Szabolcsi and den Dikken 1999

Tomioka 2005).[10] However, it seems to be true that WI inducers are always quantificational elements. Thus, I will basically assume that WIs are induced by quantificational elements.

3.4.2. Mosi as a Scope Marker

Given the above discussion, the question naturally arises whether the WI inducer *mosi* is also a quantificational element or not. In the following discussion, I will show that there are some reasons to believe that *mosi* is a quantificational element.

3.4.2.1.The Basic Observations

One of the prominent properties of *mosi* is its freedom of location inside a conditional clause. For example, in an example like (149), *mosi* can appear at any of the underlined positions.

(149) Quinn-wa [___Virginia-ga___Stillman-ni___tegami-o
Q-top V-nom S-dat letter
____ dasita-ra] naki-dasu daroo.
send-cond cry-start will
"Quinn will start crying if Virginia send a letter to Stillman."

[10] Tomioka, for example, suggests that WIs should not be explained purely by syntax. He argues for a semantic/pragmatic approach to WIs. Szabolcsi also argues for a semantic account of WIs.

It is also worth noting that *mosi* cannot be licensed outside of a conditional clause. For instance, in (150) the conditional clause is fronted to the beginning of the sentence, and *mosi* is left outside of it. A *mosi* that is stranded out of a conditional clause normally creates a wild unacceptability. Thus, *mosi*'s distribution is free but not totally free. It is free inside a conditional clause. As we have seen in the earlier discussion, *mosi* must be licensed by a conditional clause, and thus must be inside the licensing conditional clause.

(150) *[Virginia-ga Stillman-ni tegami-o dasita-ra]
V-nom S-dat letter-acc send-cond
Quinn-wa *mosi* naki-dasu daroo.
Q-nom *mosi* cry-star will
Intended interpretation: "Quinn will start crying if Virginia sends a letter to Stillman."

Note also that *mosi* must be licensed by a conditional clause, but must be licensed by a clause-mate conditional marker. Let's take a look at the following pair of sentences. (151b) is severely degraded compared to (151a). In (151b) *mosi* is not in the same clause as the conditional marker *-ra*. On the other hand, in (151a), they are in the same clause.

(151) a. Quinn-wa [Virginia-ga sono-ringo-o *mosi*
Q-top V-nom that-apple-acc mosi
tabeta-**ra**] okori-dasu daroo.
eat-cond get-angry-start will
"Quinn will get angry if Virginia eats that apple."

b. *Quinn-wa [Stillman-ga [Virginia-ga
Q-top S-nom V-nom
sono-ringo-o *mosi* tabeta-to] itta-**ra**]
that-apple-acc mosi eat-comp say-cond
okori-dasu daroo.
get-angry-start will
"Quinn will get angry if Stillman says that Virginia ate that apple."

This contrast, thus, suggests that *mosi* must be licensed by a clause-mate conditional.[11]

[11] The clause-mate condition, however, can be relaxed in certain environments. For example, when so-called bridge expressions (Kato 1985) intervene between *mosi* and the conditional complementizer, the clause-mate condition can be violated as illustrated in (i).

(i) Virginia-ga sono hon-o *mosi* yonda-koto-ga are-**ba…**
V-nom that book-acc mosi read-fact-nom exist-cond
"If Virginia has an experience of reading that book…"

Interestingly bridge expressions can also mediate the relation between negation and negative polarity items in Japanese such as *-sika*, "only", which also is subject to the clause-mate condition in a normal situation (Kato 1985).

(ii) a. Quinn-wa ringo-*sika* tabe-**nakat**-ta.

There is another interesting property of *mosi* that is related to its relative freedom of positioning, namely, there are some differences in interpretation of elements that are under and outside the scope of *mosi*.[12] Let us take examples containing a numeral quantifier. All the examples in (152) share basically the same meaning. However, in the examples containing *mosi*, depending on *mosi*'s position, the interpretation is different. The difference is related to what is focused in the conditional clause. In (152b) both *gakusee* "student", the quantifier or action of eating can be focused as in (152a), which does not contain *mosi*. On the other hand, in (152c), in its most natural interpretation, *gakusee* is not focused, but the quantifier or the action can be focused. Thus (152c) can be paraphrased using clefts such as (152c.i) or (152c.ii). They are most naturally available but (152c.iii) is somewhat

---

Q-top apple-only eat-**not**-past

"It was only apples that Quinn ate."

b. *Quinn-wa [Virginia-ga ringo-*sika* tabe-***ta***-to]iwa-**naka**-ta.

Q-top V-nom apple-*only* eat-past-comp say-neg-past.

"Quinn said it was only apples that Virginia ate."

c. Taroo-wa [ringo-*sika* tabe-***ta***-koto]-ga nai.

T-top apple-only eat-past-fact-nom not

"It is only apples that Taroo has an experience of eating."

Note that *sika* is also an example of a focus marker. The similarity between *mosi* and *sika* can also support the claim that *mosi* is a focus marker. However, it is not at all clear in exactly which contexts the clause-mate condition can be relaxed. Tomohiro Fujii (p.c.) suggests that in environments that resemble raising or control constructions (so-called restructuring environment), the clause-mate condition can be relaxed. In this study, however, I will leave this point open.

[12] This was suggested by Hiromu Sakai (p.c.). I would like to thank him for directing my attention to this type of phenomenon.

degraded. A tricky part of this is that if one puts a heavy focus stress on *gakusee*, the interpretation (152c.iii) becomes readily available.[13] However, the most natural interpretations is (152c.i) or (152c.ii) depending on the focus intonation inside the scope of *mosi*. We can support this point by using negation. In (152b) we can negate any of the elements to the right of *mosi*. However, it is odd if someone says, "no it is teachers not students" after an utterance like (152c). This oddity does not arise in the case where someone negates the elements to the right of *mosi* as in "it is 5 not 3" or "it is if they leave not if they come". These interpretive differences suggest that elements on the right-hand side of *mosi* are readily counted as the member of alternative set in the sense of Rooth (1985). Note that this negation test also can be side stepped by manipulating the intonation. However, the important point for us is

[13] This might be related to a widely acknowledged phenomena, the expansion of the domain of focus (Chomsky 1972; Jackendoff 1972 among many others). For example, in English, if emphatic stress falls on the object, the sentence can be interpreted in several different ways. This is illustrated in the following example. In (i) the emphatic stress falls on the object NP *the car* but other elements than the object, including the whole sentence, can be focused.

(i) John fixed THE CAR

a. What did Quinn fix?

b. What did Quinn do?

c. What happened?

Taking this expansion of focus as a general phenomenon, it is not surprising that something on the left of *mosi* or the whole clause containing the focus marker can be interpreted as focused as well as the ones on the right. The same focus expansion can, indeed, be seen in examples containing other focus particles in Japanese (see Aoyagi 1998; Kuroda 1965 among others for detailed discussion on this point).

even though the effect is not so strong, still there are differences in the most natural interpretations depending on the position of *mosi*. Exactly the same pattern holds true for (152d). Here, only the verb is on the right of *mosi*, and thus only the action described by the verb can be focused.

(152) a. gakusee-ga 3-nin kita-ra …
student-nom 3-cl$_{(human)}$ come-cond
"If three students come…"

b. *mosi* gakusee-ga san-nin kita-ra
mosi student-nom 3-cl$_{(human)}$ come-cond

c. gakusee-ga *mosi* san-nin kita-ra
student-nom *mosi* 3-cl$_{(thing)}$ come-cond

i. "If it is 3 (students) (rather than 4 or 5) that comes… "

ii. "If it is coming of 3 (students) that studends do …"

iii. ?"If it is students that comes (three of them)"

d. ringo-o mit-tu *mosi* tabere-ba
apple-acc 3-cl$_{(thing)}$ *mosi* eat-cond

i. "If it is 3 (apples) (rather than 4 or 5) that you eat… "

ii. ?"If it is the eating of 3 (apples) that you do …"

iii. ?"If it is apples that you eat (three of them)"

Summarizing the discussion so far, we have seen that the position of *mosi* affects the interpretation of a conditional sentence. The basic pattern is that elements

on the right hand side of *mosi* can be easily focused, but those on the left hand side are not.

3.4.2.2. Disjunction Scope in English and Focus Particles in Japanese

Now the question is how we can analyze the behavior of *mosi*. To gain clues for the analysis of *mosi*, I will pick up an example from English that show similar behavior to *mosi*: the so-called disjunctive scope marker *either*.

Under the assumption that the conjunction or the disjunction is a scope-taking element (see Moltmann 1992; Munn 1993 among others for detailed discussion), there have been various studies on *either* in the literature (Larson, 1985; Schwartz 1999 among many others). Most previous studies on *either* have been devoted to capturing the correlation between its distribution and its interpretation. Let us summarize the most basic results of the previous studies.

The most basic distributional property of *either* can be summarized as the following: *either* can occur in a clause that contains the disjunction *or*. More precisely, *Either* must be in a tensed clause that contains *or* but it can go outside of an infinitival clause containing *or* (Han and Romero 2004; Larson 1985; Schwarz 1999). This distributional pattern is illustrated in (153) and (154). (153) illustrates the fact that *either* can be inside or outside of an infinitival clause containing disjunction. On the other hand, (154) clearly shows that it cannot be outside of the finite clause.

(153) a. Quinn pretended [to *either* be looking for a burglar **or** a thief].

b. Quinn *either* pretended [to be looking for a burglar **or** a thief].

(154) a. Quinn believes [that Stillman said [that Virginia was *either* drinking **or** playing video games]].

b. Quinn believes [that Stillman said [that *either* Virginia was drinking **or** playing video games]].

c. ??Quinn believes [that Stillman said *either* [that Virginia was drinking or playing video games]].

d. ??Quinn believes [that Stillman *either* said [that Virginia was drinking **or** playing video games]].

e. ***Either*** Quinn believes [that Stillman said [that Virginia was drinking **or** playing video games]].

Now let us turn to the interpretive aspects of *either*. Larson claims that *either* is a scope indicator for the disjunction. He cites the examples in (155) and (156). Larson points out that there are three possible interpretations regarding the disjunction in (155) (Larson 1985; Rooth and Partee 1982). One interpretation that is represented in (155a) is the so-called *de dicto* reading in which Mary is searching for a servant and would be satisfied with any individual x that meets the description "x is a maid or x is a cook". (Larson 1985: 218). Another reading of (155b) is the *de re* reading. Under the *de re* reading there is a particular individual who is either a maid or a cook such that Mary is seeking that individual (Larson 1985:218). The third reading in (155c) is the following: Mary is looking for an individual x satisfying the description "x is a maid" or else she is looking for any individual x satisfying the description "x is a cook" (Larson 1985:218).

(155) Virginia is looking for a maid or a cook.

a. Virginia is looking for ((a maid) or (a cook)).

b. for some x, a maid or a cook, Virginia is looking for x.

c. Virginia is looking for (a maid) or Virginia is looking for a cook.

Larson points out that (156a) has all three of the readings in (155). In (156b), on the other hand, the interpretation in (155a) is not available, i.e., *or* cannot take the wide scope. Also in examples (156c) and (156d), the narrow scope readings become unavailable.

(156) a. Virginia is looking for *either* a maid **or** a cook.

b. *Either* Virginia is looking for a maid **or** a cook.

c. Virginia is *either* looking for a maid **or** a cook.

d. Virginia *either* is looking for a maid **or** a cook.

The examples in (156) strongly suggest that the scope of *or* is affected by the position of *either*. If the *either* occurs adjacent to the disjunctive constituent, the narrow scope reading becomes available. However, if the *either* is displaced from *or*, the narrow scope interpretation becomes unavailable. Considering these interpretive aspects of *either* Larson concludes that it is a scope marker of disjunction.

The crucial similarities between *mosi* and *either* are: (i) there is a correlation between their positioning and their interpretation, i.e., both elements seem to indicate the upper-bound of the scope of the operator, either *or* or conditional; and (ii) they

have a local relation to their licensors, i.e., both of them are basically in a clause-mate relation to their licensors. Based on these similarities, it is plausible to understand that *mosi* has a similar function as *either*, i.e., marking the scope of its licensor. This, in turn, means that these elements are scope-bearing elements.

Having seen the discussion so far, the real question is how we can capture the distribution of the scope markers. At this point, however, I would like to just assume the following two points. First, these scope markers are adverbs resembling the English focus particle *even* (Hendriks 2003), and their distribution of *either* is regulated by whatever principle regulates the distribution of *even*. I would like to leave the actual mechanism open, and just pursue the descriptive generalization.

3.4.3. Intervention Effects

In the discussion so far, we have seen that *mosi* can be analyzed a scope-bearing element, more specifically the scope marker of conditional operator. Now, we have a clue for why *mosi* induces WI effects. The simplest answer is that *mosi* is a quantificational element that becomes a harmful intervener if it appears between the moved phrase and its underlying position. In other words, *mosi* induces intervention effects in exactly the same way as other scope-bearing elements.

4. Approaches to Adjunct (Non)Islands

Thus far, we have established that conditional clauses in Japanese are not islands. A closer look at other adjunct domains suggests this claim is not limited to conditional clauses. Rather, it is generally true for all adjunct clauses (for the related

discussion, see Saito & Fukui 1998; Ishii 1997; Mihara 1994 among others). Let us consider some examples. In the following, I present examples of *Because Clauses* and *After Clauses* as well as RCs, Coordinate clauses and Complement clauses.[14] These are to be understood as baseline examples.

(157) Because clauses

Quinn-wa [Virginia-ga dono-gakusee-ni present-o
Q-top V-nom which-student-dat present-acc
ageta-**node**/-**kara**] nakidasita-no?
gave-because/-because cried-Q?

Lit. "Which student does Quinn get angry because Virginia gave a present to?"

a. Scrambling

Dono-gakusee-ni$_1$ Quinn-wa [Virginia-ga t$_1$
which-student-dat Q-top V-nom
present-o ageta-**node**/-**kara**] nakidasita-no?
present-acc gave-because/-because cried-Q?

[14] Examples of coordination are cited here because relativization in Japanese is normally immune to island constraints except the Coordinate Structure Constraint, as we have seen in the discussion in section 2 of this chapter. Thus, examples of relativization in (1b) and (158b) should be compared with the example in (1b) and (1d).

b. Relativization

[$_{NP}$[$_{CP}$ Quinn-ga [Virginia-ga gap$_1$ present-o

Q-nom V-nom present-acc

ageta-**node/-kara**] nakidasita]gakusee$_1$]

gave-because/-because cried student

Lit. "The student who Quinn cried because Virginia gave a present to."

(158) Temporal Clause: *After Clause*

Quinn-wa [Virginia-ga dono-gakusee-ni present-o

Q-top V-nom which-student-dat present-acc

ageta **ato**] nakidasita-no?

gave **after** cry-Q

Lit. "Which student did Quinn cry after Virginia gave a present to?"

a. Scrambling out of After clause

Dono-gakusee-ni$_1$ Quinn-wa [Virginia-ga t$_1$

which-student-dat T-top H-nom

present-o ageta **ato**] nakidasita-no?

present-acc gave **after** cry-Q

b. Relativization out of After clause

[$_{NP}$[$_{CP}$ Quinn-ga[Virginia-ga t$_1$ present-o ageta

Q-nom V-nom present-acc gave

ato]nakidasita] gakusee$_1$]

after cry student

Lit. "The student who Quinn cried after Virginia gave a present to"

(159) a. Scrambling out of Relative Clauses

*Dono-gakusee-ni$_1$ Quinn-wa [$_{NP}$[$_{CP}$ Virginia-ga t$_1$

which-student-dat Q-top V-nom

t$_2$ ageta] present$_2$]-ga suki-nano?

gave present-nom like-Q

Lit. "Which student does Quinn like the present that Virginia gave to?"

b. Relativization out of Coordinations

*Quinn-ga [$_{NP}$[$_{CP}$[$_{IP}$ Virginia-ga [$_{VP}$[$_{VP}$ Stillman-ni

Q-nom V-nom S-dat

hon-o age] [$_{VP}$ t$_1$ [$_{VP}$ present-o age]]-ta]]

book-acc give present-acc give -past

gakusee$_1$]-ni atta.

student-dat met

Lit. "Quinn met the student who Virginia gave a book to Stillman and gave a present to."

c. Scrambling out of Complement Clauses

Stillman-ni$_1$ Quinn-wa [Virginia-ga t$_1$ present-o

S-dat Q-top V-nom present-acc

ageta-to] omotteiru.

gave-comp think

"Quinn thinks that Virginia gave a present to Stillman."

d. Relativization out of Complement Clauses

[$_{NP}$[$_{CP}$ Quinn-ga [$_{CP}$ Virginia-ga t_1 present-o

Q-nom V-nom present-acc

ageta-to] omotteiru] gakusee$_1$]

gave-comp think student

"The student that Quinn thinks that Virginia gave a present to"

The examples in (157) and (158) are as good as extraction out of complement clauses, as in (159c) and (159d). In particular, they sound much more acceptable when compared to extraction out of RCs or Coordination. Based on this comparison, we take it to be the case that extraction out of adjunct clauses is generally good in Japanese.

The aim of this section is to investigate why adjunct islands are not operative in Japanese. The first point I would like to make is that the non-islandhood of adjunct clauses is not a property specific to Japanese language. Instead, I show that at least two other languages (Korean and Malayalam) pattern like Japanese. Furthermore, I show that English and Spanish also exhibit adjunct non-islands in some limited environments (Etxepare 1999, and Hornstein 2001). Based on these facts, I discuss two possible approaches to adjunct islands. The first approach suggests the possibility that the 'islandhood' of adjunct clauses is dependent upon the choice of predicates in the matrix clause and the positions of the adjunct clauses in a sentence. Etxepare (1999) and Hornstein (2001) show that in Spanish and English, extraction out of preposed adjunct clauses is generally better than extraction out of post-verbal adjunct clauses. We discuss whether the same holds true in Japanese. The second approach is

a parametric approach to adjunct islands. We pay attention to the fact that all the three languages share a typological property, so-called wh-in-situ. We discuss a possibility that adjunct non-islands can be derived from this particular property of these languages. In the following discussion, the argument is mainly based on data from conditional clauses, but essentially the same holds true for other adjunct clauses.

4.1. Adjunct Islands and Left/Right Asymmetry

4.1.1. Spanish and English

Etxepare (1999) points out that wh-phrases can be extracted from conditional clauses in Spanish if the conditional clause is in the complement position of a reporting verb, such as *say*, and if the conditional clause is located immediately after the complementizer. Conversely, if a conditional clause is in the complement of a verb like *desire, interpret and mention*, wh-extraction is not possible. He cites the following examples from Spanish.

(160) a. Qué $libro_1$ dijiste [que [si Ricardo leia alguna vez] abandonaria la Linguistica]?
"Which book did you say that if Ricardo ever read he would abandon linguistics?"

b. *Qué libro quieres [que [si algun lee]] abandone la linguistica?
"Which book do you desire that if anyone read, he would abandon linguistics?"

Hornstein (2001) points out that the same pattern holds true in English. He cites the examples in (161). As evident in these examples, the possibility of extraction is affected by the type of matrix verb in the same way as in Spanish.

(161) a. Which book$_1$ did you say [that [if Quinn ever read t_1] he would abandon linguistics]?

b. *Which book$_1$ did you desire [that [if anyone read t_1] he would abandon linguistics]?

Hornstein further showed that if the conditional clause is located after the embedded verb, extraction becomes impossible, even if the conditional clause is inside the complement clause of reporting verbs.

(162) *Which book$_1$ did you say [that Quinn would abandon linguistics [if he ever read t_1]]?

Their accounts are based on the theory of sideward movement (Nunes 1995; Hornstein 1999, 2001). Etxepare and Hornstein argue that the pattern of wh-extraction is expected under their theory of sideward movement. In this study, however, let us put the technical details aside, and investigate the descriptive generalization they formulated. Informally, their findings can be summarized as the following. First, extraction is possible only if the matrix verb is a reporting verb, such as *say*. Second, extraction is possible only if a conditional clause is located higher

than the embedded subject. In their terms, if an *if*-clause is adjoined to the FP generated between the embedded CP and IP. Thus, as Hornstein (2001) points out, if the conditional clause is located at the post-verbal position, extraction is not possible; we call this second generalization the Left/Right Asymmetry.

4.1.2. Scrambling from Conditionals and the Left/Right Asymmetry

Let us now examine what the descriptive generalization can tell us about the adjunct non-islands in Japanese. The important point of the argument is if the two descriptive generalizations that Etxepare and Hornstein formulated are generally true cross-linguistically, we would observe the same pattern in languages other than Spanish or English. More specifically, based on their descriptive generalizations, we predict that preposed adjunct clauses are islands if they are in the complement clause of non-reporting verbs, such as *desire*, *mention* or *interpret*. Furthermore, we predict that if the adjunct clauses are located post-verbally, extraction is prohibited in general. Next, we examine these two points.

First let us test whether the extraction out of adjunct clauses in Japanese are affected by the type of matrix verb. In order to test this, however, we must consider one caveat. The type of the complementizer that is selected by the non-reporting verbs that Etxepare mentions are not the same as the complementizer selected by *say*-type verbs in Japanese. *Say*-type reporting verbs in Japanese normally select the declarative complementizer -*to*; however, non-reporting verbs, such as *mention*, *interpret* or *desire*, require the nominalizing complementizer -*koto*. Thus, we have to make sure that the nominalizing complementizer does not independently create island

effects. The examples in (163) show that scrambling out of *koto*-clauses does not create any detectable difficulty. Specifically, comparing it with scrambling out of complement clauses or RCs, it becomes clear that it does not show island effects induced by RCs (for further details of *koto*-clauses see Uchibori 2000 and Watanabe 1996).

(163) a. Scrambling out of Koto-clauses

Dono-gakusee-ni$_1$ Quinn-wa [Virgina-ga t$_1$

which-student-dat Q-top V-nom

present-o ageta-koto]-ni kizuita-no?

present-acc gave-fact-dat realized-Q

"Which student did Quinn realize that Virginia gave a present to?"

b. Scrambling out of Complement Clauses

Dono-gakusee-ni$_1$ Quinn-wa [Virginia-ga t$_1$

which-student-dat Q-top V-nom

present-o ageta-to] itta-no?

present-acc gave-comp said-Q

"Which student Quinn said that Virginia gave a present to?"

c. Scrambling out of Relative Clauses

*Dono-gakusee-ni$_1$ Quinn-wa [$_{NP}$[$_{CP}$ Virginia-ga t$_1$

which-student-dat Q-top V-nom

ageta]present]-ga sukina-no

gave present-nom like-Q

Lit. "Which student does Quinn like the present that Virginia gave to?"

We established that there is no problem with scrambling out of *koto*-clauses. Now, let us test the prediction of Etxepare-Hornstein generalization. The prediction is that extraction out of conditional clauses in the complement of reporting verbs is acceptable, but extraction from the complement of such verbs as *mention*, *interpret* and *desire* is degraded. The relevant Japanese examples are in (164). (164a) is an example of reporting verb, *iu* "say." (164b) is an example using the verb *nozomu* "desire". (164c) is example of scrambling out of multiply embedded complement clauses and (164d) is scrambling out of RCs, which are to be understood as the baseline example.

(164) a. *Reporting Verbs*

Dono-gakusee-ni$_1$ Quinn-wa [Stillman-ga
which-student-dat Q-top S-nom
[$_{cond}$ mosi Virginia-ga t_1 present-o ageta-ra]
mosi V-nom present-acc give-cond
naki-dasu-daroo-to] itta-no?
cry-start-will-comp said-Q
"Which student did Quinn say that if Virginia gave a present to Stillman would cry?"

b. Desire

Dono-gakusee-ni$_1$ Quinn-wa [Stillman-ga

which-student-dat Q-top S-nom

[$_{cond}$ mosi Virginia-ga t_1 present-o ageta-ra]

mosi V-nom present-acc give-cond

naki-dasu-koto]-o nozonde-iru-no?

cry-start-fact-acc desire-be-Q

"Which student does Quinn desire that if Virginia gives a present to Stillman cries?"

c. Multiply Embedded Complement Clause

Dono-gakusee-ni$_1$ Quinn-wa [Stillman-ga

which-student-dat Q-top S-nom

[Virginia-ga t_1 present-o ageta-to] itta-to]

V-nom present-acc gave-comp said-comp

omotteiru-no?

think-Q

"Which student does Quinn think that Stillman said that Virgia gave a present to?"

d. Multiply Embedded Relative Clause

*Dono-gakusee-ni$_1$ Quinn-wa [Stillman-ga

which-student-dat Q-top S-nom

[$_{NP}$[$_{CP}$ Virginia-ga t_1 t_2 ageta] present$_2$]-ga

V-nom gave present-nom

sukida-to] itta-no?

like-comp said-Q

Lit. "Which student did Quinn say that Still likes the present that Virginia gave to?"

There are two points to be noted. First, because of the basic word order of Japanese, the type of examples that we are looking at contains center-embedding structures. Thus, we have to prepare an appropriate baseline that contains a multiply center-embedding construction such as (164c) or (164d). Because of the complexity of the center-embedded structure, it is more difficult to judge them than others. However, if we compare (164a and b) to an appropriate baseline, like (164d), we can see that extraction out of conditionals are much more acceptable than extraction out of RCs, and if compared to (164c), they are as good as extraction out of complement clauses.

Second, because of the head-finality of Japanese, the complementizer comes at the end of each embedded clause, and we do not have any indicator of the left-edge of embedded clauses. Thus, to ensure that conditional clauses are located in higher positions they are located right below the embedded subject. We can locate conditional clauses to the left of the embedded subject too. However, in this case the sentence is ambiguous with respect to the position where the conditional clause is

interpreted. It can be either in the matrix clause or embedded clause. Still, under the intended interpretations in which conditional clauses are interpreted in the embedded clause, the same judgements seem to come out. The relevant examples are presented in (165). However, we do not know if these are really good test cases because of the ambiguity we noted.

(165) a. *Reporting Verbs*

Dono-gakusee-ni$_1$ Quinn-wa [[$_{cond}$mosi Virginia-ga
which-student-dat Q-top mosi V-nom
t_1 present-o ageta-ra] Stillman-ga
present-acc give-cond S-nom
naki-dasu-daroo-to] itta-no?
cry-start-will-comp said-Q

"Which student did Quinn say that if Virginia gave a present to Stillman would cry?"

b. Desire

Dono-gakusee-ni$_1$ Quinn-wa [[$_{cond}$ mosi
which-student-dat Q-top mosi
Virginia-ga t_1 present-o ageta-ra]
V-nom present-acc give-cond
Stillman-ga naki-dasu-koto]-o nozonde-iru-no?
S-nom cry-start-fact-acc desire-be-Q

"Which student does Quinn desire that if Virginia gives a present to Stillman cries?"

These examples suggest the following two points. First, extractability is not determined by the type of matrix predicate. Thus, long-scrambling out of conditionals is acceptable either from the complement of reporting verbs or that of *desire*-type verbs. The comparison between the examples in (164) and (165) suggest this point. Acceptability of these examples, thus, disconfirms the first generalization of Etxepare and Hornstein. Second, the position of conditional clauses does not affect their extractability either. The examples in (164) may be counter examples against the second generalization such that conditional clauses must be located in a position high enough, where they are adjoined to FP. At this point, we do not have clear tests that can diagnose where conditional clauses are located in examples like (164a) or (164b). It is not clear if conditionals are adjoined to FP or not. This is because it is not clear where Nominative subjects are generated in Japanese (Fukui 1995; Kuroda 1988; Saito 1982: Takezawa 1987 among others). It may be in the specifier of VP or it may be in the specifier of IP. However, what is clear from the examples in (164) is that conditionals are located below the subject. This is in contrast to the English and Spanish examples. Thus, it is plausible to assume that they are adjoined to VP rather than a projection higher than IP. Consequently, extractability in these examples does not straightforwardly follow from the Etxepare-Hornstein generalization, which states that an element can be extracted from a conditional clause that is adjoined to the FP.

However, the data from Japanese are not necessarily counterexamples to the cross-linguistic generalization of Etxepare-Hornstein. Rather, it is possible that the acceptability of scrambling out of conditionals is attributable to some special properties of Japanese, such as head-finality. Thus, to test the cross-language

generality of the generalization, we have to check other languages that show adjunct island effects. We return to this point later.

Let us turn to the second generalization, the Left/Right Asymmetry. The important point that Hornstein is raising is if a conditional clause is located to the right of the verb, extraction is not possible. The relevant minimal pair is presented in (166).

(166) a. Which book$_1$ did you say [that [if he ever read t_1] Quinn would **abandon** linguistics].

b. *Which book$_1$ did you say [that Quinn would **abandon** linguistics [if he ever read t_1]].

We can capture the contrast in (166) in two ways. One is that extraction is not allowed because the conditional clause is located to the right of the verb. The other is that extraction is not allowed because the adjunct is adjoined to VP rather than IP. At this point, it is difficult to test these two hypotheses. It is not clear whether they make different predictions. Simply taken, it as if the adjunct is located right of the verb, extraction is not possible.

What are the predictions of the hypotheses above when extended to Japanese adjunct clauses? We predict that an element cannot be extracted out of an adjunct clause if it is located to the right of the verb. On the other hand, if it is located to the left of the verb, extraction will be allowed. The problem is whether Japanese has such an operation that can locate an adjunct clause to the right of the verb. One possible

test case would be the 'so-called' right-dislocation construction. In right-dislocation, we can locate constituents to the right of the verb. In any of the other constructions, it does not seem to be possible to generate constituents to the right of the verb. Let us examine some examples of right-dislocation. (167b) is an example of right-dislocated complement clauses.[15] A prominent property of this construction in Japanese is that a special particle is attached to the matrix verb. If right-dislocation is applied, the special particle, so-called *shuu-josi* "end particle" is attached to the matrix verb.

(167) a. Quinn-wa [Virginia-ga Stillman-ni present-o
Q-top V-nom S-dat present-acc
ageta-to] itta.
gave-comp said
"Quinn said that Virginia gave a present to Stillman."

b. Quinn-ga $___1$ itta-yo, [Virginia-ga Still-ni
present-o said-part ageta-to]$_1$.

As we have seen in various places, complement clauses normally allow scrambling. However, once right-dislocation applies, scrambling is no longer licit. The minimal pair in (168) clearly shows this point.

[15] The original position of the right-dislocated constituent is indicated by the underscore __.

(168) a. present-o$_1$ Quinn-wa [Virginia-ga Stillman-ni t_1
present-acc Q-top V-nom S-dat
ageta-to] itta.[16]
gave-comp said
"Quinn said that Virginia gave a present to Stillman."

b. *present-o$_1$ Quinn-wa $__{}_2$ itta-yo, [Virginia-ga
Stillman-ni t_1 ageta-to]$_2$.

In (168b), the accusative NP is extracted out of the right-dislocated complement clause, and the sentence is unacceptable. This unacceptability is surprising given that, in Japanese, even moved constituents are not islands (Kikuchi et al. 1994) as the examples in (169) show. In (169b) the accusative NP is scrambled out of the scrambled complement clause. Both examples in (169) are acceptable, but (169b) is even easier to interpret because it avoids the interfering effect from multiple center-embedding.

[16] In this example, I am using a sentence containing the scrambling of an accusative NP. This is simply because the dative NP can be interpreted as the matrix argument, and thus may create unnecessary ambiguity. However, scrambling of the dative argument does not make any difference (in acceptability?). To keep the examples in this study as similar as possible, I continue using dative verbs in the embedded clause.

(169) a. present-o$_1$ Quinn-wa [$_{CP}$ Stillman-ga
present-acc Q-top S-nom
[$_{CP}$ Virginia-ga otooto-ni t_1 ageteta-to]
V-dat brother-dat gave-comp
itta-to] omotteiru.
said-comp think
"Quinn thinks that Stillman said that Virginia gave a present to her brother."

b. present-o$_1$ Quinn-wa [$_{CP}$[$_{CP}$ Virginia-ga t_1 ageta-to]$_2$ Stillman-ga t_2 itta-to] omotteiru.

The sharp contrast between (169b) and (168b) indicates that something special is going on in the right-dislocation construction.

Turning our attention to adjunct clauses, exactly the same pattern holds. As the examples in (170) indicate, scrambling out of conditional clauses in their original position is much more acceptable than that out of the right-dislocated conditionals.

(170) a. Quinn-wa [Stillman-ga __$_1$ naki-dasu-daroo-to]
Q-top S-nom cry-start-will-comp
itta-yo, [mosi Virginia-ga sono-gakusee-ni
said-part mosi V-nom that-student-dat
present-o ageta-ra]$_1$(-ne).
present-acc give-cond(-part)
"Quinn said that Still will cry if Virginia gives a present to that student."

b. *present-o$_1$ Quinn-wa [Stillman-ga __$_2$
naki-dasu-daroo-to] itta-yo, [mosi Virginia-ga
sono-gakusee-ni t_1 ageta-ra]$_2$(-ne)

c. present-o$_1$ Quinn-wa [Stillman-ga [mosi
present-acc T-top J-nom mosi
Virgnia-ga sono-gakusee-ni t_1 ageta-ra]
V-nom that-student-dat gave-cond
naki-dasu-daroo-to] itta.
cry-start-will-comp said

Given these examples showing that the right-dislocated constituents are islands, it seems that the simple version of the Left/Right Asymmetry is supported. However, the story is not so simple. As we have seen, it seems something special is going on in the examples of right-dislocation. Therefore, in order to draw any conclusions, we have to investigate the syntax of right-dislocation and confirm that

there is no special factor that can independently block extraction out of dislocated constituents.

A possible account for the islandhood of the dislocated phrase is the following. Extraction is not allowed because the dislocated constituent is not part of the sentence.[17] Let us examine this possibility.

The intuition behind this account is that the right-dislocated constituent is a special element that is not a part of the structure of the other part of the sentence. Put differently, it is disconnected from the matrix clause. Theoretically, this intuition might be captured by postulating that the right-dislocated structure has its own root in a sentence, in the a manner similar to the 'so-called' parenthetical constructions under certain analyses (McCawley 1982).[18] This analysis allows us to analyze right-dislocation constructions in the same way as parenthetical constructions. A possible account for the islandhood of the dislocated phrase is the following. If the right-dislocated constituent has its own root, then movement out of it means an element is moved out of a "sentence" to the beginning of another "sentence". As far as I know, such inter-sentential movement operation is not allowed. Thus, according to this account, extraction out of the dislocated phrase is also not allowed.

There are, however, at least two problems with this approach. First, it is not clear whether Japanese has parenthetical constructions. We do not have any good diagnoses that can distinguish parentheticals from the other constructions. Thus, we

[17] This possibility is suggested by Ilhan Cagri. I thank her for leading my attention to this possibility.

[18] For further details of multiple rooted structures and shared constituency, see Citko (2005), Moltmann (1992), Wilder (1999) among others. Guimarães (2004) also contains a detailed discussion on this point.

cannot test whether they are similar or not. Second, there is evidence against a structural disconnection between the main clause constituents and right-dislocated constituents. One such argument comes from binding connectivity and subjacency effects and another from an example of discontinuous constituency.

It is reported in the literature that right-dislocated phrases show binding connectivity effects and subjacency effects. Both are typical of constructions involving movement. Tanaka (2001) shows various patterns of connectivity effects exhibited by right-dislocation constructions. Examine the following pair of examples. (171) is the case of Binding Condition C. The right-dislocated constituent shows a binding connectivity effect in (171c) in which the R-expression in the relative clause is bound by the right-dislocated subject pronoun. The example is unacceptable.

(171) a. *Kare$_1$-ga [$_{NP}$[$_{CP}$ Virgnina-ga Quinn$_1$-ni okutta]
he-nom V-nom Q-dat sent
tegami]-o mada yonde-inai.
letter-acc yet read-neg
"He$_1$ has not yet read the letter that Virginia sent to Quinn$_1$."

b. [$_{NP}$[$_{CP}$ Virginia-ga Quinn$_1$-ni okutta]tegami]-o$_2$
V-nom Q-dat sent letter-acc
kare$_1$-ga t$_2$ mada yonde-inai.
he-nom yet read-not

c. *__$_1$ [$_{NP}$[$_{CP}$ Virginia-ga Quinn$_2$-ni okutta]
V-nom Q-dat sent
tegami]-o$_3$ mada yonde-inai-yo, kare$_2$-ga$_1$.
letter-acc yet read-neg-part he-nom

This connectivity effect, as well as other connectivity effects that Tanaka is citing, potentially shows that the dislocated element is moved rightward, and holds a syntactic relation with the matrix clause.

Second, right-dislocation shows a certain locality effect that can be presumably subsumed under the Subjacency Condition. (172c) exhibits severe unacceptability. This parallels with scrambling out of RCs in (172b). As opposed to right-dislocation out of RCs, extraction out of complement clauses does not show such degradation.

(172) a. Quinn-ga [$_{NP}$[$_{CP}$ Virginia-ga Stillman-ni
Q-nom V-nom S-dat
ageta]hon]-o nusunda.
give book-acc stole.
"Quinn stole the book that Virginia gave to Stillman"

b. *Stillman-ni$_1$ Quinn-ga [$_{NP}$[$_{CP}$ Virginia-ga t$_1$
ageta]hon]-o nusunda.

c. ?*Quinn-ga [$_{NP}$[$_{CP}$ Virginia-ga __$_1$ ageta]hon]-o
nusunda-yo, Bill-ni$_1$.

d. Quinn-ga [Virginia-ga $___1$ hon-o ageta-to]
Q-nom V-nom book-acc gave-comp
itta-yo, Stillman-ni$_1$.
said-part S-dat
"John said that Mary gave a book to Bill."

These examples potentially show that dislocated phrases are moved rightward. However, Tanaka argues that there is another explanation. He argues that right-dislocation is a special case of scrambling with ellipsis under identity. According to him, there is an elided structure that maintains the parallelism with the matrix clause. The dislocated element is actually scrambled inside the hidden clausal structure. Thus, Tanaka argues that right-dislocation and scrambling show close parallelisms, and these locality and connectivity effects are induced by scrambling rather than rightward movement. Although Tanaka does not touch on the relation between the matrix clause and dislocated phrases[19] and does not argue against the structural

[19] Endo (1996) argues that right-dislocation constructions are derived from coordinated clauses via movement of the dislocated constituent and deletion. Thus, his analysis is basically similar to the one that Tanaka is exploring. In his (whose?) analysis, the relation between the matrix clause and the dislocated constituent is clear. However, because his analysis is appealing to a scrambling-like leftward movement of the dislocated constituent, it will face the same problem as Tanaka's as we will see shortly.

However, if we follow Endo's approach and assume that right-dislocation constructions have the structure of Coordination, it is possible to analyze the islandhood of the dislocated phrases as a violation of the Coordinate Structure Constraint. An obvious prediction of this approach is if we extract elements in an across-the-board fashion, extraction is allowed. Although this is potentially an

disconnection either, this approach is compatible with the structurally disconnected analysis of right-dislocation. Under this analysis, the structure of a right-dislocated sentence looks something like (173).

(173) $[_{IP}$ NP $___1$ V]-yo, $[_{IP}$ NP_1 $[\sout{_{IP}\ NP\ t_1\ V}]]$.

A closer look at the examples of right-dislocation, however, reveals that this approach is not perfect, although it can capture many of the basic properties of the construction. A potential counter-argument against the scrambling approach comes from the possibility of discontinuous constituency. The example in (174b) suggests that right-dislocation can disconnect a relative clause from its head noun.[20]

(174) a. Quinn-ga $[_{NP}[_{CP}$ Virginia-ga katta] hon]-o
Q-nom V-nom bought book-acc
yonda.
read
"Quinn read the book that Virginia bought."

b. Quinn-ga $[_{NP}$ $___1$ hon-o] yonda-yo, $[_{CP}$ Virginia-ga katta$]_1$.

---

interesting possibility, we are not adopting Endo's analysis because of the problem that the analysis faces. If you're not adopting it, then is all this necessary?

[20] This is pointed out to me by Tomohiro Fujii (p.c.). I thank him for letting me know about this type of example.

To the best of my knowledge, there are two constructions that can split a relative clause and its head noun in Japanese: right-node raising and right-dislocation. However, no other constructions allow such detachment of relative clauses. For example, neither scrambling nor clefting results in an acceptable sentence if a relative clause alone is moved. (175a) is an example of scrambling and (175b) is an example of clefting.

(175) a. *[$_{CP}$ Virginia-ga katta]$_1$ Quinn-ga [$_{NP}$ t_1 hon]-o
V-nom bought Q-nom book-acc
yonda.
read
"Quinn read the book that Virginia bought."

b. *Quinn-ga [t_1 hon]-o yonda-no-wa,
Q-nom book-acc read-comp-top
[$_{CP}$ Virginia-ga katta]$_1$-da
V-nom bought-cop
Lit."It is that Virginia bought that Quinn read the book."

The contrast between scrambling and right-dislocation cannot be captured by Tanaka's scrambling and ellipsis approach. The simplest way to reconcile Tanaka's observations and the contrast between scrambling and right-dislocation is to assume that right-dislocated phrase is indeed rightward moved as in the extraposition construction or heavy-shift construction in English. As it is known that English extraposition from NP also can detach a relative clause from its head noun (Johnson

1986; Coopmans & Roovers 1986 among others). Thus, we can assume that right-dislocation in Japanese is a variant of rightward movement that is found in English.

(176) [$_{NP}$ Virginia reviews __$_1$] have been published [$_{CP}$ which criticized Quinn's book]$_1$.

If right-dislocation is a type of rightward movement, then the connectivity effects should be taken as that the main clause R-expression and the dislocated pronoun in (171c) has a direct c-command relation at some point of derivation. Thus it supports the position that the right-dislocated phrases take part in the main clause structure in some way, i.e., they are not really disconnected. This entails that the multiple root analyses that potentially disallow the communication between clauses with different roots cannot capture the direct relation between the main clause element and the dislocated phrase.

If we do not appeal to a discourse/pragmatic explanation for which we do not have any clear arguments at this point[21], the examples that we have seen so far may suggest that Left/Right-Asymmetry is an appropriate generalization that can account for the islandhood of dislocated constituents. In particular, given that derived position islands (Merchant 2001; Takahashi 1993; Wexler and Culicover 1981 among others) are not operative in Japanese as we have seen, the Left/Right-Asymmetry seems to be the right generalization. In other words, it seems that moving to the right of the verb

[21] Kuno (1973, 1978) is investigating the interpretive aspects of right-dislocation. However, it is not discussed how the interpretive properties of right-dislocation is related to the extractability.

is a crucial factor for a constituent to be an island. Now the question is how we can explain this generalization. The explanation of Left/Right-Asymmetry is beyond the scope of this study. Thus, I would like to leave this problem open for the future research.

4.2. A Parametric Study on Adjunct Islands

In the previous discussion, we concluded that the Left/Right-Asymmetry is the right generalization cross-linguistically. However, we have also seen that, as long as adjunct clauses are generated to the left of the clause, extraction out of adjunct clauses are allowed in Japanese even though their position is lower than what Etxepare and Hornstein argue, i.e., the position between the embedded subject and the complementizer. In this subsection, we try to explain this difference between Japanese and Spanish /English.

4.2.1. Languages that Allow Extraction out of Adjuncts

At the outset of this discussion, we have seen that at least two languages show the same pattern as Japanese with respect to the extraction out of adjunct clauses. Let us first take a look at data from Korean and Malayalam and try to figure out what is similar among these languages.

The following are example of scrambling out of conditional clauses in Korean.

(177) Conditional Clauses

Quinn-un [manyak Virginia-ka Stillman-hanthey
Q-top [cond-adverb V-nom S-dat
senmwul-ul cwu-myen] wul-keya.
present-acc gave-cond] cry-will.
"Quinn will cry if Virginia gives a present to Stillman."

a. Scrambling of Referential NP

Bill-hanthey$_1$ John-un [manyak Mary-ka t_1

B-dat J-top [cond-adverb M-nom

senmwul-ul cwu-myen] wul-keya.

present-acc gave-cond]cry-will.

b. Scrambling of wh-phrase out of conditionals

Etten-haksayng-hanthey$_1$ Quinnn-un [manyak

Which-student-dat1 Q-top [cond-adverb

Virginia-ka t_1 senmwul-ul cwu-myen wul-ul-ka?

V-nom present-acc gave-cond] cry-will-Q

Lit. "Which student will Quinn cry if Virginia gives a present to?"

(178) Relative Clauses

Quinn-un [$_{NP}$[$_{CP}$ Virginia-ka Stillman-hanthey cwu-n

Q-top V-nom S-dat gave-adnom]

senmwul-ul coahay.

present]-acc like.

"Quinn like the present that Virginia gave to Stillman."

a. Scrambling out of Relative Clauses

*Stillman-hanthey$_1$ Quinn-un [$_{NP}$[$_{CP}$ Virginia-ka t_1

B-dat J-top M-nom

cwu-n senmwul-ul coahay.

gave-adnom] present]-acc like.

b. WH-scrambling out of Relative Clauses

*Etten-haksayng-hanthey$_1$ Quinn-un

Which-student-dat Q-top

[$_{NP}$[$_{CP}$ Virginia-ka t$_1$ cwu-n senmwul]-ul coahay?

V-nom gave-adnom]present]-acc like

Lit. "Which student does Quinn like the present that Virginia gave to?"

(179) Complement Clause

Quinn-un [$_{CP}$ Virginia-ka Stillman-hanthey senmwul-ul

Q-top V-nom S-dat present-acc

cwuessta-ko] saynggakha-n-ta.

gave-comp think

"John thinks that Mary gave a rsent to Bill"

a. Scrambling out of Complement Clauses

Stillman-hanthey$_1$ Quinn-un [Virginia-ka t$_1$

S-dat Q-top [V-nom

senmwul-ul cwuessta-ko] saynggakhay.

present-acc gave-comp] think

b. Wh-scrambling out of Complement Clauses

Etten-haksayng-hanthey$_1$ Quinn-un [$_{CP}$ Virginia-ka

Which-student-dat Q-top V-nom

t$_1$ senmwul-ul cwuessta-ko saynggakhay?

present-acc gave-comp] think

The following are relevant data from Malayalam.

(180) a. Conditional Clauses

Quinn [Virginia Stillman-inu sammaanam

Q-nom V-nom S-dat present-acc

kodu-thaal] karayum.

give-cond cry-will

"Quinn will cry if Virginia gives a present to Stillman."

b. Scrambling out of Conditional Clauses

Stillman-inu$_1$ Quinn [Virginia t_1 sammaanam

S-dat Q-nom V-nom present-acc

kodu-thaal] karayum.

give-cond cry-will

(181) a. Relative Clauses

[$_{NP}$[$_{CP}$ Quinn [$_{CP}$ Virginia Stillman-inu t_1

Q-nom V-nom S-dat

kodu-thennu] viswasikkane] pustakam$_1$

gave-comp believe-adnom book

"The book that John believes that Mary gave to Bill"

b. Scrambling out of Relative Clauses

*Stillman-inu$_1$ [$_{NP}$[$_{CP}$ Quinn [$_{CP}$ Virginia t_1 t_2

S-dat Q-nom V-nom

kodu-thennu] viswasikkane] pustakam$_2$

gave-comp believe-adnom book

(182) a. Complement Clauses

Quinn [$_{CP}$ Virginia Stillman-inu sammaanam
Q-nom V-nom S-dat present-acc
kodu-thennu] viswasikkunu
gave-comp believe
"Quinn believes that Virginia gave a present to Stillman."

b. Scrambling out of Complement Clauses

Stillman-inu$_1$ Quinn [$_{CP}$ Virginia t_1 sammaanam
S-dat Q-nom V-nom present-acc
kodu-thennu] viswasikkunu
give-comp believe

As we can see, both of these languages show exactly the same pattern as Japanese. In all these cases, extraction out of conditional clauses is much better than that out of relative clauses.

Based on the data we have, we can conclude that there are at least two types of languages. In the first type, adjunct clauses are islands if they are not adjoined to FP in the embedded clause. In the second type, adjuncts are not islands as long as they are generated to the left to the verb.[22]

[22] If we consider languages such as Swedish or Danish, we can have the other type of language in which even relative clauses, as well as adjunct clauses, are not islands (Engdahl 1986, 1982). In all the other languages at hand, relative clause islands are respected. Considering the fourth type of languages is indeed interesting, but it is beyond the scope of this study. Thus, I leave this problem open for future research.

4.2.2. Parametric Features

The question is what is the feature that distinguishes languages like English or Spanish and Japanese, Korean and Malayalam. Are there any typological features that can distinguish these two types of languages? Do the cross-linguistic similarities or differences with respect to the extraction out of adjunct clauses systematically follow from those features? Even if we can find some significant similarities among these three languages, if these similarities do not have any connection to the extraction out of conditional clauses, the problem of extractability remains unresolved. Thus, it is desirable to find typological features that are shared by these languages and are related in some way to extraction out of adjunct clauses. In the following discussion, I try to answer the first question, but the answer for the second question remains open.

One of the prominent features that Japanese, Korean and Malayalam share is head-finality. Can we derive the extraction patterns from the head-finality? There is reason, however, to believe that head-finality is not the crucial feature that derives the adjunct island asymmetry, i.e., there are head-final languages that show adjunct island effects in the same way as Spanish and English. Basque is one such language.[23] In Basque, a conditional clause is an island if it is in the complement of non-reporting verbs such as *desire* or *want*, or if it is located to the right of the embedded verb. If they are located in a higher position, overt wh-movement is allowed. We can see this pattern in the following examples in (183).

[23] I would like to express my gratitude to Leticia Pablos for native speaker's judgment about the Basque sentences.

(183) a. Zein liburu pro **esan** zenuen [Rikardok$_1$
which-book say aux-past R-erg
irakurtzen (baldin) bazuen] pro$_1$
read if if-aux
hizkuntzalaritza utziko zuela?
linguistics leave-future aux-comp
"Which book do you say that if Rikardo read he would abandon linguistics?"

b. *Zein liburu pro **esan** zenuen pro$_1$
which book say aux
hizukunzelaritza utziko zuela [Ricardok$_1$
linguistics leave-future aux-comp R-erg
irakurtzen (baldin) bazuen]?
read if if-aux

c. *Zein liburu pro **nahi** zenuen [Rickardok$_1$
which book want aux-past R-erg
irakurtzen (baldin) bazuen] pro$_1$
read if if
hizkuntzalaritza utziko zeula?
linguistics leave-aux aux-comp
"Which book do you desire that if Rikardo read he would abandon linguistics?"

As is evident from the examples, matrix verbs precede complement clauses in Basque. Thus, Basque is not a strongly head-final language like Japanese. However, the language allows head-final configurations, and still extraction from adjunct clauses patterns very much like Spanish or English. Therefore, we can plausibly conclude that head-finality is not the right feature, or we cannot explain the paradigm only by head-finality.

Looking at these languages that allow extraction from adjunct clauses, it becomes clear that there are other similarities, in addition to the head-finality. In particular, these languages share such properties as scrambling, wh-in-situ and an indeterminate system, as well. These properties, in particular, seem to distinguish these three languages from languages like Basque, English or others. Thus, it is worthwhile taking a look at these properties and check whether we can successfully derive the cross-linguistic differences between two types of languages.

Let us first review these properties. As it is well known, both in Korean and Malayalam, wh-phrases remain in-situ[24], while they can be scrambled to the beginning of a clause, as illustrated in (177) through (182). Examples of simple wh-questions are in (184).

[24] Note that Jayaseelan (2001) suggests that non-cleft wh-question in Malayalam involves IP-internal focus movement even though the surface word order is the same as basic word order, SOV. In this study, however, we basically treat Malayalam as wh-in-situ language in the sense that wh-phrases do not obligatorily move to the sentence initial position, and they are pronounced in their base position.

(184) a. Korean Wh-Question (Beck & Kim 1997)

Suna-ka muôs-ûl sa-ss-ni?

Sun-nom what-acc buy-past-Q

"What did Suna buy?"

b. Malayalam Wh-Questions (Jayaseelan 2001)

nii entə aaNə tiinn-atə?

you what is ate-nominalizer?

"What did you eat?"

Second, in all the three languages, Japanese, Korean and Malayalam, wh-words are used as so-called indeterminate pronouns (Hiraiwa 2002; Kuroda 1965; Nishigauchi 1990; Shimoyama 2001; Watanabe 1992a, 1992b, 1996, 2001, 2002, 2003). In these languages, wh-words are used in various ways depending on the focus particles they attach to. The full paradigms of the indeterminate system in these languages are summarized in the table below. As we can see, if a wh-phrase is associated with the conjunction marker, which is also used to express *also* in Japanese, then they are interpreted as universal quantifiers. If they are associated with the disjunction marker, they are existential quantifiers.

(185) Indeterminate system

Japanese			Korean			Malayalam		
Wh	**Wh-conj**	**Wh-disj**	**Wh**	**Wh-conj**	**Wh-disj**	**Wh**	**Wh-conj**	**Wh-disj**
dare	dare-mo	dare-ka	nwukwu	nwukwu-wa	nwukwu-na	aar	aar-um	aar-oo
(who)	(anyone)	(somebody)	(who)	(anyone)	(somebody)	(who)	(anyone)	(somebody)
nani	nani-mo	nani-ka	mwuess	mwuess-kwa	mwuess-ina	ent	ent-um	ent-oo
(what)	(anything)	(something)	(what)	(anything)	(something)	(what)	(anything)	(something)
doko	doko-mo	doko-ka	eti	eti-wa	eti-na	eŋŋ	eŋŋ-um	eŋŋ-oo
(where)	(anywhere)	(somewhere)	(where)	(anywhere)	(somewhere)	(where)	(anywhere)	(somewhere)
itu	itu-mo	itu-ka	ence	ence-na	encenka	eppoozh	eppoozh-um	eppoozh-oo
(when)	(always)	(sometime)	(when)	(always)	(sometime)	(when)	(always)	(sometime)

The other three languages we have seen (English, Spanish and Basque) do not share all of these properties, though they share some of them. Taking English as an example, it is a head-initial language and does not have the indeterminate system that Japanese, Korean or Malayalam does. In English, wh-phrases and quantifiers are formed in different ways. For example, a 'person' wh-phrase is *who*, but the universal quantifier for the person is not *who-and*, or *who-also* but *someone*, an independently used nominal element like *one* and a determiner like *some* is used to form an existential quantifier. The same holds true for Basque (Hualde and Urbina 2003).

As we have shown in the case of the head-finality, we can also show that either the indeterminate system or wh-in-situ alone is not a crucial feature for the extractability out of adjunct clauses. For example, Russian has the same type of

indeterminate system as Japanese, Korean and Malayalam (Watanabe 2002), but Russian shows adjunct island sensitivity (Stepanov 2001). Crucially, Russian is wh-movement language. On the other hand, the possibility of wh-in-situ is not sufficient to derive the pattern of adjunct non-islandhood either. Malay allows wh-in-situ as well as overt wh-movement and partial wh-movement (Cole and Hermon 1998). However, Malay overt wh-movement exhibits adjunct island effects (Cole and Hermon 1998). Finally, the availability of scrambling does not seem to be the crucial factor either. As we have discussed, Russian shows adjunct island effects for overt movements even though Russian is indeed a language that allows scrambling (Bailyn 2001, 2002 among many others).

These cross-linguistic patterns suggest that the combination of these features (head-finality, wh-in-situ, indeterminate system and scrambling) seems to be crucial in deriving the adjunct (non)-islandhood. In other words, if we take the strongest position based on the data at hand, we can conclude that only the languages that have these four properties allow extraction out of adjunct clauses.

A note is in order. As it shall be clear from the data we have presented and discussion thus far, this conclusion is probably too strong. We have not checked other properties that are shared by these languages. However, the prediction of the theory of cross-language differences adopting this conclusion is quite clear. If there is a language that shares all these four properties, it should allow extraction out of pre-verbal adjunct clauses.

The remaining question is how we can derive the adjunct non-islandhood from these parametric features. In so doing, we have to answer how these features are related to extraction. At this point, I leave this problem open.

5. Conclusions

In this chapter, we have discussed the syntax of RCs and Conditional clauses. By applying various grammatical tests, we have revealed the internal syntax and external syntax of these two constructions.

We have also tackled the problem of the non-islandhood of Japanese adjunct clauses. Based on a cross-linguistic study, we found that there is a cluster of parametric features that seems to be responsible for the non-islandhood of adjunct clauses. However, the question of how we can derive the cross-linguistic patterns has been left open.

CHAPTER 3. RELATIVE CLAUSE PREDICTION IN JAPANESE

1. Introduction

Current studies in sentence processing suggest that human sentence processing is incremental, in the sense that the processor incorporates input words into a grammatical analysis as soon as they are encountered, and thus structural commitments are made and the interpretation of sentences is constructed from left to right without delay (Marslen-Wilson 1973). With this as a background, a widely adopted assumption is that the human sentence processor is strongly incremental, i.e., the sentence processor builds a fully connected structure as the processor receives each input word, where the processor does not allow unstructured or partially structured input that is stored as unconnected pieces (Frazier and Rayner 1987; Gibson 1991; Gorrell 1995; Inoue and Fodor 1995; Stabler 1994). Recently, it has been suggested in the literature that a strongly incremental parser requires a powerful predictive mechanism exploiting "extra-lexical" knowledge (Lombardo and Sturt 2002; Sturt and Crocker 1996; Schneider, 1999; among others). The aim of this study is to show that the human sentence processing mechanism is indeed equipped with a powerful predictive mechanism, through experimental studies on various aspects of relative clauses in Japanese. We will argue that the predictive mechanism in the human sentence processor has the ability to project the sentence structure that goes beyond the local structure that can be constructed based on the information available from each lexical item.

2. Necessity of extra-lexical knowledge in sentence processing

The necessity of extra-lexical knowledge by an incremental parser is typically illustrated by examples like (1). In (1), a strongly incremental parser incorporates the adjective “steeper,” into the current representation without waiting for other lexical items to become available. In order to do so, however, the parser has to build the structure of the NP and IP categories, which do not have overt heads at the point where the word “steeper” is encountered. The parser has to project an NP node to host the adjective in this environment. Furthermore, the IP node is necessary because the NP that is predicted by the adjective must be attached as a specifier of IP. At the point of "steeper" the most likely position for the NP in this partial context is the subject of an embedded clause. Building this structural skeleton requires the projection of IP based on the predicted NP node, for an attributive adjective must be licensed by an NP. While it is plausible to assume that the presence of an attributive adjective can set a prediction for the upcoming NP node, how the IP node be predicted is not at all obvious. This is so because there is no direct grammatical relation between an adjective and an IP node[25]. In other words, there is no requirement that an adjective must be licensed by an IP node or vice versa. The only conceivable requirement imposed by an adjective in a sentence like (1) is that it must be associated with an NP. Thus, based solely on the grammatical requirement of an adjective, there is no reason that the parser should project up to an IP node upon encountering an adjective. Hence,

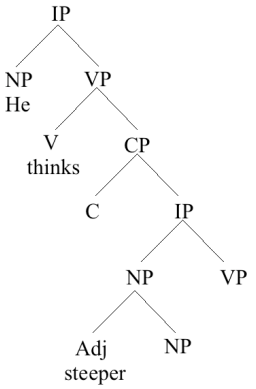
[25] Prediction of the CP node is not so problematic because a verb like “think” takes a clausal complement. Therefore, it is plausible to assume that the selectional information of the verb allows the prediction of CP.

this example, in turn, suggests that there are situations in which incremental structure building has to commit global syntactic inferences that can project the structure that goes beyond the local structure that can be projected using grammatical requirements stored in an input lexical item.

(1) a. He thinks [$_{CP}$[$_{IP}$[$_{NP}$[$_{Adj}$ steeper] prices] have come about because …]]

(Lombardo and Sturt 2002: 138)

b. .



A theoretical question that an example such as (1) raises is what kind of parsing algorithm can successfully process the sentence without sacrificing incrementality. In the history of psycholinguistics, various parsing algorithms have been proposed in an effort to formulate a psychologically plausible sentence processing mechanism. Taking one example, the so-called left-corner strategy (Abney and Johnson 1991; Aho and Ullman 1972; Johnson-Laird 1983; Resnik 1992; Stabler 1994) has often been considered as a psychologically plausible model of sentence

processing in the sense that it enable us to maintain incremental sentence processing with a reasonable power of prediction of the upcoming structure. The basic intuition behind the left-corner strategy is the following. Initially, structure is built in a bottom-up fashion using the information from an incoming lexical item. After any constituent is completed, however, its parent node is built and potential daughter nodes of the completed constituent are predicted.

Although, the left-corner algorithm maintains incremental sentence processing, even a sentence processing mechanism incorporating this algorithm cannot incrementally parse the sentence illustrated in (1). This is so because the left-corner algorithm creates a prediction for an upcoming node based on the bottom-up information of the input lexical items. In (1), the left-corner strategy allows the parser to predict the NP node and its potential siblings, based on the grammatical information from the adjective *"*steeper.*"* However, the prediction of the IP node is beyond its reach for the reasons we have discussed above. Therefore, a sentence like (1) raises an interesting theoretical question about the nature of a psychologically plausible parsing algorithm or the nature of the predictive mechanism of the human sentence processor.

From a different perspective, an example like (1) raises another theoretically important question, i.e., whether the human sentence processing mechanism should be understood as incremental at all. If the human parser does not maintain incrementality, and thus waits until a crucial lexical item appears in the input, the problem of non-local projection does not arise in the first place. All else being equal, such a powerful mechanism should be avoided if we do not have a really strong

motivation. The so-called head-driven model or delay model in more general terms (Abney 1989; Pritchett 1991a, 1992a) allows us to handle a sentence like (1) without postulating such powerful predictive mechanisms.

If structure building is driven entirely by lexically encoded information that is recovered from overt material in the input, then the processing of sentences like (1) requires a certain delay. This is so because the lexical item *"steeper"* does not provide any specific cue for the IP node as we have observed. Put differently, if the parser delays building the structure until the grammatical information of the input lexical items become available, it should not incorporate the attributive adjective into the structure immediately upon encountering it. The parser should then construct the sentence structure when it encounters the host noun for the attributive adjective, for instance. Given this alternative scenario, we are forced to ask the question whether the parser delays the processing of a sentence until crucial lexical items appear in the input or whether it builds the structure dynamically by exploiting detailed extra-lexical syntactic knowledge with a powerful predictive mechanism.

3. On the Prediction of Japanese Relative Clauses

For the two issues that we have discussed so far, Japanese may provide a good testing ground. Because of the strong head-finality of its basic sentence structures, the online sentence processing of Japanese creates situations similar to (1) in almost every sentence.

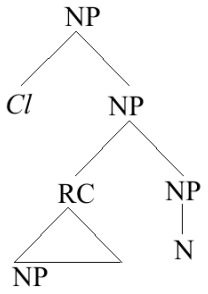
In Japanese, the head of each phrase follows all other elements in the phrase, both modifiers and arguments. Accordingly, the sentence processor is not normally

provided with reliable cues about the upcoming sentence structures until it encounters the head at the end of each phrase (Hirose 1999; Inoue 1991; Mazuka and Lust 1988, 1990; Mazuka and Itoh 1995; Miyamoto 2002, 2003). Among such head-final constructions in Japanese, relative clauses have attracted interest from researchers attention because they create notorious garden path effects (Hirose 1999; Inoue 1991; Mazuka and Lust 1988, 1990; Mazuka and Itoh 1995; Miyamoto 2002, 2003; Nakamura 1999/2000, 2003; Yamashita et al. 1993; Yamashita 1995). It is generally assumed in the literature that garden path effects are created because in Japanese an upcoming relative clause structure is extremely difficult to predict. Let us discuss this point more in detail.

It has been observed that the parser shows a bias to construct a simple clause or a complement clause structure whenever possible (Inoue 1991; Mazuka and Lust 1988, 1990; Mazuka and Itoh 1995; Miyamoto 2002, 2003; Yamashita et al. 1993; Yamashita 1995). Thus, a relative clause structure is not counted as the initially preferable structure. From this initial preference for a complement clause analysis, it follows that when a disambiguating cue from the relative head or morphology of the verb is provided, a garden path effect may arise because the parser has to reanalyze the initial complement clause structure to a relative clause structure. The garden path effect is caused because no syntactic or morphological cues for the presence of the relative clause structure are provided at an earlier point in parsing, unlike relative clauses in English where the presence of relative pronouns can mark the beginning of a relative clause. This implies, however, that this garden path effect may be resolved if an upcoming relative clause structure can be predicted in advance.

A careful examination of the syntactic properties of Japanese noun phrases, however, reveals that a certain class of modifiers of NPs, the so-called numeral classifiers, may provide an unambiguous cue for an upcoming relative clause structure. The structure in (2) illustrates this schematically. As we shall see in more detail below, in a specific context where a classifier cannot be associated with its linearly adjacent NP, it must be associated with the head of a relative clause. Under this circumstance the material that intervenes between the classifier and the host NP must be a relative clause.

(2)



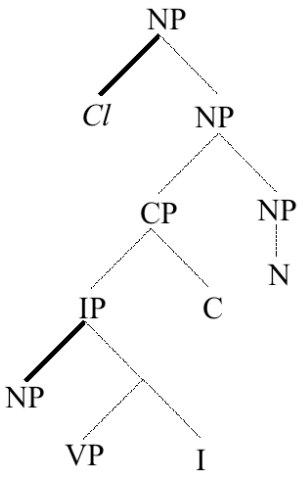
This relation between the classifier and the relative head NP suggests, in turn, that if the parser can employ this information in the course of parsing, it can create a prediction for the upcoming relative clause structure. In other words, in certain environments the classifier can provide an unambiguous cue for an upcoming relative clause during online sentence processing.

Although it is a plausible consideration that numeral classifiers may indicate an upcoming relative clause structure to the parser, how the parser creates the prediction is not at all obvious. This is because classifiers do not have any direct grammatical relation to relative clauses. A classifier modifies a noun phrase but it

does not modify a relative clause. In the same way a relative clause modifies a noun phrase, but it does not modify a classifier. Thus, even though a relative clause can intervene between a classifier and its host NP, it does not imply that the classifier and the relative clause have any direct relation to each other. They are both associated with a noun phrase but not with each other. In other words, the presence of a classifier does not grammatically imply the presence of a relative clause.

With the discussion so far in mind, let us consider how a strongly incremental parser might project an upcoming relative clause structure. If the parser is strongly incremental, it has to project the following structural skeleton based on the encounter with a numeral classifier that cannot be associated with its adjacent NP in order to incorporate these elements into a sentence structure. First, the position occupied by the classifier has to be projected. Classifiers are licensed by an NP, so an NP node has to be projected. Second, an IP node is required, where the NP containing the classifier is incorporated. Furthermore, a CP node must be projected in which the IP is inserted. Finally, the CP has to be connected to the NP as a relative clause. In (3), this process is illustrated (dotted lines indicate predicted nodes).

(3)



If this is the process that the parser employs, several problems arise. Essentially they are the same problems as we have seen in the discussion of the English example. First the problem of non-headed structures arises as we have seen in (1). In (3), IP, CP and NPs have to be projected without the cues from their heads. Second, it is not clear how the classifier, a modifier of the NP can specifically provide a cue for the relative clause structure, as we have discussed. The lexically encoded information of the classifier does not provide specific cues for the structure of relative clauses. The requirement that any numeral classifier demands is that it must be associated with an appropriate noun phrase. Given that a numeral classifier does not have any conceivable direct relation to relative clauses, it is not plausible to assume that the presence of a classifier directly cues the presence of an upcoming relative clause. Thus, how the relative clause structure is predicted is not obvious from the information encoded in/with each lexical item such as numeral classifiers. Finally, if a prediction for an upcoming relative clause is ever possible, the parser has to be equipped with a powerful predictive mechanism. This predictive mechanism must allow the parser to project the NP, IP, CP nodes and how the CP is connected to the NP, upon encountering a situation in which a classifier cannot be associated with its adjacent NP.

Now it should be clear that these problems do not arise if the parser delays the processing of sentences until the crucial lexical item appears in the input. The parser just builds the structure when the unambiguous cues from the relative head NP or the embedded verb become available. Hence, in this view, headless predicted projections

are not needed. This, in turn, means that the parser does not need to have the powerful predictive mechanism while the incrementality is sacrificed.

The goal of this study is to investigate the two points we have discussed above through experimental studies on Japanese relative clauses. Through a detailed experimental examination of various aspects of relative clauses in Japanese, we will provide supporting evidence for the position that the parser maintains the strong incrementality and that it is equipped with a powerful predictive mechanism of the sort that we have briefly discussed above. There are two major findings in this study. The first is that the parser is able to create a prediction for an upcoming relative clause structure upon encountering a numeral classifier that is not semantically compatible with its adjacent NP. This finding provides evidence for the view that the parser can project structure beyond the highly local structures that can be projected based on the input lexical information. Secondly, we will show that, in addition to the prediction for a relative clause structure, the parser can compute at least one consequence of this prediction, namely the islandhood of the relative clause. We will see that computing the islandhood of relative clauses requires abstract and detailed syntactic inferences. Three experimental results are reported that additionally support the existence of a powerful predictive parsing mechanism and that provide further evidence against delay models of sentence processing.

4. The Processing of Japanese Relative Clauses

Now, let us briefly summarize more concretely the basic issues in the processing of Japanese relative clauses that we are concerned with. Since Inoue's (1991) studies on Japanese parsing, it has generally been agreed that the beginning of embedded clauses in Japanese is hard to detect on-line. Japanese lacks obvious markers of the beginning of an embedded clause. Comparative syntactic studies between Japanese and English help us to understand this point more clearly. Let us cite some clear cases. In Japanese, a complementizer appears at the end of each clause. On the other hand, in English it comes at the beginning of the embedded clause. Japanese lacks relative pronouns or relative complementizers that can mark the left-edge of embedded clauses, while English has them (Fukui 1995; Kuroda 1988). Therefore, it seems that almost no reliable cues are provided for the beginning of embedded clauses that the parser can make use of.

Recently, Miyamoto (2002), following Inoue's (1991) insight, experimentally showed that Case Markers can induce clause boundaries in Japanese. Miyamoto argues that the second Nominative NP in (4) can mark the onset of the embedded clause. A Nominative NP typically marks the onset of a tensed clause. From this fact, it follows that the second occurrence of the Nominative NP implies the presence of a tensed clause, and therefore it can tell the reader that an embedded clause begins at this position.[26]

[26] Note that the lexical semantics of Nominative NPs may affect the clause boundary induction. See Muraoka and Sakamoto 2003 for a detailed discussion on this point.

(4) Obasan-**ga** [$_{RC}$ yoboyobo-no toshiyori-**ga** guuzen-ni kousaten-de mita]
Woman-**Nom** feeble old-man-**Nom** by chance intersection-Loc saw
onnanoko-ni isoide koe-o kaketa.
girl-Dat hurry called.
"The woman hurriedly called the girl who the feeble old-man saw at the intersection by chance."

Miyamoto's study shows that sometimes the parser can detect the beginning of an embedded clause. However, as far as relative clauses are concerned, it does not help much.

A number of previous studies report that a sequence of NPs is likely to be interpreted as the arguments of a single verb whenever possible (Inoue 1991; Mazuka and Lust 1988, 1990; Mazuka and Itoh 1995; Miyamoto 2002, 2003; Yamashita et al. 1993; Yamashita 1995). For example, in (5a) the parser initially processes the three NPs with different Case makers as being associated with a single verb, in this case "ageta," *gave*, since the sequence of NPs, [NP-Nominative, NP-Dative, NP-Accusative], is typically associated with a ditransitive verb. Because of this preference, if the verb's argument structure is not compatible with those NPs and the structure turns out to be a relative clause, a garden path effect arises because the parser has to reanalyze the sentence structure as illustrated in (1a).

(5) a. Marco-ga Kitty-ni ringo-o ageta.
Marco-Nom Kitty-Dat apple-Acc gave
"Marco gave an apple to Kitty"

a. Marco-ga Kitty-ni ringo-o tabeta inu-o miseta.
Marco-Nom Kitty-Dat apple-Acc ate dog-Acc show
"Marco showed the dog that ate an apple to Kitty."
(Inoue 1991; Mazuka and Itoh 1995)

The situation is similar even if the cues for the beginning of the embedded clause are provided by Case markers. The second occurrence of a Nominative NP can induce a clause boundary. However, a Nominative NP is equally compatible with a complement clause and a relative clause. Based on the observations in the literature it is plausible that the parser's preference for interpreting NPs as arguments of a single verb forces the complement clause analysis rather than the relative clause analysis even though the beginning of the embedded clause is signaled. What we can conclude from the discussion so far is that the relative clause structure is one of the least preferable structures for the parser's initial analyses. Moreover, it seems to be clear that the parser cannot normally construct a relative clause structure until crucial information such as the head of the relative clause becomes available.

So far, ways of unambiguously marking the left-edge of embedded clauses have not been reported in the literature. Here we show that classifiers in Japanese have the potential to provide an unambiguous cue for an upcoming relative clause. In an example like (6), a numeral classifier bearing Genitive Case, which we will refer

to as a Genitive Numeral Classifier, must be unambiguously associated with the relative head NP, "hon" [book]. Although there is an NP, “gakusee” [student] adjacent to the classifier, it cannot be associated with the classifier because of the grammatical constraints on Genitive Numeral Classifiers that we will review shortly.

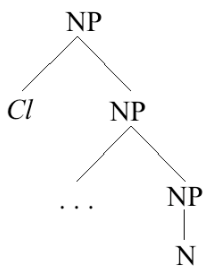
(6) 3-satu-no gakusee-ga yonda hon.
3-classifier (printed matter) student-Nom read book
“3 books that the student read.”

This type of numeral classifier has the following general grammatical requirements.

(7) a. A Genitive Numeral Classifier must be associated with a structurally adjacent Host NP.
b. A Numeral Classifier and its host NP must be semantically compatible.

The structure of a noun phrase containing a Genitive Numeral Classifier that can satisfy the above requirements is something like (8). In (8) the Genitive Numeral Classifier is left adjoined to the NP that it modifies.

(8)



This analysis implies that as long as the structural adjacency condition is satisfied, any modifiers or arguments of the head noun can intervene linearly between the Genitive Numeral Classifier and the host NP. As we can see in the examples in (9), any modifiers or arguments of the head noun including a relative clause can indeed intervene between a Genitive Numeral Classifier and its host NP without changing the relation between them.

(9) a. Inserting an adjective

[$_{NP}$ 3-satu-no [$_{NP}$ nagai hon]]

3-cl(printed matter)-Gen long book

"Three long books"

b. Inserting a possessive NP

[$_{NP}$ 3-satu-no [$_{NP}$ Marco-no hon]]

3-cl(printed matter)-Gen Marco-Gen book

"Marco's three books"

c. Inserting a modifier

[$_{NP}$ 3-satu-no [$_{NP}$ gengogaku-no hon]]

3-cl(printed matter)-Gen linguistics-Gen book

"Three books about linguistics"

d. Inserting modifiers and arguments

[$_{NP}$ 3-satu-no [$_{NP}$ Marco-no nagai gengogaku-no hon]]

3-cl(printed matter)-Gen Marco-Gen long linguistics-Gen book

"Marco's three long books about linguistics"

e. Inserting a relative clause

[$_{NP}$ 3-satu-no [$_{NP}$[$_{RC}$ Marco-ga yonda] hon]]

3-cl(printed matter)-Gen Marco-Nom read book

"Three books that the student read."

Now let us see the following examples. Both of the examples in 0) are acceptable sentences. In (10a) the numeral classifier is associated with the subject NP of the relative clause. On the other hand, in 0b) the classifier must be associated with the relative head NP because the linearly adjacent NP, the subject of the relative clause, is semantically incompatible. The only possible host NP is the NP headed by the relative head noun, i.e., the whole NP containing the relative clause. The structures of each example are in illustrated in Fig. 1.

(10) a. 3-nin-no gakusee-ga Yonda Hon

3-cl(human)-Gen student-Nom Read Book

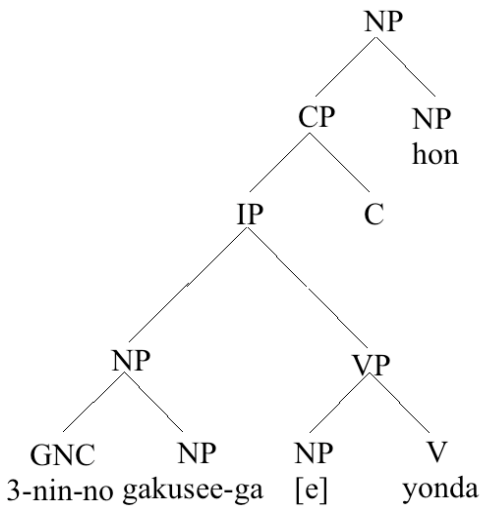
"The book that the three students read"

b. 3-satu-no gakusee-ga Yonda Hon

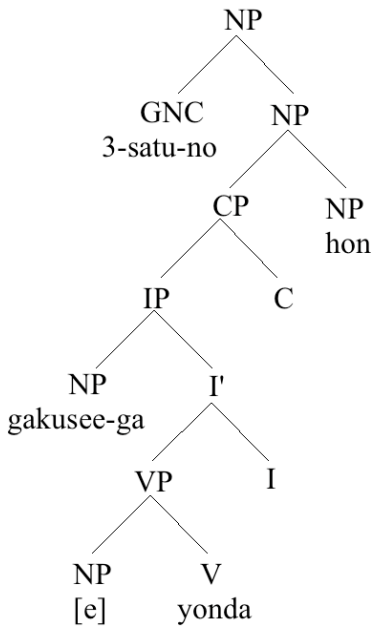
3-cl(printed matter)-Gen student-Nom Read Book

"The three books that the student read"

Fig. 1. A.



B.



From the discussion so far, it should be clear that when the classifier semantically mismatches with an immediately following nominative NP the only possible structure is a relative clause. Thus, if the parser can notice that a mismatching Classifier can be licensed only in a relative clause structure, it may be possible to predict an upcoming relative clause. However, here the very problem that we have discussed earlier arises, namely, how specifically the relative clause structure can be predicted. The examples that we have examined show that Numeral Classifiers do not have any lexical requirement to have a relative clause as an intervening element. As we have seen, the intervening element can be other modifiers or arguments of the associated noun. The only requirement that the Genitive Numeral Classifier has is that it must be licensed by a structurally adjacent Noun Phrase that is semantically compatible. In short, although the mismatching Genitive Numeral Classifier may provide a cue for an upcoming relative clause structure, there is a gap

between the information available from the mismatch between the genitive numeral classifier and its adjacent NP and the actual structure of the relative clause that the parser has to build. If the parser can fill this gap and can project the upcoming relative clause structure based on the cue, it means that the parser can dynamically exploit the relevant grammatical information without using the information from the heads.

In what follows, we will report the results from three experiments. The results of the experiments will show that the parser makes use of the indirect cue from the mismatching numeral classifier and creates a prediction upon encountering a mismatch between a Classifier and a Nominative NP. These results suggest that the parser is strongly incremental and that the parser is equipped with a powerful predictive mechanism.

The first experiment, an off-line sentence fragment completion experiment, is designed to answer the following two questions. The first question is whether native speakers of Japanese generate a relative clause structure when they are provided with mismatching numeral classifiers. If the mismatching numeral classifier can provide a cue for an upcoming relative clause, then Japanese speakers should complete sentence fragments with relative clause structures. If, on the other hand, the mismatching numeral classifier is not used as a cue for an upcoming relative clause structure, Japanese speakers should not complete the sentence fragments with a relative clause, rather, they should complete sentences with complement clauses following their general tendencies.

5. Experiment 1

An off-line sentence fragment completion test was conducted as an initial test of whether Japanese speakers are able to use numeral classifiers as a cue to generate relative clauses, using a task without time restrictions. Participants were presented with sentence fragments consisting of a sequence of 3 NPs in a paper-and-pencil task and asked to write completions for the sentences that seemed natural to them. The experiment manipulated two factors in a 2 x 3 factorial design. The first dependent variable that we are concerned with is the semantic compatibility between the numeral classifiers and the adjacent nominative NPs, in order to assess whether speakers would recognize a mismatching classifier-noun sequence as a cue for a relative clause boundary.

Participants

120 Japanese speakers participated in the experiment, all of whom were students at Shizuoka University or Shizuoka Sangyo University, Shizuoka, Japan or at Meiji Gakuin University, Tokyo, Japan. All gave informed consent and were paid the equivalent of $5.00 for their participation in the experiment, which lasted about 30 minutes.

Design

The experiment followed a 2 x 3 factorial design, which manipulated the match between a genitive numeral classifier and a following nominative NP (match vs. mismatch) and the case-marking and position of a wh-phrase that appeared earlier

in the fragment than the (mis-)matching NP (initial-dative vs. medial-dative vs. initial-nominative). The wh-phrase manipulation was included in order to investigate the impact of classifier mismatches on the processing of filler-gap dependencies, an issue that is more relevant to the later discussion on the island effect of relative clauses, and thus is not discussed in detail at this point. A sample set of experimental conditions is shown in Table 1.

Table 1 Sample Set of Experimental Conditions for Experiment 1

a. Classifier match/initial dative wh-phrase

Dono-seeto-ni	tannin-wa	3-nin-no	tosioita	sensee-ga
which -student-Dat	*class-teacher-Top*	*3-Cl(human)-Gen*	*aged*	*teacher-Nom*

b. Classifier mismatch/initial dative wh-phrase

Dono-seeto-ni	tannin-wa	3-satu-no	tosioita	sensee-ga
which -student-Dat	*class-teacher-Top*	*3-Cl(book)-Gen*	*aged*	*teacher-Nom*

c. Classifier match/medial dative wh-phrase

Tannin-wa	dono-seeto-ni	3-nin-no	tosioita	sensee-ga
class-teacher-Top	*which -student-Dat*	*3-Cl(human)-Gen*	*aged*	*teacher-Nom*

d. Classifier mismatch/medial dative wh-phrase

Tannin-wa	dono-seeto-ni	3-satu-no	tosioita	sensee-ga
class-teacher-Top	*which -student-Dat*	*3-Cl(book)-Gen*	*aged*	*teacher-Nom*

e. Classifier match/initial nominative wh-phrase

Dono-seeto-ga	tannin-ni	3-nin-no	tosioita	sensee-ga
which -student-Nom	*class-teacher-Dat*	*3-Cl(human)-Gen*	*aged*	*teacher-Nom*

f. Classifier mismatch/initial nominative wh-phrase

Dono-seeto-ga	tannin-ni	3-satu-no	tosioita	sensee-ga
which -student-Nom	*class-teacher-Dat*	*3-Cl(book)-Gen*	*aged*	*teacher-Nom*

The experimental materials consisted of 18 sets of 6 conditions distributed among six lists in a Latin Square design. Each participant saw exactly one of the lists of 36 target items intermixed with thirty-six filler items in a random order. The filler items included a variety of forms of numeral classifiers and wh-phrases, and were thus intended to prevent participants from noticing the regularity in the form of the

target items. The length of the fragments and the anticipated complexity of the completions was matched across target and filler items.

Some notes on the materials are in order. First, in this and all subsequent experiments in this article, the critical nominative embedded subject NP always denoted a human. The match between this NP and the preceding classifier was varied by manipulating the animacy of the classifier. The matching conditions used the classifier for humans *–nin*, and the mismatching conditions used classifiers for inanimate objects such as books. It was necessary for the mismatching classifiers to be inanimate, in order to avoid a potential ambiguity associated with mismatching genitive numeral classifiers that denote humans. As shown in (11), an animate classifier may sometimes be understood as referring to the possessor of a following inanimate NP. Japanese speakers find this interpretation easier to obtain when the genitive classifier is followed by a demonstrative article such as *ano* 'that', but it is at least marginally available even without the demonstrative. In contrast, since inanimate objects are highly implausible possessors, the same ambiguity does not arise with inanimate classifiers. Accordingly, (12) is judged to allow only a reading where the classifier is construed with the head of the relative clause. Therefore, the mismatching classifiers were always inanimate, in order to preserve the status of the classifier mismatch as a cue for a relative clause. Second, in order to maintain the naturalness of the fragments, we inserted adjectives between the classifiers and the nominative embedded subject NPs. All the adjectives can modify only these embedded subjects, and thus semantically they are compatible only with animate NPs.

(11) 3-nin-no hon-ga hihan-sita gengogakusha

3-$Cl_{(human)}$-Gen book-Nom criticized linguist

Interpretation 1: A linguist who the book(s) by the three authors criticized.

Interpretation 2: Three linguist who a book/book(s) criticized.

(12) 3-satu-no gengogakusha-ga hihansita hon

3-$cl_{(book\ like\ things)}$-Gen linguist-Nom criticized book

Interpretation1: Three books that the linguist criticized.

Interpretation2: #The book that the linguists in the three books criticized.

Results

This study yielded a total of 2082 codable sentence fragment completions. Table 1 summarizes the effect of the classifier match factor on the frequency of production of relative clauses in the fragment completions. Table 2 summarizes the counts and percentages of the types of completions across conditions. The impact of both the classifier match factor and the wh-phrase type factor on the production of filler-gap dependencies is presented later. A completion was classified as containing a 'matching relative clause' if it included a relative clause whose head was semantically compatible with the numeral classifier in the sentence-initial fragment. All other completions were classified as 'other'. This included both completions that contained no relative clause, and completions that contained a relative clause whose head did not semantically match the classifier.

Table 2 Rates of relative clause completions in Experiment 1.

	Construction Types				
	Matching Relative Clause		Other		Total
Conditions	N	%	N	%	
Classifier Match	1	0.1	1018	99.9	1019
Classifier Mismatch	851	80.05	212	19.95	1063
					2082

Table 3 Rates of relative clause completions in Experiment 1 across 6 conditions

	Construction Type				
	Matching Relative Clause		Other		
Conditions	N	%	N	%	total
GNC Matching/ Scrambling	0	0	314	100	314
GNC Mismatching/ Scrambling	221	70.1	94	29.8	315
GNC Matching/ Non Scrambling	1	0.2	375	99.7	376
GNC Mismatching/ Non Scrambling	335	87.9	46	12.07	381
GNC Matching/ Wh-Nominative	0	0	329	100	329
GNC Mismatching/ Wh-Nominative	295	80.3	72	19.6	367
					2082

As shown in Table 1, the match between the classifier and the following NP had a large impact upon the proportion of matching relative clauses that participants generated. In the conditions with mismatching classifiers 80% of completions contained a matching relative clause, whereas there was just a single instance (0.1%) generated in the conditions with matching classifiers. A χ^2 test showed that the proportion of relative clause completions was significantly different between the matching and mismatching conditions ($\chi^2(1)=1375.84$, $p < .01$). This pattern holds

true across all conditions. In the Matching Scrambling Condition there were no matching relative completions, in the Matching Non Scrambling Condition there was just one, and in the Matching Wh-Nominative Condition there were no matching relative clauses again. On the other hand, in all of the Mismatching Conditions there were many more matching relative clause completions. The Mismatching Scrambling Condition contained 70.1% matching relative clause completions, the Mismatching Non Scrambling Condition contained 87.9%, and the Mismatching Wh-Nominative Condition contained 80.3%. A χ^2 test showed that the proportion of relative clause completions was significantly different between the matching and mismatching conditions within each level (χ^2 (5)=1398.397, $p < .01$). This contrast shows very clearly that mismatching classifiers were effective in raising the proportion of relative clauses that were generated, but the 19.9% (212/1063) of trials in the classifier mismatch conditions that did not elicit a relative clause is also notable, since a relative clause should be required in order to complete these conditions in a grammatically acceptable fashion. These trials consisted of a combination of blanks, gibberish or ungrammatical completions. An example of such an ungrammatical completion is shown in (13). (13) is ungrammatical both because it lacks a host NP for the classifier and because it has only one predicate, despite having two subject NPs.

(13) Tannin-wa 3-satsu-no [tosioita sensee-ga ita-rasii].

class-teacher 3-$Cl_{book\text{-}like}$-gen aged teacher-Nom be-there-seems

"As for the class teacher, there seem to be ***three (book-like) old teachers***."

Discussion

The analysis of clause types in the sentence fragment completion data showed that the presence of a classifier mismatch dramatically increased the likelihood that a Japanese speaker would treat a subject NP as the subject of a relative clause. Mismatching classifiers elicited relative clause completions in over 80% of trials, indicating that despite the indirect nature of the cue that mismatching classifiers provide, they are nevertheless effective cues for relative clauses, at least when participants have unlimited time to complete the task. In the next experiment we proceed to the question of whether Japanese speakers are able to use the classifier mismatch cue in the more time-sensitive environment of an on-line task.

6. Experiment 2

Building upon the finding in Experiment 1 that a mismatch between a classifier and an adjacent NP provides a strong cue for relative clauses in an off-line setting, Experiment 2 was designed to investigate whether classifier mismatches are effective cues in an on-line setting.

As discussed earlier, Japanese speakers show a general bias to parse embedded subject NPs as the subject of a complement clause. Previous studies have shown that this can lead to a mild-to-moderate garden path effect if the embedded clause is subsequently disambiguated as a relative clause. There are potentially three elements that can indicate the presence of a relative clause and therefore induce a garden path effect: one is the relative head NP after the embedded verb; another is the

gap before the verb, although this cue is compromised by the fact that Japanese relatively freely allows argument omission; the other is an embedded verb without the complementizer -*to*. In the literature, some researchers have concluded that the first two elements are reliable disambiguating cues (Inoue, 1991; Mazuka, 1995; Yamashita, 1995; Yamashita, 1993), although the effectiveness of the bare verb cue has not been extensively tested. In one study of simple clause/relative clause ambiguities (Mazuka & Itoh, 1995) it has been argued that the relative head NP provides the only reliable disambiguating information. On the other hand, studies of the gapped/gapless relative clause ambiguity (Yamashita, 1995, Yamashita et al., 1993), suggest that the gap in the argument position before the verb provides disambiguating information, although the effects are typically not observed until the relative head NP position. Finally, the third possibility has not been investigated in the literature, but it is also possible that a verb without the complementizer -*to* also works as a reliable disambiguator. In a relative clause the embedded verb cannot bear the complementizer -*to*. Thus, if the parser independently notices that a clause is an embedded clause based on such information as case markers, it is possible that the bare embedded verb may provide a reliable cue for the presence of a relative clause, and thus trigger the same type of garden path effect. We expect the same garden-path effect based on any of these three factors in on-line reading times for sentences containing a relative clause whose subject is preceded by a matching numeral classifier. On the other hand, if the subject of the relative clause is preceded by a mismatching numeral classifier, and if the classifier mismatch is an effective cue in an on-line task, then we predict a different pattern. The NP that mismatches the

classifier should be unexpected and should lead to an initial slowdown in reading-times, but if the mismatch is then successfully used to recognize a relative clause boundary, then reading-times should be *faster* at the clause-final regions where the relative clause is normally disambiguated, since the garden-path effect should be avoided. On the other hand, if Japanese speakers are unable to exploit the classifier mismatch cue on-line or if it delays structure building until it reaches disambiguating information, then the classifier mismatch should not lead to facilitation in reading times at the end of the relative clause.

Participants

Sixty-three native speakers of Japanese participated in the experiment. All were students at Shizuoka University, Shizuoka Sangyo University, Shizuoka, Japan, or Meiji Gakuin University, Tokyo, Japan. They gave informed consent and were paid the equivalent of $5 for their participation in the experiment, which lasted about 30 minutes.

Materials and design

Twenty-four sets of two conditions each were used in the experiment. Both conditions contained an embedded relative clause, but varied with regard to the semantic compatibility between the subject of the relative clause and a preceding numeral classifier. In both conditions a main clause subject marked with the topic marker *–wa* was followed by a genitive numeral classifier, which in turn was followed by an adjective and a nominative NP that indicated the start of an embedded

clause. In the classifier match condition the classifier and the following NP were semantically compatible, and in the classifier mismatch condition the classifier was semantically incompatible with the following NP and was only compatible with the head of the relative clause that appeared a number of regions later. The head of the relative clause always corresponded to the direct object of the embedded verb, and thus the relative clause verb always appeared without an overt direct object. The relative clause verb was always a verb that canonically appears with both a direct and indirect object, such as a ditransitive verb like *ageru* 'give', *watasu* 'pass', or a causative verb. The relative head position and the matrix verb position were separated by two elements, an indirect object and a locative adverbial phrase. This additional material was inserted in order to separate any effects on reading times created at the relative head position from the reading time slowdown typically observed at the end of a sentence, i.e., the so-called wrap-up effect. For this manipulation, we used a verb that is canonically associated with two objects as the matrix verb.

The twenty-four pairs of items were distributed between two lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with seventy-two filler items in a random order. The filler items were matched with the target items in overall length and complexity, and were counterbalanced with regard to the distribution of genitive and other numeral classifiers in main and embedded clauses. A sample set of experimental items is shown in (14).

(14) a. Classifier Match Condition

1	2	3	4	5	6
Tannin-wa	san-nin-no	tosioita	sensee-ga	Atarasii	koochoo-ni
Class-teacher-Top	*three-cl(human)-Gen*	*Aged*	*teacher-Nom*	*New*	*president-Dat*

7	8	9	10	11	12
yorokonde	Okutta	hon-o	aru-seeto-ni	kyoositu-de	yomase-masita.
Gladly	*Gave*	*book-Acc*	*a-student-Dat*	*class-room-at*	*made-read.*

"The teacher made a student read the book that three aged teachers gladly gave to the new president at the classroom."

b. Classifier Mismatch Condition

1	2	3	4	5	6
Tannin-wa	san-satu-no	tosioita	sensee-ga	atarasii	koochoo-ni
Class-teacher-Top	*three-cl(book)-Gen*	*Aged*	*teacher-Nom*	*New*	*president-Dat*

7	8	9	10	11	12
yorokonde	Okutta	hon-o	aru-seeto-ni	kyoositu-de	yomase-masita
gladly	*Gave*	*book-Acc*	*a-student-Dat*	*class-room-at*	*made-read.*

"The teacher made a student read three books that an aged teacher gladly gave to the new president at the classroom."

Procedure

The experiment was conducted on Dell laptop computers running the *Linger* software developed at MIT (Rohde 2001-2003). Participants were timed in a phrase-by-phrase self-paced non-cumulative moving window reading task (Just et al. 1982). Sentences were presented using Japanese characters with the font MS Gothic 16 point, with target items segmented according to the scheme in (14). All sentences,

including the filler items, were presented on a single line. Stimulus segments initially appeared as a row of dashes, and participants pressed the space bar of the keyboard to reveal each subsequent region of the sentence.

In order to ensure that participants attended to the stimuli, an argument-verb matching task was presented after each trial. It was not practical to ask yes/no comprehension questions, since many of the experimental sentences could themselves be understood as questions. Following each trial, a verb was displayed on the computer screen followed by two NPs, corresponding to NPs from the experimental sentence, and participants had to indicate which of the NPs was the subject of the verb in the sentence just read by pressing one of two keys on the keyboard. In the comprehension task, the two NPs were displayed without case markers, in order to exclude the possibility of answering the question based on the case markers on the NPs. This task was adopted from Nagata (1993) and Aoshima et al. (2004) with a slight modification. In the previous studies only subject NPs were used in this task, but in our experiment both subject NPs and non-subject NPs were presented. This modification was adopted both in order to encourage participants to attend to all words in a sentence, and also because some of the filler sentences had quantificational wh-phrase subject NPs, which were not suitable for the comprehension task.

In order to familiarize participants with the subject-verb matching comprehension task, an off-line practice session was included before the experiment. In this practice session, only three instances of numeral classifiers were included among the twenty practice sentences.

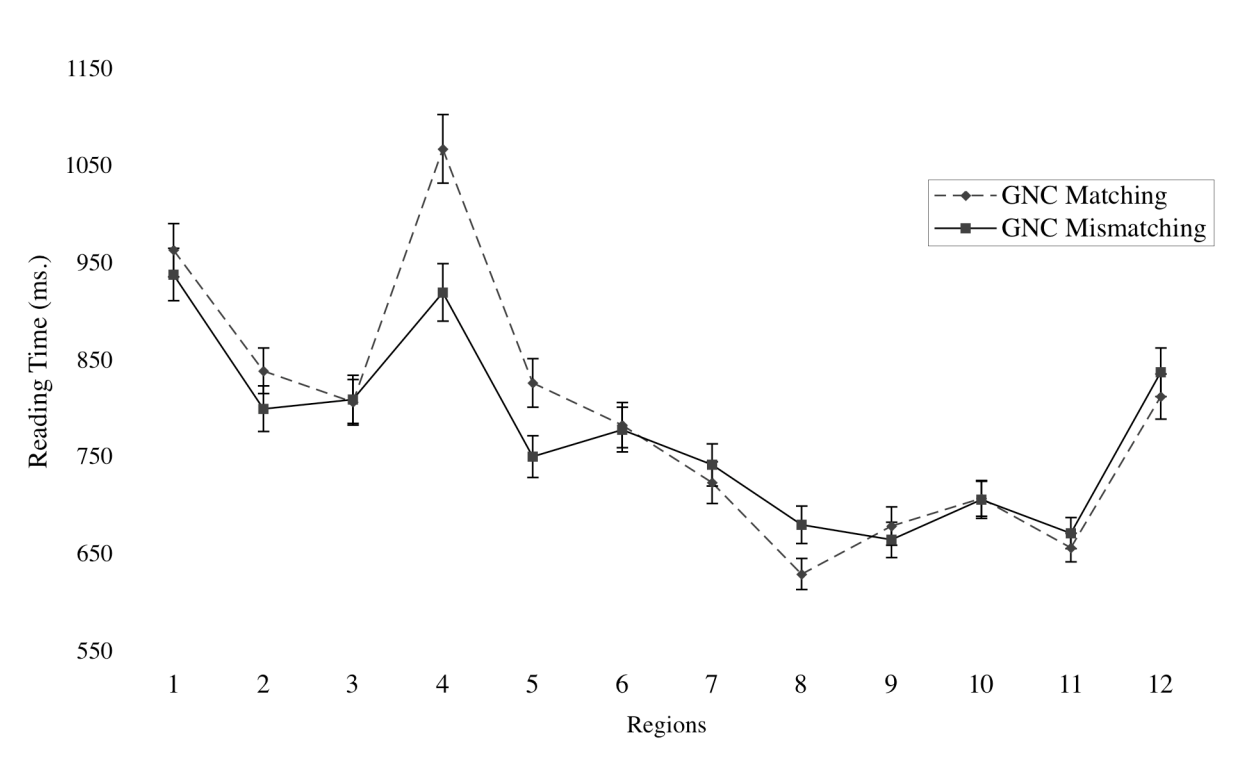
Results

Comprehension accuracy and reading times at each region were analyzed using a repeated-measures ANOVA, with semantic compatibility between numeral classifiers and the adjacent subject NP (match vs. mismatch) as a within-subjects factor. All data from participants whose comprehension task accuracy was below 70% in total were discarded (n = 9, 16 %). Reading times longer than 3000ms were discarded, affecting 1% of trials. The means and analyses presented below are based on the remaining trials.

The average comprehension accuracy among the fifty-four participants who were included in the analysis was 82.5%. Mean accuracy scores did not differ significantly between the two conditions. (Fs < 1).

Average reading times for each region are summarized in Figure 1. There were significant differences in reading times between the two conditions at Region 4 (embedded subject NP), Region 5 (adjective), and Region 8 (embedded verb), with a reversal in the pattern of difficulty between Regions 4-5 and Region 8. There were no significant differences at any other regions (all Fs < 1).

Fig. 2. Reading times per region, Experiment 2.



At the embedded subject NP in Region 4 there was a slowdown in the classifier mismatch condition ($F1(1,39) = 15.3$, MSe = 4897196, $p < .01$; $F2(1,23) = 4.10$, MSe = 3113357, $p < .05$). This effect continued to the following adjective in Region 5, where the slowdown was significant in the participants analysis but not in the items analysis ($F1(1,39) = 5.05$, MSe = 1207737, $p < .05$; $F2(1,23) = 2.18$, MSe = 768905, $p = 0.15$). At the embedded verb in Region 8, on the other hand, reading times in the classifier match condition were on average 51 ms slower than in the classifier mismatch condition ($F1(1,39) = 4.15$, MSe = 403055, $p < .05$; $F2(1,23) = 4.19$, MSe = 580423, $p < .05$). No other differences were significant.

Discussion

The aim of this study was to investigate whether classifier mismatches could serve as a cue for relative clause structures in an on-line setting, and accordingly whether the parser is able to build head-final relative clause structures incrementally, rather than deferring structure building until the end of the relative clause.

The slowdown at the embedded subject NP in the classifier mismatch condition, which persisted to the adjective in the following region, is unsurprising. A mismatching classifier-noun sequence is very rare in Japanese, and increased difficulty in reading times should be predicted relative to a matching classifier-noun sequence, irrespective of whether a predictive or a head-driven strategy is adopted for parsing NPs.

The more interesting finding of this study is the reversal in the reading-time patterns at the embedded verb in Region 8, where the classifier match condition was read more slowly than the classifier mismatch condition. The experimental conditions are identical except for the classifier manipulation in Region 2, so the reading-time difference at the embedded verb must be a consequence of the classifier match manipulation. The embedded verb provides two sources of disambiguating evidence for the relative clause structure. First, the lack of the complementizer suffix *–to* on the verb indicates that the verb is in a relative clause. Second, we consistently used verbs that select both a direct and an indirect object, and therefore the fact that the verb is not preceded by an accusative-marked direct object provides further support for a relative clause parse (cf. Yamashita, 1995; Yamashita et al, 1993). Of course, the lack of an overt direct object is not an unambiguous cue for a relative clause, since

Japanese allows null objects, typically in situations where the object is a highly accessible discourse referent. If the nominative NP in Region 4 is analyzed as the subject of a complement clause in the classifier match condition, as should be predicted based on previous results (e.g., Miyamoto, 2002; Aoshima et al., 2004), then a garden-path effect is predicted when the relative clause verb is encountered. The fact that the verb is read more quickly in the classifier mismatch condition suggests that the relative clause structure was not unexpected in this condition. This in turn suggests that the classifier mismatch was sufficient to initiate construction of a relative clause in the on-line setting.

The fact that reading times at the disambiguating embedded verb position were facilitated by the classifier mismatch is consistent with the predictions of an incremental structure building account of Japanese, in which the relative clause structure can be initiated before the embedded verb is reached. It is less easy to reconcile these results with a head-driven account, in which the presence of a head of a phrase is necessary for integrating its sub-constituents (Abney 1989; Pritchett 1991a, 1992b). Such an account would predict that construction of the embedded clause would be delayed until the embedded verb was processed, and would not readily predict facilitation at that point in the classifier mismatch condition.

The fact that Japanese speakers appear to be able to use the classifier mismatch cue to initiate building of relative clause structures begs the question of what parsing mechanism allows them to exploit this unambiguous yet highly indirect structural cue. This issue is taken up in more detail in the General Discussion section

below. First, however, we consider in more detail the content of the representation that Japanese speakers construct prior to the verb in the classifier mismatch condition.

The evidence from Experiment 2 indicates that a verb that disambiguates a relative clause structure is less surprising following a mismatching classifier-noun sequence, and we suggested that this reflects earlier initiation of the relative clause structure. However, there are a couple of limitations on this conclusion. First, since the evidence is based on a reading-time facilitation at the verb position, this provides only indirect evidence that the relative clause structure was initiated prior to the verb position. Second, the facilitation at the verb position in this experiment suggests that the relative clause verb is more easily integrated into the existing structure in the classifier mismatch condition than in the classifier match condition, but provides only limited information on the extent to which the existing structure in the classifier mismatch condition already has the structural properties of a relative clause. The next experiments address both of these limitations by investigating the impact of classifier-noun mismatches on the processing of filler-gap dependencies in Japanese, in particular whether mismatching classifiers activate the island property of relative clauses.

7. Relative Clause Islands in Japanese

Japanese, like English, allows long-distance filler-gap dependencies. For example, the dative-marked indirect object of the embedded verb *ageta* 'gave' canonically appears inside the embedded clause as in (15a), but may also appear in sentence-initial position as a result of a process known as scrambling (15b). Phrases

may also be scrambled to other positions between the sentence-initial position and the canonical position.

(15) a. [Kitty-wa [$_{CP}$ otokonoko-ga Marco$_1$-ni hon-o ageta-to] omotta].
Kitty-Top boy-Nom Marco-Dat book-Acc gave-Comp thought.

b. Marco$_1$-ni [Kitty-wa [$_{CP}$ otokonoko-ga ___$_1$ hon-o ageta-to] omotta].
Marco-Dat Kitty-Top boy-Nom book-Acc gave-Comp thought.
'Kitty thought that the boy gave the book to Marco.'

Scrambling may target both referential NPs and wh-phrases, as in (15) or in (16). Importantly, unlike in English, the position of a wh-phrase in Japanese does not indicate its scope. As shown in (17), direct and indirect questions are distinguished in Japanese by the presence of a question particle on the main verb or embedded verb, respectively, and not by the position of the wh-phrase.

(16) a. [Kitty-wa [$_{CP}$ otokonoko-ga dare$_1$-ni hon-o ageta-to] omotta-no]?
Kitty-Top boy-Nom who-Dat book-Acc gave-Comp thought-Q

b. Dare$_1$-ni [Kitty-wa [$_{CP}$ otokonoko-ga ___$_1$ hon-o ageta-to] omotta-no]?
Who-Dat Kitty-Top boy-Nom book-Acc gave-Comp thought-Q
'Who did Kitty think that the boy gave the book to?'

(17) a. [Kitty-wa Marco-ni [$_{CP}$ otokonoko-ga dare-ni hon-o ageta-ka] itta].
Kitty-Top Marco-Dat boy-Nom who-Dat book-Acc gave-Q said
"Kitty said to Marco who did the boy gave the book to."

b. [Kitty-wa Marco-ni [$_{CP}$ otokonoko-ga dare-ni hon-o ageta-to]

Kitty-Top Marco-Dat boy-Nom who-Dat book-Acc gave-Comp

itta-no]?

said-Q

"Who did Kitty say to Marco that the boy gave the book to?"

Scrambling, like wh-movement in English, can create potentially unbounded filler-gap dependencies. However, scrambling is also subject to a number of the island constraints on dependency formation that restrict wh-movement in English and other languages. In particular, both scrambling and wh-movement may create dependencies that span a complement clause boundary (18a), but may not create dependencies that span a relative clause boundary (18b) (Ross 1967; Saito 1985). Following standard linguistic terminology, relative clauses are known as 'islands' for scrambling.

(18) a. Marco$_1$-ni [Kitty-wa [otokonoko-ga __$_1$ hon-o ageta-to] itta].

Marco-Dat Kitty-Top boy-Nom book-Acc gave-Comp said

"Kitty said that the boy gave the book to Marco."

b.* Marco$_1$-ni [Kitty-wa [$_{RC}$ otokonoko-ga __$_1$ ageta] hon]-ga sukida].

Marco-Dat Kitty-Top boy-Nom gave book-Acc like.

'Kitty likes the book that the boy gave to Marco'

The islandhood of relative clauses is successfully captured in many different accounts of unbounded dependencies (Kaplan and Zaenen 1989; Kroch and Joshi 1985; Pollard and Sag 1994; Steedman 1996 among many others), and the choice among these accounts does not matter for the argument in this paper. However, as an example, Chomsky's well-known *subjacency condition* (Chomsky 1973) states that filler-gap dependencies may not span more than one bounding node, where the bounding nodes are assumed to be NP and S. A dependency that spans a relative clause boundary violates this constraint, because the relative clause contains an S category and the combination of the relative clause with its head forms an NP category, and thus at least two bounding nodes are crossed.

Whatever account of relative clause islands ultimately proves to be correct, the relevant point for our current concerns is that the islandhood of relative clauses depends on the specific structural properties of relative clauses. Therefore, the islandhood of relative clauses can be used as a diagnostic of whether detailed relative clause structures are constructed following mismatching classifier-noun sequences. We next turn to a discussion of filler-gap dependency processing in Japanese, which will provide a measure of islandhood in Japanese.

8. Processing Filler-Gap Dependencies in Japanese

Much evidence in English and similar languages indicates that when the parser encounters a fronted phrase (i.e., a filler), there is a preference to associate the fronted phrase with the first potential gap position, leading to a general bias for shorter filler-gap dependencies. This bias has been confirmed using a variety of

different measures, including the *filled gap effect* in reading-time studies (Bourdages 1992; Crain & Fodor, 1985; Stowe, 1986), plausibility manipulations in eye-tracking studies (Traxler & Pickering, 1996), ERP studies (Garnsey, Tanenhaus, & Chapman, 1989) and speeded grammaticality judgment studies (McElree & Griffith, 1998). In the filled gap paradigm illustrated in (19), for example, readers exhibit a slowdown in reading times upon encountering an overt NP (e.g. *us*) in a position where a gap had been expected. Such effects are taken to indicate that the parser actively posits a gap at the first possible position, without waiting for bottom-up evidence for an empty argument position (Crain & Fodor, 1985; Stowe, 1986).

(19) a. My brother wanted to know if Ruth will bring **us** home to Mom at Christmas.

b. My brother wanted to know who Ruth will bring us home to __ for breakfast?

The bias for shorter filler-gap dependencies observed in English could, in principle, be due to either of a couple of different sources. On the one hand, it may reflect a constraint that explicitly favors creating a gap as soon as possible after encountering a filler, irrespective of what other constraints are satisfied by positing that gap. Alternatively, the bias may reflect the fact that creation of a gap allows other linguistic constraints to be satisfied, such as confirmation of thematic role assignments, and thus the bias may reflect the parser's goal of satisfying these other constraints as soon as possible. Although it is difficult to distinguish these alternatives

using evidence from English, Aoshima et al. (2004) provide evidence from Japanese in favor of the second alternative, by showing that under certain circumstances Japanese speakers favor *longer* filler-gap dependencies.

A dative wh-phrase that appears at the beginning of a two-clause sentence in Japanese clearly does not occupy its canonical position, and is temporarily ambiguous between one of two possible gap positions. It may be associated with a main clause gap (20a), which would allow for early completion of the filler-gap dependency, but would delay confirmation of a thematic role until the final word of the sentence, due to the verb-final property of Japanese. Alternatively, it may be associated with an embedded clause gap (20b), which would delay completion of the filler-gap dependency, but would allow for earlier confirmation of the filler's thematic role, and would make it possible for the question particle to appear on the embedded clause verb. A question particle in Japanese must appear at least as high in the sentence structure as the thematic position of a wh-phrase that it is associated with.

(20) a. Dare$_1$-ni [$_{CP}$ Kitty-wa __$_1$ [$_{CP}$ Marco-ga sono-hon-o ageta-to] itta-no]?
who-Dat Kitty-Top Marco-Nom that-book-Acc gave-Comp said-Q
"To whom did Kitty say that Marco gave that book to someone?"

b. Dare$_1$-ni [$_{CP}$ Kitty-wa [$_{CP}$ Marco-ga __$_1$ sono-hon-o ageta -to] itta-no]?
who-Dat Kitty-Top Marco-Nom that-book-Acc gave-Comp said-Q
"Who did Kitty say that Marco gave the book to ?"

Aoshima et al. (2004) show using a number of different measures that Japanese speakers prefer to associate fronted wh-phrases with a gap in the embedded clause. This suggests that the driving force behind filler-gap dependency formation is the satisfaction of grammatical requirements, rather than creation of a gap as an end in itself. One of the measures used to demonstrate this bias, which we draw upon in the current study, is a Japanese counterpart of the Filled Gap Effect paradigm.

The examples in (21a-b) both contain two dative-marked NPs, one in the main clause and a second in the embedded clause. It is not unnatural for a Japanese sentence to contain two dative NPs. What is less natural is for a Japanese sentence to include two dative NPs within a single clause. Therefore, the logic of Aoshima and colleagues' Experiment 2 is that if Japanese speakers try to interpret the fronted dative wh-phrase in (21a) as an argument of the embedded clause, and if they do so as soon as they encounter the subject NP that marks the onset of the embedded clause, then they should be surprised to encounter a second, overt dative NP inside the embedded clause (i.e., *katyoo-ni*, 'assistant'). This is the counterpart of the English Filled Gap Effect observed when readers encounter an overt NP in a position where they had already posited a gap (Crain & Fodor, 1985; Stowe, 1986). The Japanese Filled Gap Effect would be seen as a slowdown in reading times at the embedded dative NP in (21a), relative to the same region in (21b), in which the first dative NP (i.e., *senmu-ni*, 'managing director') is not unambiguously scrambled, and is therefore likely to be interpreted as a main clause argument (Kamide & Mitchell, 1999). Using a self-paced reading paradigm, Aoshima and colleagues observed a significant Filled Gap Effect at the second dative NP in (21b), providing further support for the bias for

longer filler-gap dependencies in Japanese, and also providing evidence that filler-gap dependencies are constructed in advance of the verb in Japanese.

(21) a. Scrambled Condition

Dono-syain-ni	senmu-wa	syatyoo-ga	**katyoo-ni**	syookyuu-o	yakusokusita-to	iimasita-ka
Which employee-Dat	managing director-Top	president-Nom	assistant manager-Dat	raise-Acc	promised-DeclC	told-Q

'To which employee did the managing director tell that the president promised a raise to the assistant manager?'

b. Control Condition

Dono-syain-ga	senmu-ni	syatyoo-ga	**katyoo-ni**	syookyuu-o	yakusokusita-to	iimasita-ka
Which employee-Nom	managing director-Dat	President-Nom	assistant manager-Dat	raise-Acc	promised-DeclC	told-Q

'Which employee told the managing director that the president promised a raise to the assistant manager?'

9. Filler-gap Dependencies and Relative Clause Islands

There is a potential conflict between the islandhood of Japanese relative clauses and the finding that Japanese speakers favor longer filler-gap dependencies. Aoshima et al. (2004) demonstrated a bias for Japanese speakers to associate fronted ambiguous wh-phrases with an embedded clause gap, presumably because this allows for earlier satisfaction of thematic or scope requirements, and they showed that filler-gap dependencies are formed before the embedded clause verb is reached. In that study all of the embedded clauses were complement clauses, which freely allow long-distance scrambling, and thus all of the embedded clause gaps turned out to be grammatically acceptable. However, the bias for longer-distance filler-gap

dependencies creates a potential danger, since a speaker may form a long-distance dependency into an embedded clause that turns out to be a relative clause. Relative clauses are islands for scrambling (Saito, 1985), and a number of studies have shown that comprehenders avoid forming filler-gap dependencies that cross island boundaries in English (Stowe, 1986; Traxler & Pickering, 1996 among others). However, Japanese speakers may inadvertently construct filler-gap dependencies that violate an island constraint, because relative clauses typically cannot be identified until the end of the relative clause, after the point when filler-gap dependencies are constructed. In what follows, we investigate whether this occurs and, more importantly, whether the presence of a classifier-noun mismatch provides a sufficient cue to block the formation of a long-distance filler-gap dependency. Using the Japanese version of the Filled Gap Effect paradigm it should be possible to use the islandhood of relative clauses to test whether classifier mismatches activate sufficiently rich relative clause structure for inhibiting filler-gap dependency formation, and whether this structure is available prior to the relative clause verb.

We first investigate the interaction of classifier mismatches and filler-gap dependency formation using an off-line sentence fragment completion paradigm, and then turn to an on-line reading-time study using the Filled Gap Effect paradigm.

10. Experiment 1b

This study is an analysis of another dependent measure from the sentence fragment completion study described above in Experiment 1a. In Experiment 1a, we were concerned with the first dependent measure, namely an analysis of how

effective classifier-noun mismatches were as cues for relative clauses. In Experiment 1b we focus on how participants complete filler-gap dependencies, and on how this interacts with the generation of relative clauses, which are islands for long-distance dependency formation. Aoshima et al. (2004, Experiment 3) used a sentence fragment completion task as one measure of the bias for longer filler-gap dependencies in Japanese, in a study where all embedded clauses were potential complement clauses. In the current study we predict that this finding should be be replicated in cases where a matching classifier-noun sequence introduces an embedded clause, making it compatible with a complement clause analysis. In cases where the embedded clause is introduced by a mismatching classifier-noun sequence, however, it should be grammatically impossible to construct a long-distance filler-gap dependency that spans the relative clause boundary. The aim of this study is to determine whether Japanese speakers are sensitive to this constraint in an off-line generation task.

Participants, Materials and Design

The details of the experimental design are presented in Experiment 1a above. The participants were the same participants. Here we summarize the second dependent measure that we are concerned with. Materials consisted of 18 sets of six conditions, organized in a 2 x 3 factorial design, as shown in Table 4. In addition to the manipulation of the match between the embedded clause subject and a preceding classifier, we manipulated the form and position of a wh-phrase. In the initial dative condition a dative wh-phrase appeared in sentence initial position in the fragment. In this position it must obligatorily be understood as scrambled, and it is compatible

with either a main clause or an embedded clause gap. This corresponds to the condition that was most likely to yield an embedded clause gap in the studies by Aoshima et al. (2004). In the medial dative condition a dative wh-phrase appeared in second position in the fragment, between an initial topic-marked NP and the embedded subject NP. This phrase could, in principle, be interpreted either as a main clause in-situ argument or as a phrase that has undergone short-distance scrambling to the front of the embedded clause, but previous evidence suggests that speakers prefer the in-situ analysis (Kamide & Mitchell, 1999; Aoshima et al., 2004). Finally, in the initial nominative condition the fragment started with a nominative wh-phrase, which can only be interpreted as a main clause argument, due to the ban on scrambling of nominative wh-phrases in Japanese (Saito 1985).

Table 4 Sample Set of Experimental Conditions for Experiment 1

a. Classifier match/initial dative wh-phrase

Dono-seeto-ni	tannin-wa	3-nin-no	tosioita	sensee-ga
which -student-Dat	*class-teacher-Top*	*3-Cl(human)-Gen*	*Aged*	*teacher-Nom*

b. Classifier mismatch/initial dative wh-phrase

Dono-seeto-ni	tannin-wa	3-satu-no	tosioita	sensee-ga
which -student-Dat	*class-teacher-Top*	*3-Cl(book)-Gen*	*Aged*	*teacher-Nom*

c. Classifier match/medial dative wh-phrase

Tannin-wa	dono-seeto-ni	3-nin-no	tosioita	sensee-ga
class-teacher-Top	*which -student-Dat*	*3-Cl(human)-Gen*	*Aged*	*teacher-Nom*

d. Classifier mismatch/medial dative wh-phrase

Tannin-wa	dono-seeto-ni	3-satu-no	tosioita	sensee-ga
class-teacher-Top	*which -student-Dat*	*3-Cl(book)-Gen*	*Aged*	*teacher-Nom*

e. Classifier match/initial nominative wh-phrase

Dono-seeto-ga	tannin-ni	3-nin-no	tosioita	sensee-ga
which -student-Nom	*class-teacher-Dat*	*3-Cl(human)-Gen*	*Aged*	*teacher-Nom*

f. Classifier mismatch/initial nominative wh-phrase

Dono-seeto-ga	tannin-ni	3-satu-no	tosioita	sensee-ga
which -student-Nom	*class-teacher-Dat*	*3-Cl(book)-Gen*	*Aged*	*teacher-Nom*

Following Aoshima et al. (2004) sentence fragment completions were analyzed for two measures of where the wh-phrase was interpreted. First, we analyzed the distribution of question particles in the completions. All wh-phrases must be associated with a question particle that marks the scope of the question, as either a direct or an indirect question. If the question particle appears on the embedded verb, then the wh-phrase must be associated with an embedded clause gap. The converse is not true, however, as a main clause question particle may be associated with either a main clause or an embedded clause gap. Second, we analyzed the distribution in the completions of ditransitive verbs that obligatorily select a dative-marked NP. In all sentence fragments a single dative NP appeared overtly in the main clause. Although it was always possible to generate an embedded ditransitive verb in the completion, in no condition was this a grammatical requirement. Therefore, if any condition shows an increased proportion of embedded ditransitive verbs, this likely reflects an increased number of dative NPs that were interpreted as being associated with an embedded clause gap. If fronted dative wh-phrases are preferentially associated with an embedded clause gap, then we expect to

find increased numbers of completions with embedded ditransitive verbs in the initial dative/classifier match condition. However, if participants respect the islandhood of relative clauses, then the number of embedded ditransitive verbs should not be increased in the initial dative/classifier mismatch condition.

Results

First, let us examine the number of Q-particles in the embedded clauses and the matrix clauses. We excluded from the analysis the fragments that were not completed, and completions containing Q-particles in both the embedded and matrix clause. The results are summarized in Table 5.

Table 5 Summary of Question Type

	Question Type					
	Embedded			Main		
Conditions	N	%	Count	N	%	Count
GNC Matching	221	23.7	931	591	63.4	931
GNC Mismatching	18	1.7	1054	992	94.1	1054
Total	239	12.04	1985	1583	79.7	1985

In the analysis of question-type, 12.04 % of fragments (239 trials) were completed as indirect questions, with a question marker on an embedded verb only. The proportion of trials in which a question particle was provided on the embedded clause verb was 23.7% for the Matching Conditions (221/931 trials), and 1.7% for the Mismatching

Conditions (18/1054 trials). A χ^2 test showed that the proportion of embedded Q-particle completions was significantly different between the Matching and the Mismatching Conditions ($\chi^2(1) = 225.21$, $p < .01$).

A summary of question-type within each condition is in Table 6.

Table 6 Summary of Question-type

	Question-type					
	Embedded			Main		
Conditions	N	%	Count	N	%	Count
GNC Matching/ Scrambling	73	22.8	319	249	63.9	319
GNC Mismatching/ Scrambling	3	0.84	354	334	94.9	354
GNC Matching/ Non Scrambling	119	38.6	308	204	44.8	308
GNC Mismatching/ Non Scrambling	10	2.8	349	336	94.2	349
GNC Matching/ Wh-Nominative	29	9.5	304	138	81.9	304
GNC Mismatching/ Wh-Nominative	5	1.4	351	322	95.1	351
Total	239	12.04	1985	1583	79.7	1985

The proportion of trials in which a question particle was provided on the embedded clause verb was 22.8% (73/319 trials) for the Matching/Scrambling condition, 0.84% (3/354 trials) for the Mismatching/Scrambling condition, 36.8% (119/308 trials) for

the Matching/Non Scrambling condition, 2.8% (9/349 trials) for the Mismatching/Non Scrambling Condition, 9.5% (29/304 trials) for the Matching/Nominative Wh condition and 1.4% (5/351 trials) for the Mismatching/ Nominative Wh condition. χ^2 tests showed that the proportion was significantly different between Scrambling, Non Scrambling and Wh-Nominative conditions within the Matching condition (χ^2 (2)=26.64, p<.01). Pairwise comparison showed that the proportion of Q-particles on embedded verbs was higher in (i), the Non Scrambling condition than in the Scrambling condition (χ^2 (1)=15.48, p<.01 and in (ii), in the Non Scrambling condition than in the Wh-nominative condition (χ^2 (1)=19.80, p<.01). There was no significant difference among conditions within GNC Mismatching conditions.

The analyses of verb argument structure are shown in Table 7. I counted the verbs that can take dative NPs as their arguments but not as benefactive phrases.

Table 7 Summary of Verb Argument Structure

	Dative Argument Structure					
	Embedded			Main		
Conditions	N	%	Count	N	%	Count
GNC Matching	219	23.5	931	793	85.1	931
GNC Mismatching	84	7.96	1054	1015	96.2	1054
Total	303	15.2	1985	1808	91.08	1985

In the analysis of verb argument structure, 15.3% of fragment completions (303 trials) contained an embedded ditransitive verb that obligatory takes a dative argument. The

proportion of trials in which a ditransitive verb was provided in the embedded clause was 23.5% for the GNC Matching Conditions (219/931 trials), and 7.96% for the GNC Mismatching Conditions (84/1054 trials). A χ^2 test showed that the proportion of embedded ditransitive verb completions was significantly different between the GNC Matching and the GNC Mismatching Conditions ($\chi^2(1) = 93.67$, $p < .01$).

In the analysis of verb argument structure, trials in which a Q-particle was provided on the embedded verb only were separated from the trials in which a Q-particle was provided on the main clause verb. This reflects the general fact that a Wh-NP must be assigned its thematic role and licensed by a Q-particle.

The results of verb argument structure with embedded Q-particles only are summarized in Table 8.

Table 8 Dative Argument Structure: Embedded Q-particle Only

	Dative Argument Structure (Embedded Q-particle only)					
	Embedded			Matrix		
Conditions	N	%	Count	N	%	Count
GNC Matching/ Scrambling	43	40.1	107	43	58.9	73
GNC Mismatching/ Scrambling	3	60	5	2	66.6	3
GNC Matching/ Non Scrambling	59	41.5	142	72	60.5	119
GNC Mismatching/ Non Scrambling	2	12.5	16	6	60	10
GNC Matching/ Wh-Nominative	16	26.6	60	19	65.5	29
GNC Mismatching/ Wh-Nominative	3	50	6	4	80	5
Total	126	37.5	336	146	61.08	239

The proportion of embedded dative verbs was 40.1% (43/107 trial) for the Matching/Scrambling condition, 60% (3/5) for the Mismatching/Scrambling Condition, 41.5% (59/142) for the Matching/Non-Scrambling condition, 12.5% (2/16) for the Mismatching/Non-Scrambling condition, 26.6% (16/60 trials) for the Matching/Wh-Nominative condition, and 50% (3/6) for the GNC Mismatching/Wh-

Nominative condition. χ^2 tests showed that none of the differences between conditions are reliable.

The results of dative argument structure with matrix Q-particle only are summarized in Table 9.

Table 9 Dative Argument Structure: Matrix Q-particle Only

	Dative Argument Structure (Matrix Q-particle only)					
	Embedded			Matrix		
Conditions	N	%	Count	N	%	Count
GNC Matching/ Scrambling	31	15.1	204	233	97.8	238
GNC Mismatching/ Scrambling	29	8.6	336	334	98.8	338
GNC Matching/ Non Scrambling	21	15.2	138	158	98.1	161
GNC Mismatching/ Non Scrambling	15	4.6	322	322	98.1	328
GNC Matching/ Wh-Nominative	35	14.0	249	273	97.5	280
GNC Mismatching/ Wh-Nominative	26	7.7	334	325	97.0	335
Total	157	9.91	1583	1645	97.9	1680

The proportion of embedded dative verbs was 15.1% (31/204 trial) for the Matching/Scrambling condition, 8.6% (29/336) for the Mismatching/Scrambling Condition, 15.2% (21/138) for the Matching/Non-Scrambling condition, 4.6% (15/322) for the Mismatching/Non-Scrambling condition, 14.0% (35/249 trials) for the Matching/Wh-Nominative condition, and 7.7% (26/334) for the Mismatching/Wh-Nominative condition. χ^2 tests showed none of the differences between conditions were reliable.

Finally, we combined information from the analyses of Q-particles and verb argument structure, in order to provide a composite estimate of the proportion of trials in which the dative wh-phrase was interpreted in the embedded clause. The composite figure was based on the proportion of responses in which there was either a Q-particle on the embedded clause verb or an embedded verb that takes a dative argument. Table 10 summarizes this.

Table 10 Composite Estimate

	Composite Estimate Embedded Q-particles and Dative Verbs		
		Embedded	
Conditions	N	%	Count
GNC Matching/ Scrambling	116	36.3	319
GNC Mismatching/ Scrambling	6	1.6	354
GNC Matching/ Non Scrambling	178	57.7	308
GNC Mismatching/ Non Scrambling	12	3.4	349
GNC Matching/ Wh-Nominative	45	14.8	304
GNC Mismatching/ Wh-Nominative	8	2.2	351
Total	365	18.3	1985

The estimated proportion of fronted Dative wh-phrases that were interpreted in the embedded clause was 36.3% (116/319) for the Matching/Scrambling condition, 1.6% (6/354) for the Mismatching/Scrambling Condition, 57.7% (178/308) for the Matching/Non-Scrambling condition, 3.4% (12/349) for the Mismatching/Non-Scrambling condition, 14.8% (45/304 trials) for the Matching/Wh-Nominative condition, and 2.2% (8/351) for the Mismatching/Wh-Nominative condition.

A χ^2 test showed that the estimated proportion of embedded interpretations of wh-phrases was significantly higher in the Matching/Non-Scrambling condition than in the Matching/Scrambling Condition (χ^2 (1)=28.89, p<.01). Also χ^2 tests showed that the estimated proportion of embedded interpretations of wh-phrases was significantly higher in the Matching/Non-Scrambling condition than in the Matching/Scrambling Condition (χ^2 (1)= 147.09, p<.01), and in the Matching/Scrambling condition than in the Matching/Wh-Nominative condition(χ^2(1)= 22.50, p<.01). There were no other reliable differences.

Discussion

There were two main findings in this experiment. Those are summarized in (22).

(22) a. Fronted Dative Wh-phrases are not interpreted in the embedded clauses when mismatching classifiers are provided.

b. Wh-phrases are interpreted more often in the embedded clause in the Matching/Non Scrambling condition than in the Matching/Scrambling condition.

Let us discuss these findings in turn. First, (22a) is an expected result. This finding is consistent with what we have found in previous experiments, i.e., (i) the mismatching classifier creates an expectation for an upcoming RC structure, and (ii) if an upcoming RC is expected a wh-gap dependency is not created inside the RC to avoid

the violation of the RC island. As we have seen, according to every measurement, the number of wh-phrases that are interpreted in the embedded clause within the mismatching classifier conditions is smaller than in the other conditions. Thus, this finding supports our view that mismatching classifiers provide a cue for the RC structure and further that this cue is strong enough to create an expectation for the islandhood of RCs.

The second finding in (22b), however, was not expected. Specifically, the results showing that the number of wh-phrases that are interpreted in the embedded clauses was larger in the Non Scrambling Conditions than in the Scrambling Condition. This finding does not support the findings in either Kamide and Mitchell (1999) or Aoshima et al. (2004). The comparison between the Scrambling Condition and the Wh-Nominative condition shows that wh-phrases are interpreted in the embedded clauses in the Scrambling Condition, and are thus compatible with our results in the on-line experiment, i.e., readers prefer the long distance scrambling analysis of the fronted Dative wh-phrases. However, the results of the Non-Scrambling condition conflict with the results of on-line experiment. In our on-line experiment, experiment 3, we created the base line condition using the type of sentences that correspond to the sentences in the Non Scrambling condition. As we have seen in the previous section, there was no Filled Gap Effect in the base line condition. This result suggests that wh-phrases located between the matrix subject and the embedded subject are not analyzed as scrambled phrases, rather they are treated as matrix indirect objects.

We note two points about this problematic result. The first is that there may be differences between the on-line and off-line experiments. The other is the presence of classifiers.

The first point is relevant to the manner of presenting sentences in the two different experiments. In Experiment 1 and Experiment 3 we used highly similar sentences. However, in each experiment the way of presenting the sentences was different. In the on-line experiment, sentences were presented in a phrase-by-phrase fashion, and previous words disappeared from the computer screen. Readers could not re-read the previous words. On the other hand, in the off-line sentence fragment completion task, participants were able to read sentence fragments while they completed the sentences. My conjecture here is that such differences in experiments may have created the different results. However what creates those differences is totally not clear. More studies are needed to make this point clearer.

The second point is relevant to the design or our experiments. There is a possibility that the classifier creates some problems. Unlike Aoshima et al.'s study or Kamide and Mitchell's study, our experiments contain classifiers. The design of the experiment was not exactly the same as in the two previous studies. Thus, to test the validity of those studies' findings, we would have to run another experiment excluding the classifier. Therefore, it is not reasonable to conclude that Aoshima et al. and Kamide and Mitchell's study were not correct on the basis of our study. Rather, we need to run different experiments excluding the classifier. In the revised experiment, sentence fragments that do not contain classifier should be provided. If the existence of a classifier is the confound, and if the findings of Aoshima et al's

findings are real, we would find a greater number of completions that indicate that the long-distance scrambled wh-phrases are interpreted in the embedded clause and wh-phrases in the scrambling condition than in the non scrambling condition.

11. Experiment 3

The aim of Experiment 3 was to test whether mismatching classifier-noun sequences are sufficient to induce islands for filler-gap dependencies in an on-line task. If classifier mismatches can block long-distance dependency formation, this suggests that sufficiently rich structure is built to activate the islandhood of relative clauses. Furthermore, by using the Japanese Filled Gap Effect paradigm in this study, we aimed to test for evidence of relative clause structure being built in advance of the embedded clause verb.

Participants

Ninety-eight native speakers of Japanese participated in the experiment, all of whom were students at Shizuoka University or Shizuoka Sangyo University, Shizuoka, Japan, or at Meiji Gakuin University, Tokyo, Japan. They all gave informed consent and were paid the equivalent of $5 for their participation in the experiment, which lasted about 30 minutes.

Materials and design

The experimental materials consisted of twenty-four sets of 4 conditions, organized in a 2 x 2 factorial design that manipulated the semantic congruity between

the embedded subject NP and a preceding genitive numeral classifier (classifier match vs. classifier mismatch) and the position of a dative-marked wh-phrase (initial dative vs. medial dative). In the initial dative condition a dative wh-phrase appeared in sentence initial position, whereas in the medial dative condition the dative wh-phrase appeared in the second region of the sentence, between the topic-marked main clause subject and the embedded clause subject. The sentence-initial dative NP is obligatorily analyzed as scrambled, and was previously shown to be preferentially associated with an embedded clause gap (Aoshima et al., 2004). In contrast, the medial dative NP could be interpreted as either an in-situ main clause argument, or as a locally scrambled embedded clause argument. Previous findings suggest that if a Japanese NP can be interpreted as an in-situ argument, speakers prefer this analysis over a scrambling analysis (Kamide & Mitchell, 1999). Therefore, there should be no reason for a Filled Gap Effect in the medial dative conditions, and it serves as a baseline for the initial dative conditions.

The twenty-four sets of items were distributed among four lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with seventy-two filler items in a random order. The filler items were matched with the target items in overall length and complexity, and Genitive Numeral Classifiers, other numeral classifiers and wh-phrases, in main and embedded clauses were equally distributed across stimuli and fillers. A sample set of experimental materials is shown in Table 11.

Table 11 Sample set of experimental materials, Experiment 3

a. Classifier mismatch/Initial dative Condition

1	2	3	4	5	6
Dono-sensee-ni	tannin-wa	3-satu-no	[tosioita	sensee-ga	atarasii
Which-student-Dat	*class-teacher-Top*	*3-Cl(book)-Gen*	*aged*	*teacher-Nom*	*new*

7	8	9	10	11	12
koochoo-ni	yorokonde	okutta]	hon-o	kyoositu-de	yomasemasita-ka?
President-Dat	*gladly*	*gave*	*book-Acc*	*class-room-at*	*read-made-honorific-Q*

'Which student did the class teacher made read three books at the classroom that the old teacher gladly gave to the new president?'

b. Classifier mismatch/Medial dative Condition

1	2	3	4	5	6
Tannin-wa	dono-sensee-ni	3-satu-no	[tosioita	sensee-ga	atarasii
class-teacher-Top	*which-student-dat*	*3-Cl(book)-Gen*	*aged*	*teacher-Nom*	*new*

7	8	9	10	11	12
Koochoo-ni	yorokonde	okutta]	hon-o	kyoositu-de	yomasemasita-ka?
President-Dat	*gladly*	*gave*	*book-Acc*	*class-room-at*	*read-made-honorific-Q*

'Which student did the class teacher made read three books at the classroom that the old teacher gladly gave to the new president?'

c. Classifier match/Initial dative Condition

1	2	3	4	5	6
Dono-sensee-ni	tannin-wa	[[3-nin-no	tosioita	sensee-ga]	atarasii
Which-student-Dat	*class-teacher-Top*	*3-Cl(human)-Gen*	*aged*	*teacher-Nom*	*New*

7	8	9	10	11	12
Koochoo-ni	yorokonde	okutta]	hon-o	kyoositu-de	yomasemasita-ka?
President-Dat	*gladly*	*gave*	*book-Acc*	*classroom-at*	*read-made-honorific-Q*

'Which student did the class teacher made read books at the classroom that three old teacher gladly gave to the new president?'

d. Classifier match/Medial dative Condition

1	2	3	4	5	6
Tannin-wa	dono-sensee-ni	[[3-nin-no	tosioita	sensee-ga]	atarasii
class-teacher-Top	*which-student-dat*	*3-Cl(human)-Gen*	*aged*	*teacher-Nom*	*new*

7	8	9	10	11	12
Koochoo-ni	yorokonde	okutta]	hon-o	kyoositu-de	yomasemasita-ka?
President-Dat	*gladly*	*gave*	*book-Acc*	*classroom-at*	*Read-made-honorific-Q*

'Which student did the class teacher force to read books at the classroom that three old teachers gladly gave to the new president?'

All of the target conditions consisted of twelve regions, segmented into phrases as shown in Table 4. In all conditions the numeral classifier appeared in the third region, the embedded subject in the fifth region, and the second dative NP in the seventh region. An adverb appeared between the embedded dative NP region and the embedded verb region in order to make it possible to differentiate any delayed reading time slowdown effect due to the embedded dative NP from any effect caused

by the embedded verb. As in Experiment 2 the embedded verbs were typical ditransitive verbs like *ageru* 'give' or *watasu* 'pass' that normally appear with both indirect and direct objects. The head of the relative clause in the target conditions always corresponded to the direct object of the embedded verb, and thus the ditransitive verb in the embedded clause always appeared without an overt direct object.

Procedure

The experiment used a self-paced moving window task using identical parameters to Experiment 2, including the use of an argument-verb matching comprehension task and an off-line practice session before the main experiment.

Results

Comprehension accuracy and reading times at each region were entered into a repeated-measure ANOVA, with classifier match (match vs. mismatch) and the position of the dative wh-phrase (initial vs. medial) as within-subjects factors.

Data from 10 participants (10%) whose comprehension task accuracy was below 70% in total were discarded, leaving eighty-eight participants who were included in the analysis. The average comprehension accuracy among the remaining participants was 81.9%. Mean accuracy scores were not significantly different among the four conditions (Fs < 1). Reading times for individual regions that were longer than 3000ms were discarded, affecting 2.5% of trials. Average reading times are shown in Figures 3-4.

Fig. 3. Average reading times in the classifier mismatch conditions, Experiment 3.

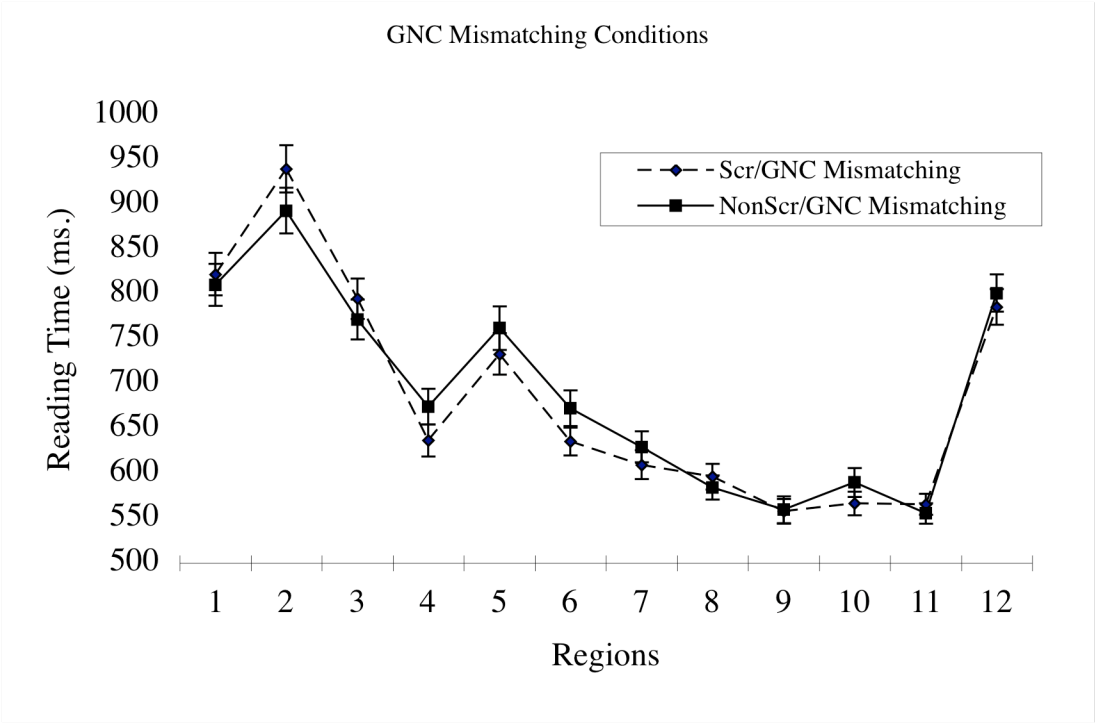
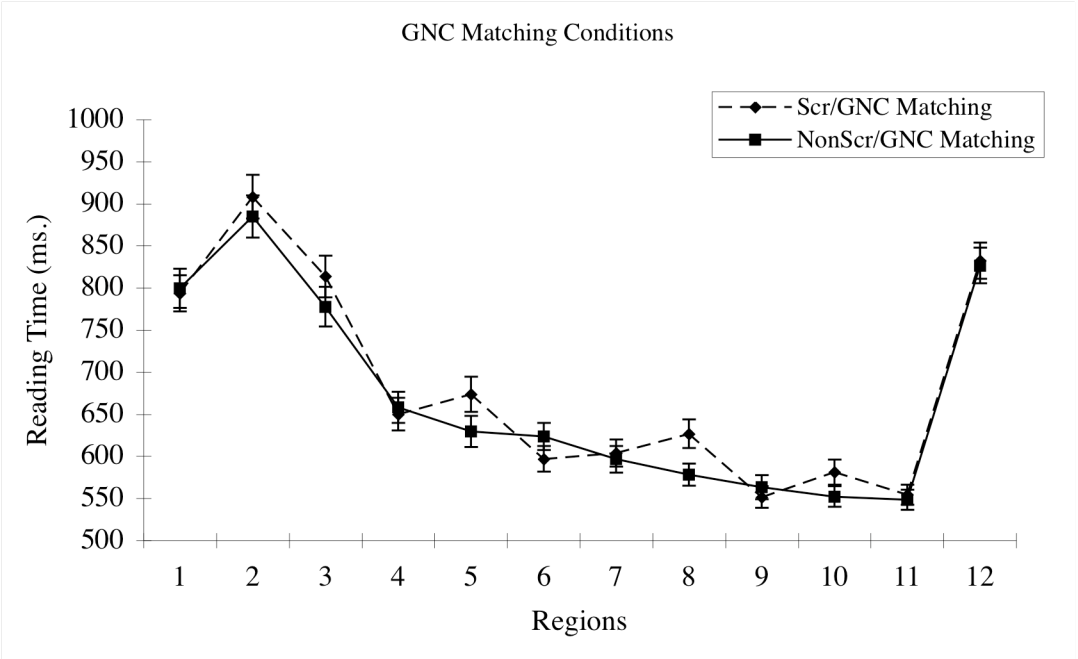


Fig. 4. Average reading times in the classifier match conditions, Experiment 3.



The reading times in regions 1 through 4 did not show any significant differences (all Fs < 1). There were either no main effects or interactions. At the embedded subject NP in Region 5 there was a main effect of classifier match, due to reading times that were 60 ms. slower in the classifier mismatch conditions ($F1(1,87) = 23.5$, MSe = 7592484, $p < .01$; $F2(1,23) = 12.5$, MSe = 6912526, $p < .01$). Pairwise comparisons showed that this effect was significant within each level of the classifier match factor (F1 (1,87) = 5.71, MSe = 1443168, p <. 01; F2 (1,23) = 4.16, MSe = 1653686, p<. 0.5). On the other hand there were no significant effects of the position of the dative wh-phrase, either in the main ANOVA or in pairwise comparisons within each level of the classifier match factor (all *F*s < 1), nor was there an interaction of the classifier match and the position of the wh-phrase.

There were no significant reading time differences in the other regions (Fs < 1).

At the position of the embedded dative NP in the classifier match conditions there was no significant effect of the position of the wh-phrase at the embedded dative NP in Region 7, there were no significant differences in reading times ($Fs < 1$). However, at the adverb position in the immediately following region, region 8, there was a significant main effect of word order in the main ANOVA ($F1(1,87) = 4.0$, MSe = 309963, $p < .05$; $F2(1,23) = 4.45$, MSe =462981, $p < .05$). However there was no significant main effect of Classifier Match factors. There was no interaction. Pairwise comparisons of the word order factor revealed that there was a significant slowdown in reading times for Scrambling/GNC Matching condition (F1(1, 87)=3.73, MSe = 404738, p < .05; F2 (1, 23)=5.35, MSe = 651668, p < .05). In the other

regions, we did not find any significant differences (Fs < 1). In the GNC Mismatching conditions, we found no significant differences at any region between the reading times for the two conditions a (Fs < 1).

Discussion

There were three main findings in this experiment: The replication of the classifier mismatch effect; a Filled-gap effect at the preverbal position in the classifier Matching condition; and no filled-gap effect in the classifier Mismatching conditions. Let us discuss them in turn.

First we replicated the classifier mismatch effect that we observed in the second experiment. This finding support the claim that Japanese readers can calculate the lexically encoded semantic information immediately upon encountering it.

Secondly we replicated the Japanese filled-gap effect observed by Aoshima et al (2004). In the Classifier Matching Condition, there is no indication of the upcoming relative clause. Thus, the parser's initial analysis should be that the embedded clause is a complement clause following its general preference for the complement clause analysis. Because it is analyzed as a complement clause, the fronted Dative Wh-phrase is likely to be interpreted in the most deeply embedded clause as Aoshima et al's study suggests. One concern with this result is the actual position of this reading time slowdown. The eighth region is the region immediately following the Dative NP region. This region is occupied by an adverb. In our design, however, this region is occupied by the same adverb in all conditions. For this reason, it is unlikely that the reading time differences were caused by some factor intrinsic to

each condition. We therefore conclude that this is a spill over effect processes that took place at the seventh region, the embedded Dative NP region.

Let us discuss the consequences of this finding. First, this result replicated Aoshima et al.'s finding in the following two senses. First the fronted Wh-Dative-NP was interpreted in the most deeply embedded clause. Second the Filled-Gap Effect appeared at a preverbal region. These two points, together with Aoshima et al.'s findings, support the position that filler-gap dependencies are constructed in advance of the verb, and that filler-gap dependency formation is motivated by the requirements of the filler. This is, in turn, provides support for the Incrementality Hypothesis. First, the sentence structure is built even before the embedded verb becomes available in the input. This suggests that the parser does not wait for the information from the crucial lexical head. Second, the fronted NP is likely to be interpreted in the embedded clause. This observation suggests that the parser tries to satisfy the requirements of the filler, the thematic requirement, as early as possible.

Now let us turn to the finding in the Classifier Mismatching Conditions. Both in the scrambling and non-scrambling conditions, we do not observe the reading time slowdown at the embedded Dative NP region or at the embedded adverb region. In other words, the filled-gap effect that took place in the Classifier Matching Conditions is not found in the Classifier Mismatching Conditions. This result indicates that the parser avoided the relative clause island violation.

Let us elaborate this point. Aside from the classifier mismatch there were no differences between these four conditions. Additionally, the Classifier Matching Conditions suggest that the indication of long-distance dependency formation comes

at the embedded adverb in region 8. Thus, if the parser created a filler-gap dependency inside the embedded clause, there should be a filled-gap effect compared to the control condition at the embedded Dative NP position or the embedded adverb position as in the Classifier Matching Conditions. The lack of a reading time slowdown suggests that the parser takes advantage of the Classifier Mismatch at the earlier point to avoid creating a filler-gap dependency inside the embedded domain, i.e., the parser detected the presence of the relative clause in advance, and further it could compute the consequence of the relative clause structure, i.e., the islandhood of the relative clause. This is so because before the embedded verb or the relative head, there is no indication of the relative clause structure other than the classifier mismatch. Everything is the same across all the conditions. Thus, we can plausibly conclude that the parser avoided the island violation because of the Classifier Mismatch that took place at the onset of the embedded clause.

One potential alternative to our interpretation of the result is that because of the classifier mismatch effect participants abandoned processing the sentence structure thereafter, and just mechanically pressed the space-bar. There are at least two reasons to question this conclusion. The first is that after the classifier mismatch region in region 5 and the potential spillover region in region 6 there were no significant reading time differences among conditions. If mechanical button pressing without reading the sentence took place, we might expect that there would have been significant differences in the reading pattern after the classifier mismatch region, significantly faster reading times would be the most likely case. However, there were no such differences in reading patterns. The second reason is that there were no

significant differences in accuracy for the comprehension questions. Again, if participants did not read the sentences and just pressed the key we should have obtained lower accuracy for the Classifier Mismatch Conditions. Nonetheless, there were no differences in accuracy among the four conditions.

Taken together, the results strongly suggest that the parser avoided the island violation by means of the presence of the classifier mismatch at the earlier point.

Now let us discuss the theoretical consequences of these findings. Overall, the results of Experiment 3 support the position that the sentence processor is both strongly incremental and still grammatically precise. The fact that the parser avoided the island violation on-line supports the view that a fully elaborated relative clause structure is built immediately after the Mismatching Classifier. Furthermore, if the human parser is incremental, it is easy to imagine that it might sacrifice grammatical precision and just build whatever structures based on the locally available information. If this were true, we would obtain the results that indicated that the parser violated the relative clause island. This is so because, as we have seen, local lexical information does not tell the parser that the current structure is a relative clause, until the embedded verb or the relative head becomes available.

In summary, our results strongly suggest that the sentence processor is strongly incremental and it is equipped with a powerful predictive mechanism. With those two properties, the human parser is designed to satisfy both incrementality and grammatical precision at the same time.

12. General Discussion

12.1. Processing of relative clauses

The goal of this paper was to take advantage of the grammatical properties of Japanese Genitive Numeral Classifiers to investigate the extent to which the sentence processor is strongly incremental and equipped with a powerful predictive mechanism that uses impoverished local information to project the full structural skeleton of an upcoming structure. We have argued that the curious grammatical property of the Genitive Numeral Classifier potentially creates a prediction for an upcoming relative clause structure when the classifier is semantically incompatible with its immediately following Nominative NP. Because a Mismatching Classifier can only be licensed in a relative clause configuration, it may provide an unambiguous cue for an upcoming relative clause structure. While, at the same time, if it is the case the parser can ever create a prediction for an upcoming relative clause structure based on the Mismatching Classifier, the parser has to employ a rich and complex syntactic inference involving multi-step syntactic computations.

Experiments 1 and 2 showed that Japanese readers make use of the Mismatching Classifier to foresee an upcoming relative clause structure. In the off-line sentence completion study in Experiment 1 the participants generated a significantly greater number of sentences containing relative clauses in the Classifier Mismatch Conditions than in the Classifier Match conditions. This result indicates that the mismatching classifier robustly biases Japanese native speakers to generate relative clauses. The on-line self-paced moving window experiment in Experiment 2 showed that Japanese readers make use of the classifier mismatch to predict the

upcoming relative clause structure. The faster reading times obtained at the embedded verb position in the Mismatching Condition indicated that readers projected the relative clause structure upon encountering the classifier mismatch. This finding also suggests that a relative clause verb, which does not bear the complementizer -*to*, provides disambiguating information, i.e., readers can notice that a sentence contains a relative clause as soon as they encounter a bare embedded verb that does not have the complementizer. Thus, we observed reading time differences at the embedded verb position, the first possible disambiguation position for relative clauses.

Experiment 1b and 3 showed that the parser avoids the relative clause island violation by means of the classifier mismatch. This suggests that the parser can compute an elaborated relative clause structure based on the impoverished cue from the classifier mismatch.

The investigations on the second dependent measure in Experiment 1b revealed that the fronted wh-dative NP is rarely interpreted in the embedded clause when the mismatching Genitive Numeral Classifier is provided. Thus, we obtained asignificantly smaller number of sentence completions that indicated that the wh-dative NP was interpreted in the embedded clause.

Experiment 3 used an on-line self-paced moving window experiment to verify that the parser can avoid island violation on-line on the basis of the cue from the classifier mismatch. In the Classifier Matching Conditions the scrambling condition showed a filled gap effect at the embedded adverb position compared to the control condition. On the other hand there was no such effect in the Classifier Mismatching Conditions. The results showed that the parser did not try to create a filler-gap

dependency inside the embedded domain. This lack of effect indicates that the parser avoided the relative clause island violation.

Putting these pieces together, our results indicate that Japanese readers project the full structural skeleton immediately when they encounter a semantic mismatch between the Genitive Numeral Classifier and an embedded Nominative NP. Crucially, the parser seems to create sufficiently rich structure in the embedded clause to prevent the fronted NP from being associated with the embedded verb. This means that the parser preserves both incrementality and grammatical precision at the same time.

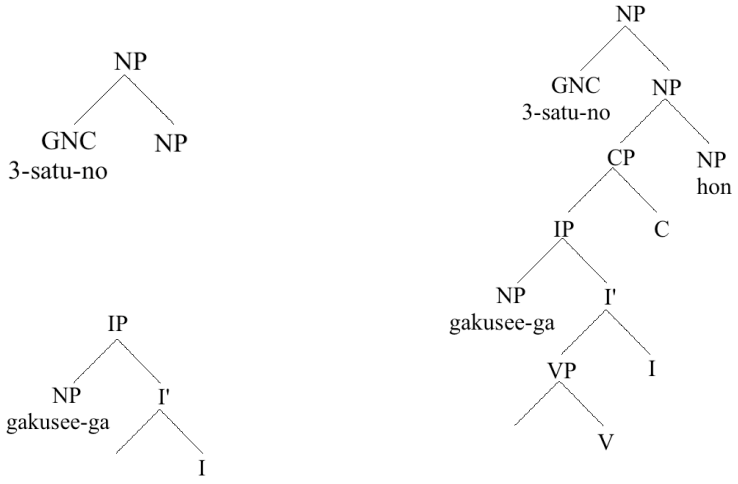
There can be, in principle, a number of ways to deal with the structural uncertainty of Japanese sentences, in the context of on-line sentence processing. Taking an example from filler-gap dependencies, if priority is placed on incrementality, a parser can blindly create a gap where the requirement of the filler can be satisfied as early as possible. In this case, the grammatical precision of structure building could be sacrificed. For the sentence processor, regardless of whether it is inside a relative clause island or not, the embedded verb is always the earliest position where a fronted constituent can receive its thematic role. On the other hand, if the parser's priority is grammatical precision, the parser might defer positing the gap inside an embedded clause until it has clear evidence that the embedded clause is not an island. Under this approach, the parser can avoid the danger of creating a dependency that might turn out to be ungrammatical. In this case, however, the incrementality of structure building would be sacrificed.

Our experimental results indicate that the parser does not sacrifice either full incrementality or grammatical precision. Rather, by dynamically exploiting rich grammatical inferences based on relatively impoverished information from the input, the parser manages to achieve incremental and grammatically precise structure building.

12.2. Incremental structure building with a powerful predictive mechanism

Although, we have emphasized the necessity of a powerful predictive mechanism that should underlie the incremental sentence processor, of the specific mechanism that the parser is equipped with has not been discussed in detail so far. Any approach in which the parser can access grammatical knowledge of Genitive Numeral Classifiers and relative clauses can, in principle, handle the parser's behavior we have seen so far (see Inoue, 1991 or Inoue and Fodor 1995 for example). Still, however, the problem of how and at what point in time grammatical information is consulted and the elements in the input are incorporated into a syntactic tree structure remains unclear. As is clearly argued in Lombardo & Sturt (2002), certain extra-lexical knowledge has to be exploited in order to achieve incremental processing of sentences that we have seen so far. The crucial step in the processing of these sentences containing Mismatching Classifiers is that the parser projects a full structural representation for the relative clause at the exact position of the classifier mismatch, as illustrated in Fig. 5.

Fig. 5.



One approach that has extensive power of prediction is the so-called left-corner parser (Abney and Johnson 1991; Aho and Ullman 1972; Babyonyshev and Gibson 1995; Gibson 1991; Schneider 1999; Stabler 1994 among others). The left-corner parser has various attractive features as a model of the human sentence processor. However, even a left-corner parser does not capture such massive prediction of upcoming material. Based on lexically encoded grammatical information, a left-corner parser can project the two independent nodes of NP and IP based on the Classifier Mismatch. Even though a left-corner parser can predict the upcoming sister nodes of the given input, as far as this prediction is dependent on the bottom-up information from each lexical item, it is not possible to project and incorporate the nodes that are necessary for the relative clause structure. Thus, for the parser to accomplish the predicted structure in Fig. 5 there must be a syntactic inference in which the locally predicted NP node and the IP node are connected to each other via a CP node, i.e., a prediction based on predicted nodes is necessary, a form of recursive prediction. A challenge for this type of predictive mechanism is how to constrain the predictive

power in such a way that the mechanism does not overgenerate possible structures. Although a mechanism that incorporates recursive prediction is currently not available, our results suggest that the human sentence processor is designed in such a way.

13. Conclusion

In this chapter we have seen evidence for strong incrementality. In the course of discussion, we have argued that to achieve the strong incrementality, the parser is equipped with a powerful predictive mechanism. We have discoverted a potential trigger for the predictive mechanism of sentence processing. What we have discovered was, the classifier mismatch, a relatively an indirect cue can be a trigger for the predictive mechanism.

We have also discovered that the representation that is built by the predictive mechanism must be rich enough to represent the islandhood of RCs. We have argued that the predictive mechanism can project the detailed structure of the RC that encodes the information of the islandhood of the RC upon encountering the classifier mismatch.

In the later chapters, we will see how a specific mechanism of sentence processing can capture the prediction made by the classifier mismatch.

CHAPTER 4. CONDITIONALS AND LONGER DEPENDENCY FORMATION

1. Introduction

The goal of this chapter is to further investigate the nature of the predictive mechanism underlying the human sentence processor.

In the previous chapter, we have seen that the parser makes use of a relatively indirect cue, the classifier mismatch effect, to predict an upcoming relative clause structure. Based on the results of a series of experiments, we have established the following claims. First, we have seen in the previous chapter that there is a potential trigger for the predictive mechanism, the classifier mismatch. The experimental results demonstrated that even an indirect cue such as the classifier mismatch can be used by the parser to predict the upcoming structure. Second, we have seen the representation that the predictive mechanism accesses. Through the investigation of the interaction between the Longer Dependency Bias (LDB) and relative clause islands, we have argued that the representation built by the predictive mechanism must be such that it can derive the island effects of relative clauses.

In this chapter, we will turn to another environment where we can test the nature of the predictive mechanism, namely conditional clauses in Japanese. This chapter is devoted to investigate the following two points. First we will see the nature of the LDB. We will see what is the driving force behind LDB. The first half of the chapter, therefore, is spent investigating this point. Second, based on the finding of the first half, we will investigate the processing of conditional clauses, where LDB is

blocked. We will see that looking into the blocking effect of LDB in conditional clauses will tell us about the nature of the predictive mechanism.

2. On the Motivations for the Longer Dependency Bias

In chapter 2, we have examined how the longer dependency is constrained. We have confirmed that LDB is best understood in terms of the human sentence processor's general property of completing a dependency as early as possible. We also confirmed that the LDB can be best captured by the constraint satisfaction model (Boland et al. 1995; Boland and Boehm-Jernigan 1998; Macdonald et al. 1994a; MacDonald et al. 1994b; Tanenhaus et al. 1989 among many others) in such a way that the grammatical requirement on the fronted element is satisfied as early as possible (Aoshima et al. 2004). Although the general direction has been made clear in the previous chapter, we have not discussed what is actually the "grammatical requirement" on the fronted phrase that needs to be satisfied as early as possible. As is pointed out in Aoshima et al. (2004), there are several possibilities. Bearing in mind that the fronted NPs are all wh-phrases bearing Dative Case in the series of experiments that we have seen, it is obvious that there are at least four possibilities: it could be the thematic requirement (theta criterion: Chomsky 1981); it could be wh-feature satisfaction (the requirement of wh-Q binding: Harada 1972; Saito 1989); it could be Case feature satisfaction (structural Case assignment: Chomsky 1981; Saito 1982, 1983; Takezawa 1987); or it could be a combination of these features.

One of the goals of this chapter is to examine which feature(s) on the fronted NP drives the formation of the longer dependency. Through two sets of experimental

studies, we will conclude that the motivation for the LDB is the establishment of an association between a wh-phrase and a Question Particle (Q-particle) as soon as possible. The first set of experiments is concerned with a manipulation of the features of the fronted NPs. We will see what happens if we change the fronted NP from a wh-NP to a non-wh-NP (referential NP). We will review the results of two experiments and show that only the wh-NP creates the LDB. The second set of experiments will be concerned with the predicted verbal morphology. Specifically, we will ask whether the LDB is created if the predicted verbal morphology cannot host a Q-particle. To examine this point, we will consider the processing of Japanese conditional constructions.

Another purpose of this chapter is to see how the structure is predicted in head-final constructions during online processing, through a detailed examination of the experimental findings on conditional clauses. I will show that the interaction between long-distance scrambling and the syntactic properties of conditional clauses gives us a clue to exactly what is predicted by what kind of element, i.e., we can see a trigger of the predictive mechanism, and the representation that the predictive mechanism accesses.

Experiments 1 and 2 are concerned with the driving force behind the longer dependency bias. In two experiments, we test the behavior of the sentence processor when a property of the fronted materials is manipulated, incorporating Aoshima et al.'s (2004), and Aoshima's (2003) Japanese filled-gap paradigm as a probe for the LDB. Specifically, we test the wh-scrambling construction, and construction containing the scrambling of referential NPs. The results of these experiments

demonstrate that the wh-feature plays a crucial role in creating the longer dependency.

Experiments 3 through 6 are concerned with the trigger and the representation employed by the predictive mechanism. The main purpose of these experiments is to test the behavior of the sentence processor when the predicted upcoming structure does not license the wh-feature of the fronted phrase. For this purpose, Japanese conditional clauses provide a good testing ground. By taking advantage of the grammatical properties of Japanese conditional clauses, we will first show what kind of bottom-up information can be a trigger for the structural prediction of conditional clauses, and what is actually predicted by the predictive mechanism. Two experiments (an online and an offline experiment) are devoted to answering these questions. The results of these experiments will reveal that the conditional adverb *mosi* "by any chance" sets a prediction for the upcoming conditional verbal morphology.

After establishing the trigger and the contents of the prediction, we will test how the parser behaves when the predicted upcoming structure is not compatible with the feature carried by the fronted phrase. In particular, by taking advantage of the fact that the conditional verbal morphology is not compatible with the Q-particle and the fact that Japanese conditional clauses are not islands, we can test whether a longer dependency is preferentially formed when the conditional verbal morphology is predicted by means of the adverb *mosi*. The other two experiments are devoted to testing this point. The results of these experiments demonstrate that when the

upcoming verbal morphology that cannot host Q-particle is predicted, the LDB is blocked.

Taken together, these two sets of experiments reveal a potential trigger for the predictive mechanism and the representation accessed by the predictive mechanism, as well as the driving force for the LDB in general.

3. The Source of the Bias for Longer Filler-Gap Dependencies in Japanese[27]

3.1. Longer Dependency Bias

To understand the logic of the experiments, it is helpful to review the motivation for the LDB again here.

First let us summarize Aoshima et al.'s (2004) basic assumptions. Their account of the LDB appeals to the following two important assumptions concerning the sentence processing mechanism. The first assumption is principle-based parsing, whereby grammatical principles are directly incorporated into the mechanism of sentence processing in order to build structure (Berwick 1991; Crocker 1994, 1996; Pritchett 1992b). The second assumption is the so-called constraint satisfaction model (Boland et al. 1995 among many others) in a general sense, in which all categories are associated with lexical constraints, and all categories can initiate attempts to satisfy those constraints. Thus, each lexical item in the input can potentially initiate the structure building needed in order to satisfy the lexical constraints that it introduces.

[27] This section is a summary of a paper with the same title (Aoshima et al., 2005) read at the 18th Annual CUNY Conference on Human Sentence Processing, at the University of Arizona, Tucson. I am grateful to Sachiko Aoshima and Colin Phillips for allowing me to include these studies in this dissertation.

Within this view, grammatical constraints such as the theta criterion or wh-feature satisfaction are understood as lexical constraints of lexical items in the input.

Adopting these two assumptions, Aoshima and colleagues argue that the interaction of the basic word order of Japanese sentences, i.e., head-finality and the possibility of long-distance scrambling plays a crucial role for creating the LDB.

Because of the basic word order of Japanese sentences, the most deeply embedded verb is normally[28] the first verb to be encountered during the processing of a sentence from left-to-right, as the schematic structure in (1) shows.

(1) [Subject [Subject Object ***Verb***-comp] verb]

[28] Note that there is always a possibility that an embedded clause is post-posed as in right dislocation construction as in (i). However, this type of sentences requires special focus interpretation and special intonation patterns. Thus it is not likely that parser's initial choice would be the right-dislocated structure rather than the simple embedded clause structure.

(i) Walt-ga itta(-yo), Sue-ga sukida-tte.
W-nom said S-nom like-comp
"What walt said was that he likes Sue."

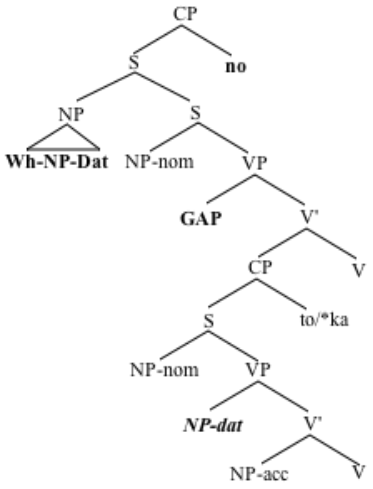
A careful examination is of course necessary about what can be the parser's initial choice and what cannot. But in this study, I continue to assume that the parser's initial choice is always a complement clause if there is no obvious cue for other alternatives.

This, in turn, means that the position of the most deeply embedded verb is the first position where grammatical constraints conveyed by each lexical item can potentially be satisfied or confirmed by an explicit lexical head.

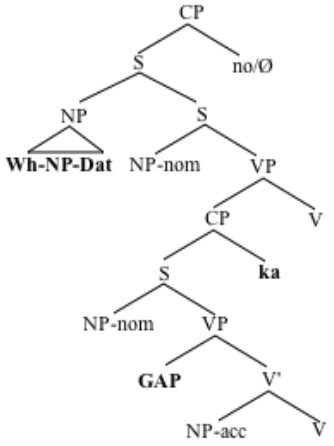
Second, if a long-distance scrambled NP is involved, it creates a temporary ambiguity with respect to its potential gap position. Given the string in (2) for example, during the online processing of a sentence, there can be two possible gap positions where the Dative Wh-phrase can potentially receive its theta role. One is the indirect object position in the matrix clause. A verb like *iu* "say" can take a dative NP as its indirect object as well as a clausal complement. Thus, a gap in the matrix clause as in (2a) is a possible option. The other possibility is to locate the gap in the embedded clause. In this case, the embedded verb is most likely to be a verb like *ageru* "give" that can take a Dative NP as its object.

(2) Wh-NP-Dat [Subject [Subject …

a. Wh-NP-Dat [Subject **[gap]** [Subject Object V] *SAY*-Q]



b. Wh-NP-Dat [Subject [Subject **[gap]** Object *GIVE*-**Q**] V-Q]



If the first option is taken, the Q-particle that must be associated with the fronted wh-phrase can only be attached to the matrix verb, and the sentence is unambiguously interpreted as a direct wh-question (if the sentence-initial clause is confirmed as the matrix clause). Japanese wh-interrogative sentences have a well-known restriction such that the Q-particle which is understood as a complementizer must c-command

the wh-phrase that it licenses (Hagstrom 1998; Harada 1972; Saito 1989). Thus, because of this grammatical constraint, the Q-particle that is associated with the fronted wh-phrase cannot be attached to the embedded verb. On the other hand, if the gap is located in the embedded clause, the Q-particle that licenses the wh-phrase can be attached to either the embedded verb or the matrix verb, depending on the interpretation of the sentence. If the Q-particle is attached to the embedded verb, the sentence is interpreted as an indirect wh-question. If it is attached to the matrix verb, the sentence becomes a direct wh-question.

The most important difference between these two options is the possibility of attaching the Q-particle to the embedded verb. If the gap is located in the matrix clause, the Q-particle cannot be attached to the embedded verb, but if it is in the embedded clause, the Q-particle can be attached to the embedded verb.

Now, let us look at this temporary ambiguity created by long-distance scrambling and head finality from the perspective of the grammatical constraint satisfaction model. From the viewpoint of the grammatical constraint satisfaction model, the fronted dative wh-phrase can initiate an attempt to satisfy its grammatical constraints. It should be clear that this element is associated with at least three types of requirements that we can rephrase in terms of grammatical features: a thematic feature; a Case feature; and a Wh-feature. If any of these features triggers structure building so that they can be satisfied, the earliest point where they are satisfied is the most deeply embedded verb position because it is the first verb to be encountered. From this, the LDB naturally follows. Because it is the most deeply embedded verb that can satisfy and confirms the grammatical features on the fronted element at the

earliest point, gap creation in the most deeply embedded position is the preferred option.

3.2. Manipulating the Trigger

In the above discussion, it becomes clear that in Aoshima et al.'s account, it is the features that are conveyed in the fronted element that drive structure building, so that these features will be satisfied. Now, let us return to our original question: what is the real motivation for the LDB? As we have seen, the fronted dative wh-NP has to carry at least three features: a theta role; case; and a wh-feature. Thus, if the LDB is driven by grammatical feature satisfaction, it is one of these features or a combination of these features that triggers the longer dependency formation. Given that if an NP is not case marked it cannot be scrambled, and that gap creation is obviously triggered by the non-canonical word order that is signaled by the order of case-marked NPs, it is clear that Case feature plays an important role in initiating structure building (see Aoshima 2004 for detailed discussion on this point). Thus, it should be the case that the motivation for the LDB is a combination of the Case feature and either the theta role or the Wh-feature, or both. Now the question is how we can test these possibilities.

Recalling that it is not the case that only wh-phrases can be scrambled in Japanese, the answer should be quite simple: we can just manipulate the wh-feature on the scrambled NP. In other words, we make a minimal pair of wh-scrambling and the scrambling of a referential NP, and we can test whether the wh-feature is the

crucial feature for the LDB or not. Let us see the basic logic by taking a look at the following examples with schematic structures.

(3) a. **Which**-NP-Dat [Subject [Subject … Verb-Q] Verb]

Dono-gakusee-ni [Walt-wa [Master Yhudi-ga

which-student-dat W-top M-Y-nom

ringo-o ageta-ka] tazuneta.

apple-acc gave-Q ask-past

"Walt asked which student Master Yuhudi gave the apple to."

b. (**That**-)NP-dat [Subject [Subject … Verb-Q] Verb]

(**Sono-**)gakusee-ni[Walt-wa [Master Yuhdi-ga

that-student-dat W-top M-Y-nom

ringo-o ageta-ka] tazuneta.

apple-acc gave-Q ask-past

"Taroo asked whether Jiroo gave the apple to that student or not."

(3a) is the same construction as Aoshima et al. originally tested. There, the fronted phrase is a wh-phrase. (3b), on the other hand, contains a fronted phrase that does not bear a wh-feature. This construction is different from (3a) only in the wh-feature of the fronted phrase. Given these two constructions, what we expect is the following. First, in (3a) we expect that a longer dependency be preferentially formed as Aoshima et al. originally observed. On the other hand, in (3b) we expect that the longer dependency should be formed if the motivation for the LDB is confirmation of the thematic role of the fronted phrase because both in (3a) and in (3b), the fronted phrase

must bear some theta roles due to the Theta Criterion (Chomsky 1981). However, if the wh-feature, or a combination of the wh-feature and the theta role is crucial for the LDB, we do not expect a longer dependency to be formed in (3b), simply because the fronted NP does not bear a wh-feature.

3.3. The Japanese Filled-Gap Paradigm and the Motivation for the Longer Dependency Bias

As we did in the earlier studies on RCs, we will adopt the so-called Japanese Filled-Gap Paradigm (Aoshima et al. 2004). In order to see the specific predictions, it is helpful to briefly review the Japanese Filled-Gap Paradigm. Also, we will see another question that is related to the design of the original Japanese filled-gap experiment that is not our main concern, but still an important question to be investigated.

3.3.1. Aoshima et al., (2004) and Aoshima (2003)

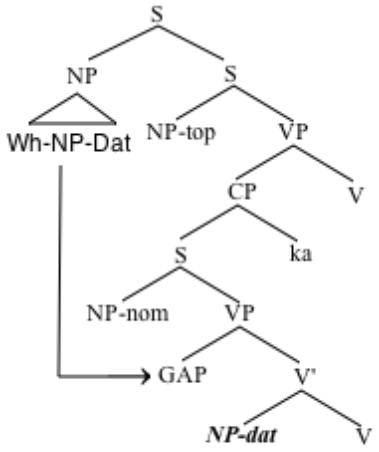
First of all, the basic intuition behind the Japanese Filled-Gap Paradigm is as follows. If a longer dependency is created, the gap should be created in the most deeply embedded clause during the processing of a sentence. If there is an additional dative phrase besides the fronted dative phrase in the embedded clause, a surprise effect occurs. To test this, Aoshima et al., made use of the structure in (4). In this structure, there is a dative NP in the embedded clause in addition to the fronted wh-dative-NP. In this structure, if the gap is created and it is associated with the fronted Wh-dative, the embedded clause has to contain two dative NPs. In Japanese such a

construction containing two dative NPs in one clause is not common. Thus, if the fronted wh-dative is interpreted in the embedded clause, Japanese readers do not expect another dative NP to be present in the embedded clause, resulting in a reading time slowdown upon encountering the second dative NP.

(4) Wh-Scrambling Condition

Dono-shain-ni senmu-wa shachoo-ga …

which-worker-**dat** senior-director-top president-nom …

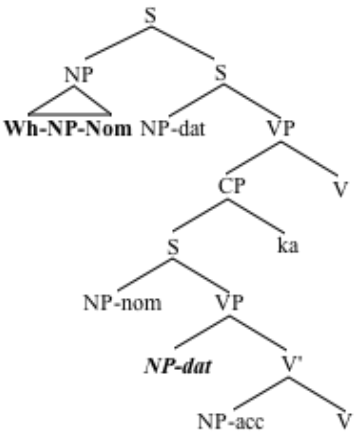


Aoshima et al. and Aoshima (2003) observed this effect by comparing the structure in (4) with two types of baseline conditions (5).

(5) a. Wh-Nominative Condition

Dono-shain-ga senmu-ni shachoo-ga…

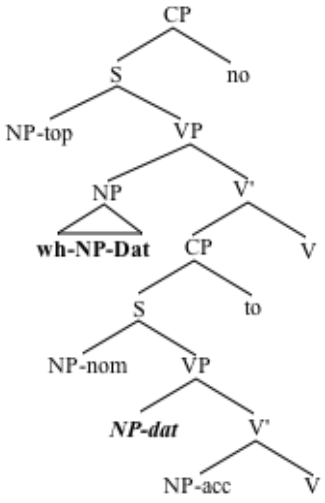
which-worker-**nom** senior-director-dat president-nom



b. Medial Wh-Dative Condition

Senmu-wa **dono-shain-ni** shachoo-ga …

senior-director-top which-worker-**dat** president-nom



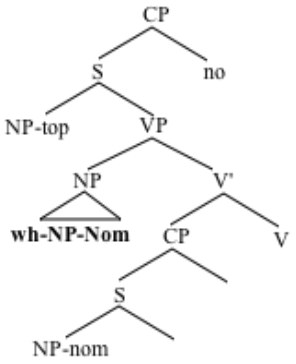
(5a) is a straightforward baseline condition. Given that Japanese does not normally allow scrambling of Nominative NPs (Saito 1985), there is no moved element involved in this structure. Therefore, we do not expect any long-distance filler-gap dependency formation, and we do not expect any surprise effects associated with moved phrases. Thus, we predict that sentences with the structure of (4) will show significant reading time slow down at the second dative NP position compared with (5a).

Aoshima (2005) obtained the same results in the comparison of (4) and (5b). (5b), however, is a little bit trickier as a base-line condition compared to (5a). (5b) contains a wh-dative phrase that is located in between two subject NPs. It has been independently observed that in this construction the wh-dative-NP is not interpreted in the embedded clause even though it is grammatically possible (Kamide and

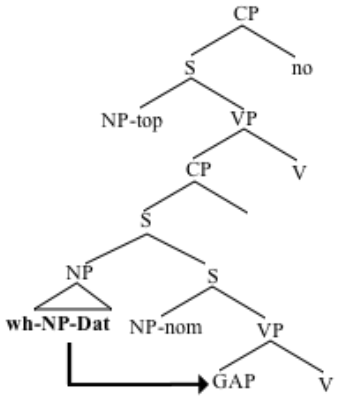
Mitchell 1999). In other words, although a string like (6) can potentially be parsed in two ways with respect to the position of the dative wh-phrase (the matrix argument analysis and embedded argument analysis respectively), somehow the parser pursues the matrix argument analysis in (6a). (6a) is the structure of the matrix analysis where the dative wh-phrase is understood as the indirect object of the matrix verb such as *iu* "say". On the other hand in (6b), the dative wh-phrase is understood as a scrambled indirect object of an embedded verb like *ageru* "give".

(6) Senmu-wa dono-shain-ni shachoo-ga…
senior-director-top which-worker-dat president-nom

a. Matrix analysis



b. Embedded analysis



For this reason, we do not expect that a dependency between the dative wh-phrase and a gap is formed in the medial wh-dative condition, and thus we do not expect any effects associated with the moved phrase either.

The results reported in Aoshima's studies confirm the prediction we have seen above. The construction containing a fronted dative wh-phrase at the beginning of the sentence showed slower reading time at the position of the second dative NP compared to both of these baseline conditions.

3.3.2. Manipulating the Features on the Fronted Phrase

To test whether the wh-feature is the crucial feature for the LDB, we can make use of the same type of baseline conditions as Aoshima et al. and Aoshima (2005) did. First, we can use the three conditions above, the wh-scrambling condition in (4), the wh-nominative condition in (5a) and the wh-medial condition in (5b), to replicate Aoshima's original results i.e., the wh-scrambling condition showed slower reading time at the second dative NP position compared to the baseline conditions. Furthermore, the comparison between (5a) and (5b) will allow us to clarify Kamide and Mitchell's observation that a medial dative-phrase is not interpreted as a scrambled embedded argument. Our specific prediction about this comparison is that there will be no significant difference in reading times at the second dative NP position between (5a) and (5b).

By manipulating the wh-feature of these three conditions, we can gain the conditions that do not contain wh-phrases. In so doing, we just have to delete the wh-determiner *dono* "which" as in (7).

(7) a. Scrambling Condition

Josee-shain-ni senmu-wa

female-worker-dat senior-director-top

shachoo-ga …

presidento-nom

b. Medial Referential NP condition

senmu-wa **Josee-shain-ni**

senior-director-top **female-worker-dat**

shachoo-ga …

presidento-nom

c. Nominative Referential NP condition

Josee-shain-ga senmu-ni

female-worker-nom senior-director-dat

shachoo-ga …

presidento-nom

Looking at the examples (7b) and (7c), it is obvious that they are the same type of constructions except for the particles on the subjects: one is a nominative Case marker and the other is a topic marker. Despite the differences in the markers on the subjects, the sentence initial NPs in both of the examples are most likely to be interpreted as matrix subjects. Thus, if we are not interested in the differences between nominative-marked NPs and topic-marked NPs, the comparison between (7b) and (7c) is not informative for our current purpose. In order to test whether the wh-feature is crucial for the LDB, it is sufficient to compare (7a) and (7b). If the wh-feature based hypothesis or the wh-feature and thematic requirement based hypothesis is correct, we do not predict any slowdown at the second dative NP position in (7a) compared to (7b). If, on the other hand, the thematic requirement hypothesis is correct, we do expect the slowdown at the second dative NP position in (7a) compared to (7b).

3.4. Experiments

3.4.1. Experiment 1

This experiment was designed to test how wh-phrases are interpreted in wh-scrambling structures, using Aoshima et al.'s (2004) and Aoshima's (2005) Japanese Filled-Gap paradigm. The aim was to determine whether a fronted wh-phrase in Japanese is preferentially interpreted in the main clause or in the embedded clause, clarifying the results from Aoshima and colleagues, and the aim was also to determine whether a medial wh-phrase is preferentially interpreted in the main clause or in the embedded clause, clarifying Kamide and Mitchell's observation. If a fronted wh-phrase is preferentially interpreted in the embedded clause, we should observe the same filled-gap effect as in Aoshima's series of studies. The same holds true for the medial wh-phrase. If it is interpreted in the embedded clause, we should observe the filled-gap effect also.

Participants

Thirty-seven native speakers of Japanese participated in the experiment. All of them were students at Hiroshima University, Japan. They were paid $5.00 or its equivalent for their participation in the experiment, which lasted about 30 minutes.

Materials and design

Twenty-four sets of three conditions each were used in the experiment, which manipulated the position of the wh-phrase (scrambled vs. medial vs. nominative). The

24 sets of items were distributed among three lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with 72 filler items in a random order. The filler items were matched with the target items in overall length and complexity.

A sample set of experimental items is shown in (8) where superscripts indicate the regions used in presenting the sentences.

(8) A sample set of experimental items

a. Wh-Scrambling

Dono-shain-ni[1] senmu-wa[2] shachoo-ga[3]

which-worker-dat[1] senior-director-top[2] president-nom[3]

kaigi-de[4] buchoo-ni[5] shookyuu-o[6]

meeting-at[4] department-director-dat[5] promotion-acc[6]

yakusoku-sita-to[7] sirase-masi-ta-ka?[8]

promice-did-comp[7] notice-hon-past-Q[8]

"Which worker did the senior director noticed that the president promised the promotion to the department director at the meeting?"

b. Wh-Medial

Senmu-wa[1] **dono-shain-ni**[2] shachoo-ga[3]

senior-director-top[1] **which-worker-dat**[2] president-nom[3]

kaigi-de[4] buchoo-ni[5] shookyuu-o[6]

meeting-at[4] department-director-dat[5] promotion-acc[6]

yakusoku-sita-to[7] sirase-masi-ta-ka?[8]

"Which worker did the senior director notified that the president promised the promotion to the department director at the meeting?"

c. Wh-Nominative

Dono-shain-ga[1] senmu-ni[2] shachoo-ga[3]

which-worker-nom[1] senior-director-dat[2] president-nom[3]

kaigi-de[4] buchoo-ni[5] shookyuu-o[6]

meeting-at[4] department-director-dat[5] promotion-acc[6]

yakusoku-sita-to[7] sirase-masi-ta-ka?[8]

promice-did-comp[7] notify-hon-past-Q[8]

"Which worker notified the senior director that the president promised the promotion to the department director at the meeting?"

In the wh-scrambling condition and the wh-medial condition, a main clause subject NP marked with the topic-marker *-wa* was followed by an embedded clause introduced by a nominative marked NP. In the wh-nominative condition, the sentence initial wh-phrase bore nominative case, and the dative marked wh-phrase followed it. The dative marked wh-phrase was followed by an embedded clause, which began with a nominative-marked NP. We adopt the well-supported supposition that the nominative marked NP provides a strong cue for the onset of an embedded clause (Inoue 1991; Miyamoto 2002 among others). The position of the embedded clause in these conditions is perfectly natural in Japanese, because Japanese is a head-final language, the main clause verb normally appears at the end of the sentence, following the embedded clause. All the conditions contained an overt dative-marked NP. The verb of the embedded clause was marked with a declarative complementizer, and the verb of the main clause with a question particle, indicating that the sentence should be interpreted as a matrix wh-question.

In all three conditions, the dative NP in the embedded clause was the second dative NP in the sentence. In the scrambled condition, the dative wh-phrase at the beginning of the sentence should also have been associated with the embedded clause, if a longer dependency is preferentially formed. If the fronted wh-phrase in the scrambled condition is associated with the embedded clause, then readers should be surprised to encounter the second dative NP.

In the other two conditions, the two sentence initial NPs were matched to the scrambled condition in the respect that there was one wh-phrase and one dative NP. However, in these two conditions, we did not expect that either of the two NPs in the matrix clause would be associated with the embedded clause. In the wh-medial condition, the wh-phrase was located after the matrix subject NP and the dative marked NP in this position was not expected to be associated with the embedded clause for the reason we have seen above (see Kamide and Mitchell 1999 for details). In the wh-nominative condition, on the other hand, the sentence initial nominative wh-phrase should not be associated with the embedded clause, since Japanese normally does not allow the scrambling of nominative NPs (Saito 1985 among others).

Procedure

The experiment was conducted on Dell laptop computers running *Linger* developed by Douglas Rohde at MIT (Rohde 2001-2003). Participants were timed in a phrase-by-phrase self-paced non-cumulative moving-window reading task (Just et al. 1982). The segmentation indicated with superscripts in (8) was used in the

presentation. The complementizer and Q-particle were presented together with the verbs because both the complementizer and Q-particle are bound morphemes in Japanese. Sentences were presented using Japanese characters. Stimulus segments initially appeared as a row of dashes, and participants pressed the space bar of the keyboard to reveal each subsequent region of the sentence.

In order to guarantee that participants attended to the stimuli, a subject-verb matching task was presented after each trial. This task was adopted from Nagata (1993), and similar to the task used in the study by Miyamoto and Takahashi (2000, 2003). A verb was displayed on the computer screen followed by two agent NPs, corresponding to the topic-marked NP and the nominative-marked NP in the target sentence, and participants had to decide which of the NPs was the subject of the verb in the sentence just read by pressing one of two keys on the keyboard. In the comprehension task, the two agent NPs were displayed without case-marking, in order to exclude the possibility of case-based question answering strategies.

In order to familiarize participants with the subject-verb matching task, an off-line practice session was included prior to the self-paced reading task.

Data Analysis

Analyses were conducted on comprehension task response accuracy and reading times. All data from participants whose comprehension task accuracy was below 80% in total were discarded.

Comprehension task accuracy

Among the 37 participants included in the analysis, the average comprehension accuracy was 82.5%. Mean accuracy scores did not differ significantly across the three conditions (all Fs < 1).

Self-paced reading

The reading time analysis yielded the following results. Reading times for all three conditions are shown in Fig. 6, and a comparison of the scrambling condition and the wh-nominative condition is shown in Fig. 7, and a comparison of the wh-medial condition and the wh-nominative condition is shown in Fig. 8.

At the third region, there were significant differences between reading times in the wh-scrambling and wh-nominative condition ($F1(1, 36)=6.48$, $p<0.01$; $F2(1, 23)=2.86$, $p=1$). Other than that, at all regions prior to the fifth region, there were no significant differences in reading times among all three conditions. At the following regions we observed a Filled-Gap Effect, reflected in longer reading times for the embedded dative NP, only in the scrambled word order.

At the embedded dative NP (region 5), there was a significant main effect of word order type ($F1 (1, 36)=3.79$, $p < 0.05$; $F2 (1, 23)=4.03$, $p < 0.05$). Pairwise comparison revealed that the dative NP was read significantly more slowly in the Wh-scrambling condition than in the Wh-nominative condition ($F1 (1, 36)=3.79$, $p < 0.05$; $F2 (1, 23)=4.03$, $p < 0.05$). On the other hand, the comparison between either the Wh-scrambling condition and the Wh-medial condition, or between the Wh-

medial condition and the Wh-nominative condition showed no corresponding slowdown (all Fs < 1). There were no other significant differences in reading times.

Fig. 6.

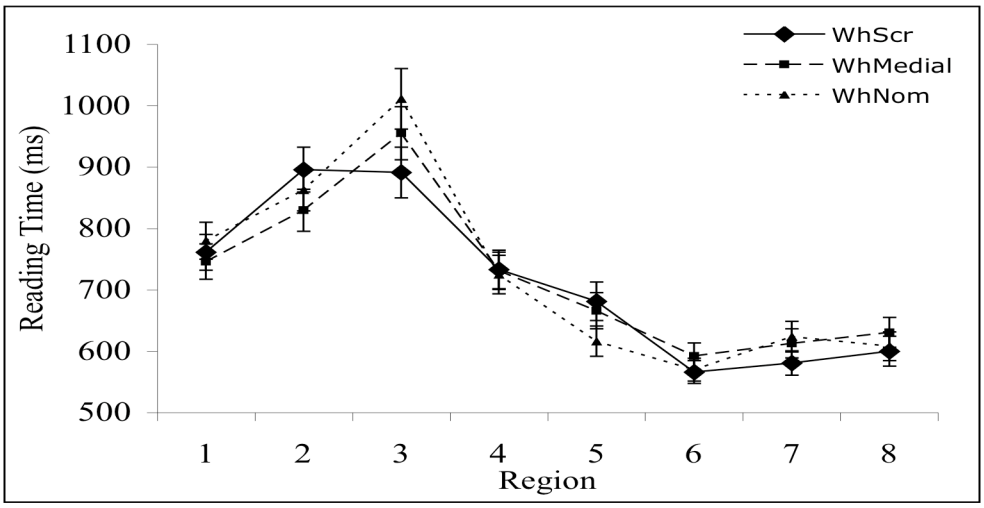


Fig. 7.

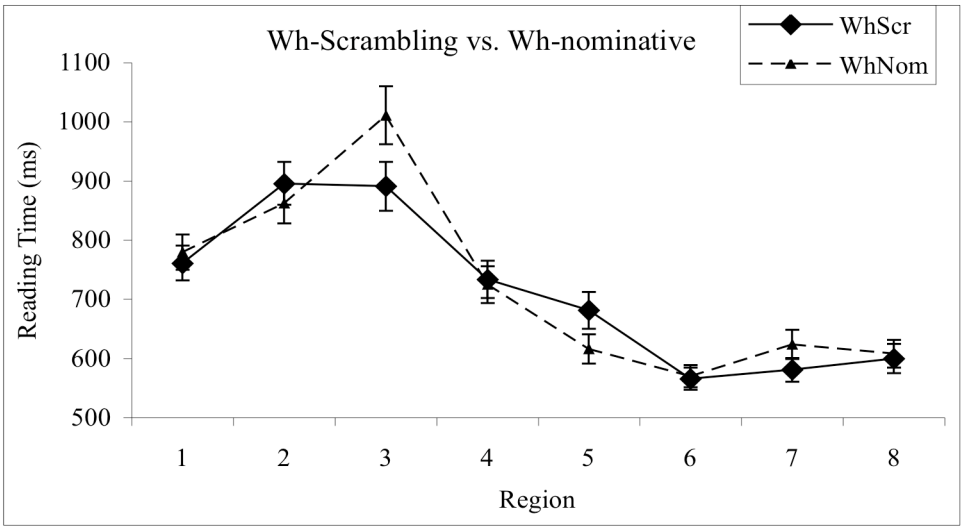
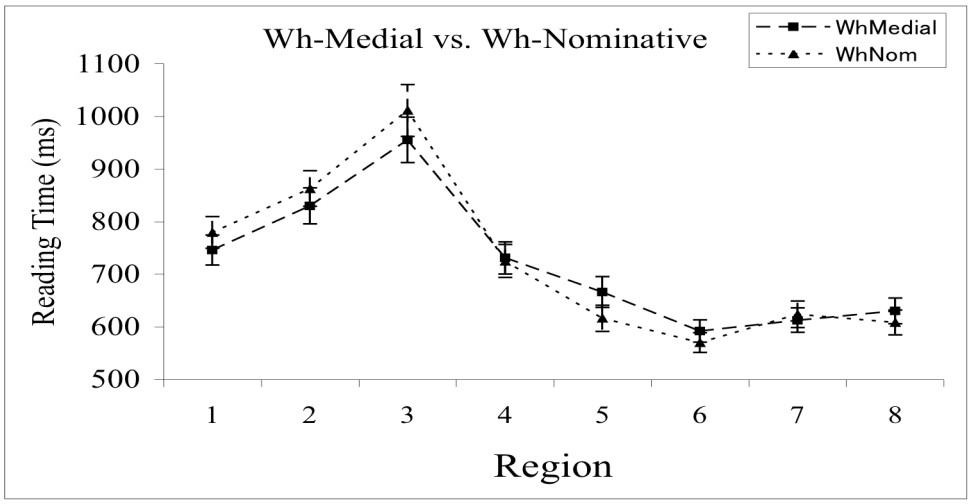


Fig. 8.



Discussion

The main result of this experiment is that a slowdown in reading time was observed in the wh-scrambling condition at the embedded dative NP (region 5), relative to the reading time for the same NP in the Wh-nominative condition. Aoshima et al.'s (2004) results are thus replicated by this result. Following Aoshima et al., we interpret this slowdown as the Japanese counterpart of the Filled Gap Effect (Crain, 1985; Stowe, 1986).

The comparison between either the Wh-scrambling condition and the Wh-medial condition, or between the Wh-medial condition and the Wh-nominative condition does not show a slowdown at the second dative NP position. The comparison between the wh-scrambling and wh-medial conditions is problematic. The results that we found in the relative clause experiments are not replicated. We do not have a clue for why the difference between the wh-medial condition and wh-scrambling condition was not significant. On the other hand, the comparison between

the wh-medial condition and the wh-nominative condition suggests that the dative wh-phrase in the medial position is not interpreted in the embedded clause, replicating Kamide and Mitchell's finding and compatible with Aoshima's results.

The significant slowdown at the embedded subject position (region 3) in the wh-nominative condition is apparently problematic. However, this problem is only apparent. In this condition, both subject NPs are Nominative NPs, but in the wh-scrambling condition the first NP is a topic-marked NP. As is observed in various places, a sequence of two nominative NPs generally creates difficulty in reading (Babyonyshev and Gibson 1995; Gibson 1998; Lewis 1996; Miyamoto 2002; Uehara and Bradley 2002 among others). Thus, it is likely that this slowdown reflects the fact that the condition contains two nominative subjects.

Finally, the Wh-nominative condition did not show a slowdown at the second dative position compared to the other conditions. This result indicates that the referential dative NP located in between two subject NPs is not interpreted as a scrambled NP. This again, supports the claim by Kamide and Mitchell that a dative NP in that position is not associated with the embedded clause. This result indeed supports Kamide and Mitchell's position, but there is a concern regarding the difference between the medial dative wh-phrase and medial referential dative NP. As we have seen, we did not observe a significant difference in reading time between the wh-scrambling condition and the wh-medial condition at the second dative NP position. Because we did not observe a significant difference between the wh-medial condition and the wh-nominative condition either, we cannot state anything definitely, but the above observation may suggest that if the medial dative NP bears

the wh-feature, it may potentially be associated with the embedded clause. However, at this point, we do not have any strong evidence that supports this view.

3.4.2. Experiment 2

This experiment was designed with the goal of investigating how long-distance scrambled referential NPs are processed. It is important to see long-distance scrambled referential NPs, in order to understand what is the motivation for the LDB. The findings in the previous experiment suggest that if a fronted phrase is a wh-phrase, the LDB is created and the Japanese Filled Gap Effect is observed as a consequence. The question here is whether the same result is obtained if the fronted NP does not bear a wh-feature. If the fronted referential NP creates a LDB, it suggests that the wh-feature is not the most crucial factor for the LDB. On the other hand, if the LDB is not created by a fronted referential NP, it suggests that the wh-feature plays a more important role for creating the LDB than the other features that the fronted NP conveys, or at least it plays a crucial role. It may be the case that two features are equally important for LDB.

Participants

Twenty-seven native speakers of Japanese participated in the experiment. All of them were students at Hiroshima University. They were paid $5.00 or its equivalent for their participation in the experiment, which lasted about 30 minutes.

Materials and Design

Experimental materials consisted of 24 sets of sentences with two conditions each, which we refer to as the scrambled condition and the control condition, respectively. (9) shows one set of conditions used in the experiment.

(9) A sample set of experimental conditions

a. Scrambled Condition

Josee-shain-ni[1] senmu-wa[2] shachoo-ga[3]

female-worker-dat[1] senior-director-top[2] president-nom[3]

kaigi-de[4] buchoo-ni[5] shookyuu-o[6]

meeting-at[4] department-director-dat[5] promotion-acc[6]

yakusoku-sita-to[7] osieta.[8]

promice-did-comp[7] noticed[8]

"The senior director notified to a female worker, that the president promised a promotion to the department head at the meeting."

b. Control Condition

Senmu-wa[2] **Josee-shain-ni[1]** shachoo-ga[3]

senior-director-top[2] **female-worker-dat[1]** president-nom[3]

kaigi-de[4] buchoo-ni[5] shookyuu-o[6]

meeting-at[4] department-director-dat[5] promotion-acc[6]

yakusoku-sita-to[7] osieta.[8]

promice-did-comp[7] noticed[8]

"The senior director notified to a female worker that the president promised the promotion to the department head at the meeting."

In both conditions, an embedded clause follows the two main clause NPs at the start of the sentence. The embedded clause contains an overt dative marked NP. The embedded verb is marked with a declarative complementizer.

In both conditions, the dative NP in the embedded clause was the second dative NP in the sentence. In the scrambled condition, if the fronted dative NP is associated with the embedded clause, readers should be surprised to encounter a second dative NP, giving rise to a Filled Gap Effect. On the other hand, if the sentence initial dative NP is not associated with the embedded clause, we do not expect the surprise effect.

In the control condition, the two sentence-initial NPs were matched with the scrambled condition in the respect that there was one dative NP and one nominative NP. As we have discussed earlier, in this condition, we have no expectation that either of these NPs should be associated with the embedded clause. The sentence initial nominative NP cannot be analyzed as a scrambled phrase, and the dative NP in

this position, in between the matrix and embedded subject, is not associated with the embedded clause, as Kamide and Mitchell and our previous experiment suggest. Therefore, the second dative NP in the control condition should not be interpreted as if it were scrambled from the embedded clause.

Procedure

The self-paced reading procedure and the comprehension task were identical in format to that used in Experiment 1.

Data analysis

Analyses were conducted on comprehension task response accuracy, item accuracy, and reading times. All data from participants whose comprehension task accuracy was less than 80% were discarded (n = 4, 12.9 %).

Results

Comprehension accuracy and reading times at each region were entered into a repeated-measures ANOVA, with word order as the within-subject factor.

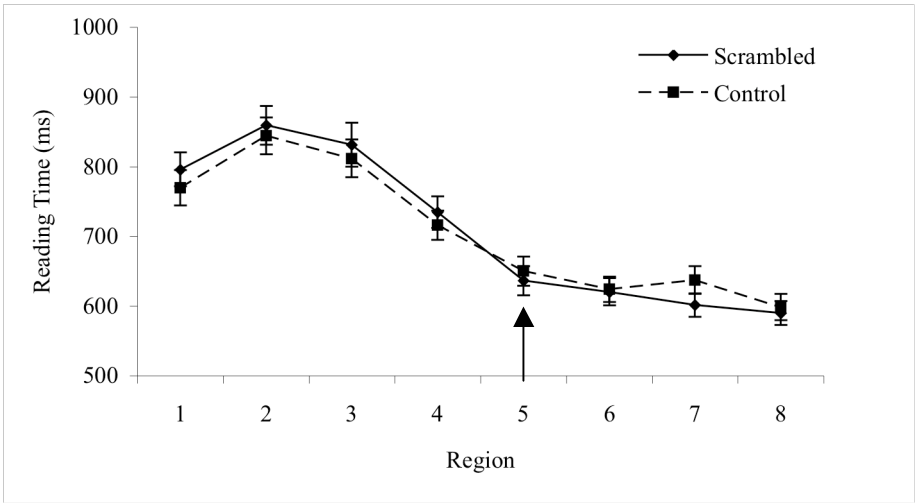
Comprehension task accuracy

Among the participants who were included in the analysis, the average comprehension accuracy was 85.5%. The average correct response percentage did not differ significantly between the two conditions ($F < 1$).

Self-paced reading

Reading times for all regions are shown in Fig 4. At all regions there were no significant differences between reading times in the scrambled and control conditions (Fs < 1).

Fig. 9.



Discussion

The main result of this experiment is that a slowdown in reading time was not observed in the scrambled condition at the embedded dative NP (region 5), relative to the reading time for the same NP in the control condition. We interpret this absence of a slowdown as showing that the fronted NP is not associated with the embedded clause, and therefore did not result in a Filled Gap Effect. This result support Kamide and Mitchell's claim, and also strengthens the view that a nominative NP is not preferentially interpreted as a scrambled phrase, because Nominative NPs cannot be long-distance scrambled in Japanese.

3.5. General Discussion

For our purposes, the most important finding in these two experiments is the contrast between the scrambling of wh-phrases and that of referential NPs with respect to Filled Gap Effect. As should be clear from the discussion above, when the fronted material is a wh-phrase, a Filled Gap Effect is observed. However, if the fronted NP is a referential NP, the Filled Gap Effect is not observed. This contrast suggests that wh-feature on the fronted NP is an important pre-condition for the LDB.

Looking at this contrast from the perspective of the Grammatical Constraint Satisfaction model, the result can be interpreted as follows. The driving force behind longer dependency formation is wh-feature satisfaction. In other words, the parser tries to associate a fronted wh-phrase with the licensing Q-particle as early as possible. As we have discussed, the verb in the most deeply embedded clause is the first verb that can host a Q-particle, so the parser tries to create a gap in the most deeply embedded clause if the fronted phrase is a wh-phrase.

This result clearly contradicts the prediction under the view that the parser's structure building is driven by thematic role satisfaction (Pritchett 1992 among many among others). If the parser tries to satisfy the thematic requirement on the fronted NP, we do not expect the contrast between wh-scrambling and referential NP scrambling because for both cases the most deeply embedded verb is the first position where the thematic relation can be confirmed. On the other hand, the result is compatible with the view that the combination of the wh-feature and thematic role is crucial for the LDB. The above two experiments do not tease these two possibilities apart. Later in this chapter, we will see that there is an environment where we can test

which is the more appropriate view. To test the prediction of these two possibilities (the wh-feature alone or the combination of the wh-feature and the thematic requirement) we have to find an environment in which either the theta role can be assigned but wh-feature cannot be satisfied in the embedded clause or the wh-feature can be satisfied but the theta role cannot be assigned. Shortly we will see that Japanese conditional clauses are indeed an environment that is compatible with the first possibility where the thematic requirement can be satisfied but the wh-feature cannot be satisfied.

Finally let us briefly discuss why a medial wh-phrase is not associated with the embedded clause. The question is why the thematic requirement on the fronted phrase does not seem to drive longer dependency formation. Here I would like to point out one possibility, namely that the thematic role of an NP can be read off from the Case marker. Japanese NPs, in most cases, bear overt Case particles. These Case particles can encode which Case an NP is assigned. However, such information as grammatical functions or thematic roles seems to be imposed on the overt Case particle too. For example, if there is a sequence of NPs that bear Nominative Case and Dative Case respectively, readers can guess which verb may follow. If there is no other material than, say, an Accusative marked NP as in (10a), a verb like *ageru* "give" is most likely to follow these NPs. If, on the other hand, clausal material follows these NPs, the matrix verb is most likely to be a verb like *iu* "say". In both of these cases, however, the theta role of the dative NP is something like the goal or the recipient. Thus, just looking at the Case maker can tell readers a lot about thematic role of the NPs.

(10) a. Walt-**ga** Sue-**ni** ringo-o **ageta.**

W-nom S-dat apple-acc gave

"Walt gave an apple to Sue."

b. Walt-ga Sue-ni [Subject …V-to] itta.

W-nom S-dat V-comp said

"Walt said to Sue that …"

Now, looking at this account from the perspective of the Grammatical Constraint Satisfaction model, it follows naturally that the thematic role does not drive longer dependency formation. When the parser hits the dative NP and the nominative NP, the parser creates a scrambling structure (Aoshima et al. 2004 and Aoshima 2003 among many others). At this point, the structure with the basic word order in Japanese is recovered in which the nominative NP is followed by the dative NP. As we have seen, the thematic role of the dative phrase can be reliably guessed. Thus, the thematic requirement is virtually satisfied at this point, and there is no reason that the parser has to do something more to resolve the thematic requirement of the fronted NP.

Although this is an important discussion to think about the nature of structure building, this point is not directly related to our main purpose, i.e., the motivation for the LDB and the nature of predictive mechanisms. So I will not go into further details about this topic and leave this part informal. For more detailed discussion in the reader is referred to Aoshima et al., (2006).

4. Conditionals and Long-Distance Dependency Formation[29]

This section investigates the following two points. First, we will try to resolve the remaining problem that arose in the previous section about what is the real driving force behind the LDB: whether it is the wh-feature on the fronted NP or whether it is the wh-feature and the thematic role. The second issue is to investigate the nature of the predictive mechanism. In particular, we will investigate what is the trigger of the mechanism and what is the representation that is accessed. In order to tackle these two issues, we will take a detailed look at Japanese conditional clauses.

4.1. The Question

Let us first briefly review our problem. There are several possible motivations for the LDB. We have considered the following possibilities: the thematic requirement, Case licensing, the association of wh-phrases with Q-particles, or a combination of these. Based on two experiments, we drew the tentative conclusion that the wh-feature of the fronted NP is crucial for the LDB. However, the experiments could not tease apart whether the wh-feature is the crucial feature or whether it is the combination of the wh-feature and the thematic requirement that is crucial. This is the primary problem that we are trying to resolve. We pointed out that to test the predictions of each possibility we would need to have an environment where either the theta role can be assigned but the wh-feature cannot be satisfied in

[29] This section is a summary of a poster with the same title (Yoshida et al., 2006) presented at the 19th CUNY Conference on Human Sentence Processing, at the City University of New York, New York. I am grateful to the co-authors Sachiko Aoshima, Hajime Ono and Colin Phillips for allowing me to include this material in this work.

the embedded clause or where the wh-feature can be satisfied but not the theta role. We will see that Japanese Conditional Clauses are appropriate testing grounds for this.

4.2. Japanese Conditionals from a Real Time Perspective

Let us briefly summarize the properties of Japanese Conditional Clauses and why they are appropriate for testing the two competing possibilities. In addition, let us point out another and more important problem for us that Japanese Conditionals raise, namely the problem of the nature of predictive mechanisms. The details of the syntax of Japanese Conditionals are explored in Chapter 4. In this section, we will see three properties that are most directly related to our study.

First, let us point out the non-islandhood of Japanese conditionals. Despite being adjuncts, Japanese conditionals are not islands. Taking a look at the following examples, it becomes clear that scrambling out of conditionals is not degraded compared to the other cases such as extraction from RCs. (11) summarizes examples of extraction out of three different domains: Conditionals, RCs and Complement Clauses. In these examples, scrambling of dative marked wh-phrases is shown because it is most relevant for our purposes, but the same pattern holds for the scrambling of referential NPs.

(11) a. *Dono-gakusee-ni*$_1$ Walt-wa [**(mosi)** Sue-ga t_1
which-student-dat W-top mosi S-nom
purezento-o ageta-**ra**] naki-dasu-no?
present-acc give-cond cry-start-Q
Lit. "Which student will Walt cry if Sue gives the present to?"

b. *Dono-gakusee-ni$_1$ Walt-wa [$_{NP}$[$_{RC}$ Sue-ga t_1
which-student-dat W-top S-nom
ageta] purezento]-ga suki-na-no?
gave present-nom like-be-Q
Lit. "Which student does Walt like the present that Sue gave to?"

c. Dono-gakusee-ni$_1$ Walt-wa[$_{CP}$Sue-ga t_1 purezento-o
which-student-dat W-top S-nom present-acc
ageta-to] omotte-iru-no?
gave-comp think-ing-Q?
Lit. "Which student does Walt think that Sue gave the present to?"

The fact that NPs can be extracted out of conditional clauses means that the extracted NP can receive its thematic role from the embedded verb in the conditional clause.

Second, conditional clauses can be introduced by the optional conditional adverb *mosi* "by any chance" as (11a) shows. Importantly, this adverb can only be licensed by conditional clauses. Thus, if *mosi* is present, the clause introduced by *mosi* must be a conditional clause.

Third, conditional clauses are uniquely distinguished from other clauses by their verbal morphology. In other words, if a clause is a conditional clause, the verb has to have a conditional marker. Conditional markers are suffixes such as *-ra*, *-nara*, *-ba*, *-naraba* and others. These suffixes correspond to the complementizer *if* in English.

The second and the third properties consequently derive the following property that if a clause is introduced by *mosi*, it must have one of the conditional markers on the embedded verb. To capture this relation between mosi and the conditional adverb, we can rephrase their relation in a way like *mosi* is licensed by the conditional marker.

Most importantly, the conditional verb, or the verb bearing the conditional marker cannot host question particles. In a complement clause, the verb can bear a Q-particle if the embedded clause is an interrogative clause. However, the question particle cannot co-occur with the conditional markers even in a case where a wh-phrase is contained in the conditional clause.

(12) a. [mosi Walt-ga nani-o tabeta-ra(*-ka)]

mosi W-nom what-acc eat-cond(*-Q)

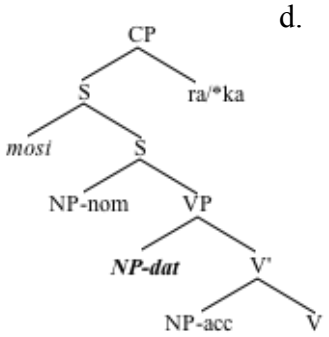
Lit. "if Walt eats what"

b. [$_{CP}$ Walt-ga nani-o tabeta-ka]

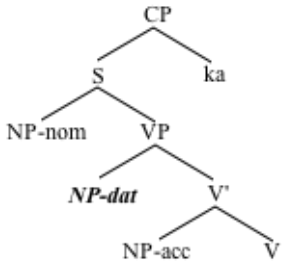
W-nom what-acc ate-Q

Lit. "… what Walt ate"

c.



d.



We can understand this property morpho-syntactially. The question particle and the conditional markers are in complementary distribution. We can capture this complementary distribution by assuming that the question particle and the conditional marker occupy the same slot in the complementizer system.

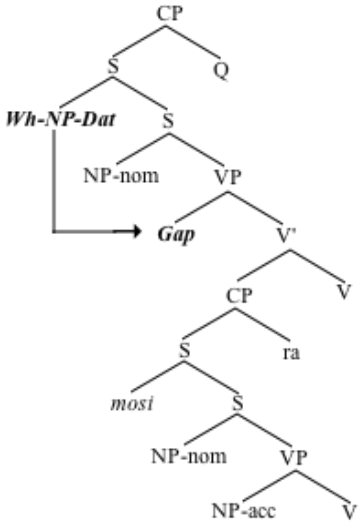
Now it should be clear that these four properties of conditional clauses can provide us with a good testing ground for the question we raised, i.e., which feature is the driving force for the LDB. From the perspective of real time sentence processing, conditional constructions containing wh-scrambling such as (11a) create a situation in which the thematic feature of the fronted wh-phrase can be satisfied at the embedded verb position in the conditional clause, but the wh-feature cannot.

First, in an example like (11a), the long-distance scrambling of the wh-phrase creates a temporary ambiguity with respect to the position of the gap, i.e., the matrix

clause analysis and the embedded clause analysis, as we have seen before in the case of complement clauses. The structure of each analysis is in (13).

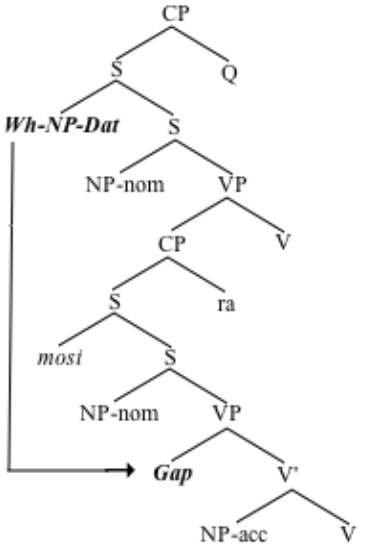
(13) a. Matrix Clause Analysis

Wh-NP-dat Subject [*gap*] [mosi Subject Object V-ra] V-Q



b. Embedded Clause Analysis

Wh-NP-dat Subject [mosi Subject [*gap*] Object V-ra] V-Q



In the same way as in the embedded clauses in our previous studies, the embedded verb is the first verb to be encountered during the processing of a sentence. In an example like (11a), the conditional clause is located after the matrix subject.

The crucial difference from the earlier cases is, however, that in the conditional clauses the embedded verb cannot host the Q-particles. This means that the wh-Q association cannot be established in the conditional clauses. Furthermore, if the conditional clause is introduced by the adverb *mosi*, it must be licensed by the conditional-marked verb, which is not compatible with a Q-particle. Thus, if the presence of *mosi* can signal a conditional-marked verb, and if the speakers can compute this consequence immediately, it becomes clear at the position of *mosi* that the wh-Q association cannot be established in the conditional clause. Assuming that *mosi* can signal the upcoming conditional structure, we can make the following prediction with respect to formation of the longer dependency. If the LDB is motivated by the thematic requirement of the fronted NP, a longer dependency can be formed because the thematic requirement can be satisfied in the conditional clause due to the non-islandhood of conditionals. On the other hand, if the wh-feature drives longer dependency bias, we expect that LDB will be created in a sentence like (11a). As we have discussed, a conditional verb cannot host a Q-particle. Thus, as long as the parser does not attempt yet another embedding, such that an interrogative complement clause is constructed inside the conditional clause, the wh-Q association cannot be established in the conditional clause. Given the difficulty of multiply center-embedded structures in Japanese (Babyonyshev and Gibson 1995 among others), it is not likely that the parser should try to construct another embedded clause inside the conditional clause. Thus, in this case, the LDB is not created due to the fact that wh-Q association cannot be established in the embedded clause.

The account above incorporates two important assumptions. First, the adverb *mosi* can set a prediction of the upcoming conditional clause structure. Second, the prediction or the predicted structure refers to the conditional verbal morphology. Only if these two conditions are fulfilled does the account above hold.

If the LDB is not observed, then the parser makes a prediction for the upcoming conditional verb. If the LDB is observed, then either the prediction is not made, or the hypothesis that the wh-feature is the motivation for the LDB is wrong. If the latter is the case, we are forced to reconsider the interpretation of the experimental results we have seen earlier.

4.3. Experiments

There are four experiments that investigate the points outlined in the previous subsection. Two of them, one offline and one online experiment, are devoted to testing whether the adverb *mosi* helps the parser to predict an upcoming conditional clause. Building upon the results of these first two experiments, the remaining experiments were designed to probe the behavior of the parser in sentences that contain fronted dative wh-phrases and conditional clauses.

4.3.1. Methodology and Predictions

In order to understand the logic behind the experiments, it is always helpful to review the specific predictions along with the methodologies that we employ as we have done before.

The methodologies involved in the first set of experiments are a sentence fragment completion task and a phrase-by-phrase self-paced reading task. In the offline experiment, we provided pairs of fragments of sentences. One member of the pair contained the adverb *mosi*, and the other contained standard adverbs that do not have any special relation to specific constructions. If the presence of *mosi* helps native speakers of Japanese create a prediction for conditional clauses, we expect they complete the fragments with conditional clauses. On the other hand, standard adverbs are not expected to create such a bias because they do not bear a special relation to conditional clauses.

(14) a. Subject mosi Subject_____

b. Subject Adv Subject _____

Basically the same logic applies to the online experiment. If the adverb *mosi* creates a prediction for an upcoming conditional structure during online sentence processing, we expect to observe facilitation effects or surprise effects at the conditional verb position. Because only the verbal morphology can unambiguously distinguish conditionals from other embedded clauses, we expect to see a reading time differences at the conditional verb position. Specifically, we predict that the conditional verb will be read more quickly in the *mosi* condition compared to the control condition, where *mosi* is not provided.

(15) a. Subject [mosi Subject Object Verb-Cond] V

b. Subject [Adv Subject Object Verb-Cond] V

The remaining two experiments are devoted to testing the interaction of long-distance scrambled wh-phrases and the prediction for a conditional clause. The offline experiment is designed to test whether a fronted wh-phrase is preferentially associated with an embedded conditional clause or not. If the wh-phrase is associated with a conditional clause, we expect that sentence fragments should be completed with conditional clauses in which the embedded verbs are those that can take the fronted wh-phrase as one of their arguments. This bias should be observed in the adverb/scrambling condition due to the LDB, which is basically the same type of

condition that Aoshima et al. originally tested. Because there is no special adverb in this condition, it should be preferentially interpreted as a construction containing a complement clause. This can be tested by comparing the completions between the types of fragments in (16).

(16) a. Wh-NP-Dat Subject ***mosi*** Subject _____

b. Wh-NP-Dat Subject ***Adv*** Subject _____

Another online experiment is designed to test whether a fronted wh-phrase is preferentially interpreted in the embedded conditional clause or not. Adopting the Japanese Filled Gap Paradigm, we expect that if the adverb *mosi* creates a prediction for an upcoming conditional verb, and if the driving force for the longer dependency bias is the wh-feature, we do not expect to observe a Filled Gap Effect at the second dative NP position compared to the control condition. Schematically the sentence looks like (17a). However, if the driving force is the thematic requirement, a Filled Gap Effect is expected. Thus, in this way we can pin down the crucial feature that motivates the LDB. On the other hand, if *mosi* is not provided, we do not expect that the conditional clause should be predicted. Thus, we expect to observe a Filled Gap Effect in this case. A schematic representation of the relevant sentences is shown in (17b).

(17) a. Wh-NP-Dat Subject [**mosi** Subject *NP-dat* NP-acc V-cond] V-Q

b. Wh-NP-Dat Subject [**Adv** Subject *NP-dat* NP-ac V-cond] V-Q

4.3.2. Experiment 3: Sentence Fragment Completion Task

An offline sentence fragment completion task was conducted in order to test whether the presence of *mosi* creates a bias toward the generation of conditional clauses. If adverbs like *mosi* work as reliable cues for an upcoming conditional structure, we expect that there should be significantly more conditional clause completions compared to a control condition where such adverbs are not provided.

Participants

40 native speakers of Japanese participated in the experiment. All of them were students at Hiroshima University, Japan, who were paid $5.00 or its equivalent for their participation in the experiment, which lasted about 30 minutes.

Materials and design

Experimental materials consisted of 12 sets of 2 conditions each (the *mosi* condition and the control condition respectively). The fragments in these conditions consisted of a topic-marked NP, an adverb, an adjective and a nominative-marked NP. The *mosi* condition contained conditional adverbs such as *mosi* and others that are only legitimate in the environment of conditional clauses. The control condition, on the other hand, contained a degree adverb that modified the immediately following adjective that was associated with the nominative-marked NP. The nominative-marked NP was provided in order to signal the onset of an embedded clause. Previous studies have shown that a nominative NP can strongly indicate the beginning of an

embedded clause (Inoue 1991; Miyamoto 2002 among others). Thus we expected that all the completions of the target conditions should have involved biclausal structures.

The 12 sets of target items were distributed between two lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with twenty-four filler items in a random order. The filler items were designed in such a way to avoid speakers' potential strategy to use the same structure in all completions.

(18) Sample set of experimental conditions for Experiment 3

a. Mosi Condition

NP-top	*mosi*	*Adj*	*NP-nom*
Sono-buchoo-wa	mosi	hunarena	buka-ga
that-department-chief-top	mosi	uninitiated	subordinate-nom

b. Control Condition[30]

NP-top	*Adv*	*Adjective*	*NP-nom*
Sono-buchoo-wa	shooshoo	hunarena	buka-ga
that-department-chief-top	a-little-bit	uninitiated	sub.-nom

Results

This study yielded a total of 456 codable sentence fragment completions. The results from two participants had to be discarded due to the fact that more than half of their responses were left blank or ungrammatical. The completions were classified

[30] Although, in roman transcription, the number of characters or length of the adverb looks different in the two conditions, in Japanese characters using Kanji characters, the length or number of characters are perfectly matched.

according to the number of conditional clauses used to complete the sentence fragment. The results are summarized in Table 12.

Table 12

Conditions	Clause Type			
	Conditionals		Others	
	N	*%*	*N*	*%*
Mosi Condition	202	88.5	26	11.4
Control Condition	24	10.5	204	89.4
Total	226	49.5	230	50.4

Overall, 226 out of 456 completions contained an embedded conditional clause (49.5%). The proportion of conditional completions was higher in the Mosi condition (88.5%, 202/228 trials) than in the Control condition (10.5%, 24/228 trials). A χ^2 test showed that the difference was significant ($\chi^2(1)=277.95$, $p < 0.01$). This result confirms that *mosi* biases speakers to generate conditional clauses.

Completions classified as "others" in all conditions contained blanks (18.6%, 43/230), simple clauses (7.3%, 17/230) simple complement clauses (49.1%, 113/230 trials), relative clauses (6.9%, 16/230 trials), other adjunct clauses than conditionals (13.9%, 32/230 trials), embedded wh-interrogative construction (0.8%, 2/230 trials) and matrix wh-interrogative construction (0.4%, 1/230 trials).

Discussion

The main finding of this study is that Japanese speakers show a strong preference to generate conditional clauses if adverbs like *mosi* are provided. This result thus supports that claim that *mosi* sets a prediction for an upcoming conditional clause.

4.3.3. Experiment 4: Self-Paced Reading Task

Building upon the finding in Experiment 3 that *mosi* creates a strong bias for the completion of a conditional clause, Experiment 4 was designed with the goal of investigating the time-course of the prediction of conditional clauses. We informally stated that *mosi* creates a bias for conditional clause completion in the previous subsection. More specifically, this means that if *mosi* is provided, the embedded verb bears conditional morphology because only the verbal morphology can distinguish conditional clauses from the other embedded clauses. Now, from the perspective of real-time sentence processing, we expect the following reading-time pattern. If *mosi* sets a prediction for an upcoming conditional verb, a conditional verb in a sentence containing *mosi* should be read more easily than a conditional verb in a sentence that does not contain *mosi,* because the sentence completion measure indicated that other adverbs than *mosi* do not create a bias for a conditional clause completion and thus they are not likely to set a prediction for an upcoming conditional structure.

Participants

Twenty-six native speakers of Japanese participated in the experiment. All of them were students at Hiroshima University, Japan. They were paid $5.00 or its equivalent for their participation in the experiment, which lasted about 30 minutes.

Materials and design

Sixteen sets of two conditions each were used in the experiment, manipulating the type of adverbs (*Mosi* and *Degree Adverb*). The 16 sets of items were distributed between two lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with 72 filler items in a random order. The filler items were matched with the target items in overall length and complexity.

A sample set of experimental items is shown in (19). Again the superscripts indicate regions.

(19) a. *Mosi Condition*

Sono-kakarichoo-wa[1] **mosi**[2] hunare-na[3]
that-senior-director-top[1] **mosi**[2] uninitiated[3]
buka-ga[4] torihikisai-ni[5] songai-o[6]
subordinate-nom[4] customer-dat[5] casualty-acc[6]
ataeta-ra[7] sonoba-de[8] jishoku-o[9]
give-cond[7] immediately[8] resignation-acc[9]
negaideru-daroo.[10]
ask-will.[10]

"That senior director will immediately ask for his resignation if his somewhat inexperienced employee causes casualties to the customer."

b. *Control Condition*

Sono-kakarichoo-wa[1] **shooshoo**[2]

that-senior-director-top[1] **a-little-bit**[2]

hunare-na[3] buka-ga[4] torihikisai-ni[5]

uninitiated[3] subordinate-nom[4] customer-dat[5]

songai-o[6] ataeta-ra[7] sonoba-de[8]

casualty-acc[6] give-cond[7] immediately[8]

jishoku-o[9] negaideru-daroo.[10]

resignation-acc[9] ask-will.[10]

"That senior director will immediately ask for his resignation if his a somewhat inexperienced employee causes casualties to the customer."

In these two conditions, a main clause topic-marked subject was immediately followed by an embedded clause, which began with either an adverb of *mosi*-type or a degree adverb that modified the adjective associated with the embedded subject NP. In both of the conditions, the embedded subjects were preceded by adjectives and the embedded verbs bear conditional morphemes.

In order to ensure that any reading time effects associated with the embedded verb would not be confounded with effect at the main verb, an adverbial phrase and an accusative-marked NP were inserted in between the embedded and matrix verbs.

Procedure

The self-paced reading procedure and the comprehension task were identical in format to that used in Experiment 1 and 2.

Data analysis

Analyses were conducted on comprehension task response accuracy, and reading times. All data of participants whose comprehension task accuracy was less than 80% in total were discarded (n = 4, 13, 3 %). Reading times longer than 2500ms were discarded. This procedure affected 2.0 % of trials. The means and analyses presented below are based on the remaining trials.

Results

Comprehension accuracy and reading times at each region were entered into a repeated-measures ANOVA, with adverb type (*mosi* and control) as within-subject factors.

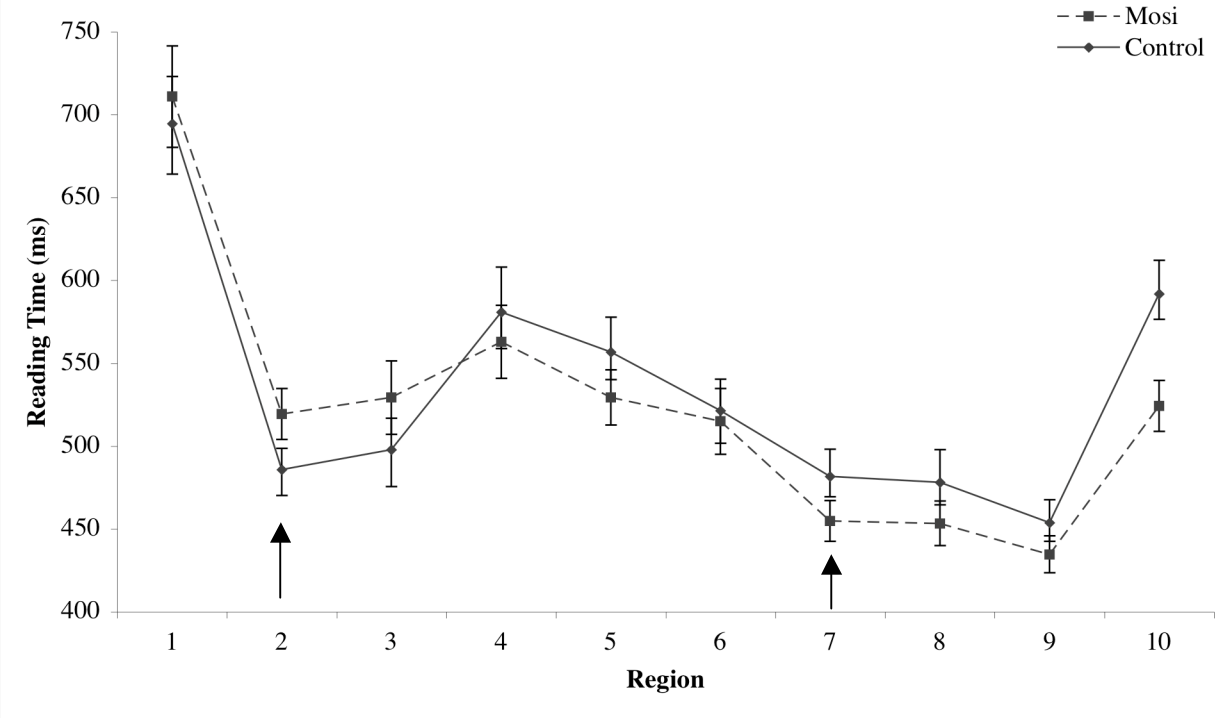
Comprehension task accuracy

Among the subjects who were included in the analysis, the average comprehension accuracy was 85.4 %. The average correct response percentage did not differ significantly across the two conditions ($F < 1$).

Self-paced reading

Reading times for all regions are shown in Fig. 10. Except for the second, seventh and tenth regions, there were no significant differences between reading times in the *Mosi* condition and the control condition (all Fs < 1).

Fig. 10.



At the adverb in region 2, the Mosi condition was read 34 ms more slowly than the control condition The effect was significant both in the participant analysis and item analysis (F1(1,25)=5.62, P < 0.05; F2(1, 15)=4.41, P < 0.05). The region 7 yielded a significant difference in reading time (the difference is 27 ms.) in participant analysis but not in the items analysis (F1(1, 25)=4.56, P < 0.05; F2(1, 15)=1.62, P < 1). This reading patter goes to the opposite direction from that observed

at the region 2. Finally, the last region, region 10, showed a significant difference in reading time (the difference is 186 ms.) both in participant and item analysis ($F1(1,25)=11.25$, $P < 0.05$; $F2(1, 15)=5.3$, $P < 0.05$).

Region 7, the critical region for us, which contains the conditional verbs, yielded a significant difference in reading time in the participants analysis but not in the items analysis. Although there is a possibility that region 8, the adverb region, reflect spill over effect from the immediately preceding region, we did not obtained a significant effect there either. However, still the control condition yielded numerically slower reading times than the *Mosi* condition. Taking this fact into consideration, we calculated the reading time differences between the two conditions by combining the two regions, region 7 and region 8. Our expectation is if these two regions crucially reflect the effect of the manipulation of *Mosi*, we would be able to gain a more powerful effect by combining these two regions. However, the result did not reach a significant effect in the item analysis ($F1(1, 25)= 4.86$, $P < 0.05$; $F2(1, 15)= 1.47$, $P < 0.2$).

Discussion

There were two main findings in this study. First, the conditional adverbs of *mosi*-type were read slower than the degree adverbs. Second, this reading pattern reversed at the embedded verb region, where the *Mosi* condition was read more quickly than the control condition. We interpret this result as the reflection of the prediction of the upcoming conditional structure by means of the adverb *mosi*. In other words, the adverb *mosi* effectively sets a prediction for an upcoming conditional

structure, and as a result the verb bearing the conditional morphology was read more easily. The slowdown at the adverb *mosi* reflects either the fact that it a rare adverb compared to degree adverbs, or the fact that conditional clauses are expected. In any case, the overall reading time profile is best understood as showing that *mosi* provides an effective cue for an upcoming conditional structures. Note that this result accords well with the result from the offline sentence fragment completion experiment. Taken together, we can conclude that *mosi* is an effective cue for an upcoming conditional structure.

Although the final region also yielded a significant effect, it is difficult to infer anything from this region because it is the matrix verb region and the end of the sentence where the so-called wrap-up effect typically takes place. For this reason, it is not so clear whether this region reflect the manipulation on the adverb at the beginning of the embedded clause.

4.3.4. Experiment 5: Self-Paced Reading Task: Japanese Filled Gap Paradigm

Now the stage is set for investigating our main issue, i.e., testing the motivation for the LDB, and the representation employed by the predictive mechanism.

Building on the findings in Experiment 3 and 4 that the conditional adverb *mosi* provides an early warning for an upcoming conditional structure, Experiment 5 was designed with the goal of investigating whether fronted wh-phrases in Japanese are preferentially associated with an embedded conditional clause or not.

The finding in the previous experiment that conditional verbs were read faster in the *mosi* condition than in the control condition indicates that the abstract syntactic structure of the conditional clause and the morphology of the conditional verb were foreseen by the parser, by means of the presence of *mosi*. However, it is not obvious whether the consequences of the conditional morphology can also be anticipated, specifically the fact that the conditional verbal morphology is not compatible with Q-particles.

What we expect is the following. If the parser calculates the consequence of the prediction of the upcoming conditional verbal morphology, and if the driving force for longer dependency formation is the wh-feature we should not observe the LDB in conditional clauses. If on the other hand, either of the above hypotheses is not true, we may observe the LDB in conditional clauses. Even if it is the case that the wh-feature is the driving force behind the LDB, if the parser cannot calculate the consequence of the prediction, it is also possible that the parser may try to associate fronted wh-phrases with the embedded clause, expecting that a Q-particle may be present on the embedded verb. On the other hand, if the driving force is the thematic feature, rather than the wh-feature, it is again possible that the parser might try to associate a fronted wh-phrase with the embedded clause in order for the fronted phrase to receive its theta role, because conditional clauses are not islands for long-distance scrambling (see Chapter 2 for details).

In order to test these competing hypotheses, this experiment adapted the Japanese Filled Gap paradigm. We used Filled Gap effects to probe for gap creation.

If the gap is created, we will observe Filled Gap Effect by providing a lexical dative phrase in the embedded clause.

Participants

Sixty native speakers of Japanese participated in the experiment. All of them were students at Hiroshima University, Japan. They were paid $5.00 or its equivalent for their participation in the experiment, which lasted about 30 minutes.

Materials and design

Twenty sets of four conditions each were used in the experiment, in a 2 X 2 factorial design, which manipulated the case-marking of the wh-phrase (Dative vs. Nominative), and the type of adverb (*Mosi* vs. Degree Adverb). The 20 sets of items were distributed among four lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with 72 filler items in a random order. The filler items were matched with the target items in overall length and complexity.

A sample set of experimental items is shown in (20).

(20) a. *Dative Mosi Condition*

Dono-joosi-ni[1] sono kachoo-wa[2] mosi[3]
which-boss-dat[1] that senior-director-top[2] mosi[3]
mijukuna[4] buka-ga[5] torihikisaki-ni[6]
uninitiated[4] subordinate-nom[5] customer-dat[6]
songai-o[7] ataeta-ra[8] sugusama[9] jishoku-o[10]
casualty-acc[7] give-cond[8] immediately[9] resignation-acc[10]
negaidemasu-ka?[11]

ask-Q[11]

"Which boss will the senior director ask for the resignation immediately if his somewhat inexperienced employee causes casualties to the customer?"

b. *Nominative Mosi Condition*

Dono-joosi-ga[1] sono kachoo-ni[2] mosi[3]

which-boss-nom[1] that senior-director-dat[2] mosi[3]

mijukuna[4] buka-ga[5] torihikisaki-ni[6]

uninitiated[4] subordinate-nom[5] customer-dat[6]

songai-o[7] ataeta-ra[8] sugusama[9] kubi-o[10]

casualty-acc[7] give-cond[8] immediately[9] resignation-acc[10]

iiwatasimasu-ka?[11]

tell-Q[11]

"Which boss will fire the senior director immediately if his somewhat inexperienced employee causes trouble to the customer?"

c. *Dative Adverb Condition*

Dono-joosi-ni[1] sono kachoo-wa[2] shooshoo[3]

which-boss-dat[1] that-senior-director-top[2] little[3]

mijukuna[4] buka-ga[5] torihikisaki-ni[6]

uninitiated[4] subordinate-nom[5] customer-dat[6]

songai-o[7] ataeta-ra[8] sugusama[9] jishoku-o[10]

casualty-acc[7] give-cond[8] immediately[9] resignation-acc[10]

negaidemasu-ka?[11]

ask-Q[11]

"Which boss will the senior director ask for the resignation immediately if his somewhat inexperienced subordinate causes troubles to the customer?"

d. *Nominative Adverb Condition*

Dono-joosi-ga[1] sono kachoo-ni[2] shooshoo[3]

which-boss-nom[1] that-senior-director-dat[2] little[3]

mijukuna[4] buka-ga[5] torihikisaki-ni[6]

uninitiated[4] subordinate-nom[5] customer-dat[6]

songai-o[7] ataeta-ra[8] sugusama[9] kubi-o[10]

casualty-acc[7] give-cond[8] immediately[9] resignation-acc[10]

iiwatasimasu-ka?[11]

tell-Q[11]

"Which boss will fire the senior director immediately if his somewhat inexperienced employee causes troubles to the customer?"

In all conditions the matrix clause material at the beginning of the sentence were followed by either *mosi* or a degree adverb. The embedded subject in all conditions was preceded by an adjective. The embedded subjects were all nominative-marked so that they could provide a strong cue for the onset of an embedded clause. In order to ensure that any reading time effects associated with the embedded verb would not be confounded with effects at the main verb, an adverbial phrase and an accusative-marked NP were inserted between the embedded and matrix verbs. The embedded verb in all conditions bore a conditional morpheme. These adjustments were adopted from experiment 4.

The only differences among conditions involved the Case-marking of the wh-phrase and the type of adverb at the beginning of the embedded clause. In the Dative conditions, the topic marked matrix subject NPs were immediately followed by the embedded clause material, either *mosi* or a degree adverb. On the other hand, in the Nominative conditions, the matrix subject was located at the beginning of the sentence, and the dative-marked NPs were immediately followed by the embedded clause material, again either *mosi* or a degree adverb.

Procedure

The self-paced reading procedure and the comprehension task were identical in format to that used in Experiment 1, 2 and 4.

Data analysis

Analyses were conducted on comprehension task response accuracy and reading times. All data from participants whose comprehension task accuracy was below 80% in total were discarded. 8 participants failed to meet this criterion (11.8%). Trials read slower than 3000ms were discarded, corresponding to less than 2.3% of trials. The means and analyses presented below are based on the remaining trials.

Results

Comprehension accuracy and reading times at each region were entered into a repeated-measures ANOVA, with Case-marking (Dative, Nominative) and adverb (*mosi*, degree adverb) as within-subjects factors.

Comprehension task accuracy

Among the 50 participants included in the analysis, the average comprehension accuracy was 85.5%. Mean accuracy scores did not differ significantly across the four conditions (all Fs < 1).

Self-paced reading

The reading time analysis yielded the following results. Reading times for the *Mosi* conditions are shown in Fig. 11 and those for Degree Adverb conditions (Adverb conditions) in Fig. 12.

Fig. 11.

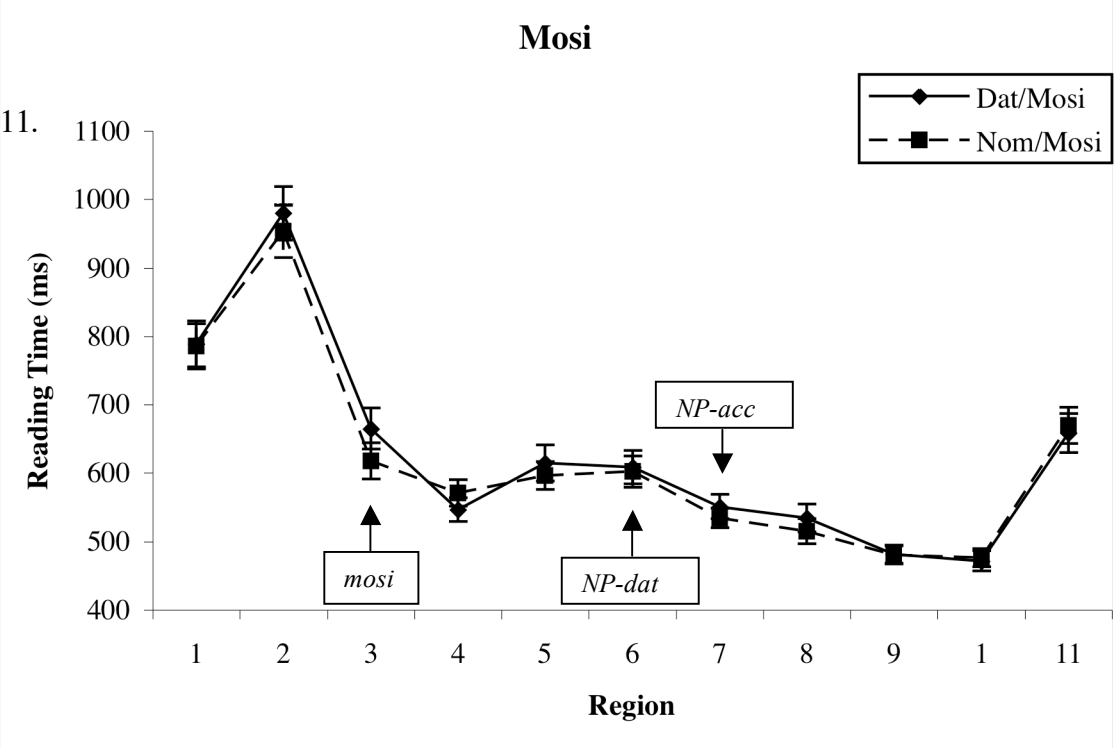
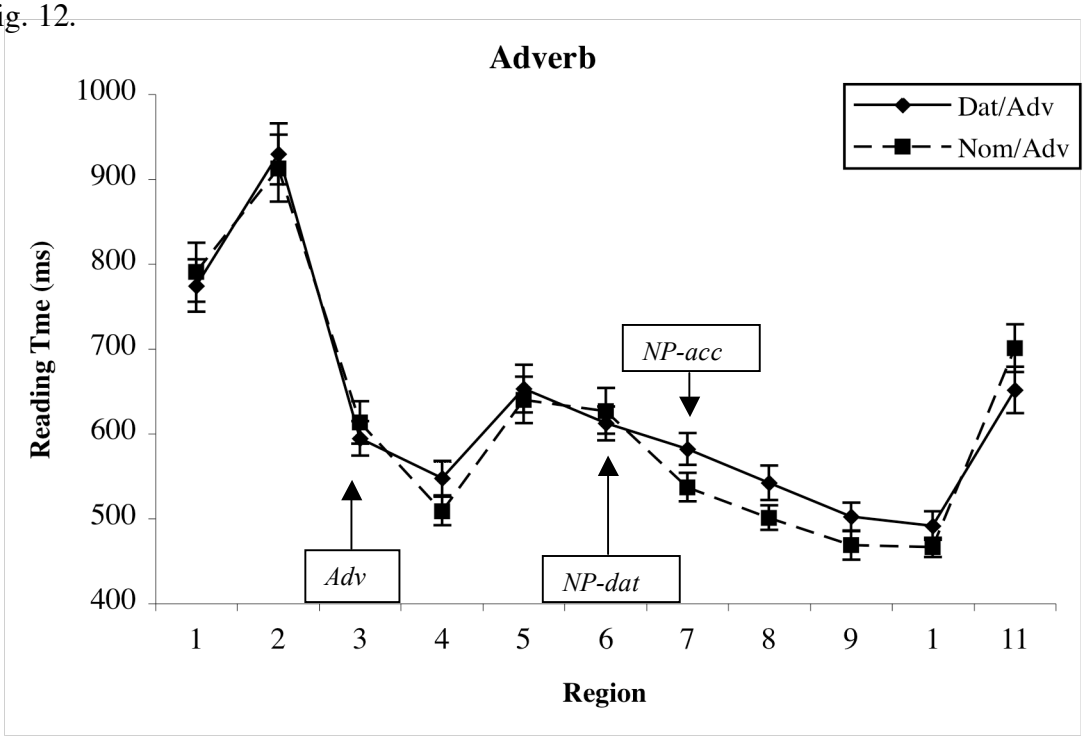


Fig. 12.



At all regions prior to the seventh region (the accusative NP region), there were no significant differences between reading times in the Dative and Nominative conditions (all Fs < 1). Comparisons between the Mosi and Adverb conditions did not yield any significant difference either (all Fs < 1).

Although there was no significant effect at the critical region (region 6) (all Fs < 1), at the following regions, region 7 and region 8, we observed a slowdown, reflected in longer reading times for the embedded dative NP only in Adverb conditions.

At the accusative-marked NP (Region 7), there was a significant main effect of Case-marking in the participants analysis, and items analysis showed marginally significant effect (F1 (1, 59) =4.17, P < 0.05; F2; (1, 19) = 4.06, P = 0.06). The

interaction of Case-marking type and Adverb type was not significant (all Fs < 1). Pairwise comparisons revealed that within the Adverb conditions the accusative-marked NPs were read significantly more slowly in the Dative condition than in the Nominative condition. The effect was significant in participant analysis, and in the item analysis there was a marginally significant effect (F1(1, 59)=6.44, P< 0.01; F2(1, 19)=3.49, P = 0.06). On the other hand, the same comparison for the *Mosi* conditions showed no corresponding slowdown (all Fs < 1).

Because there are no difference among conditions other than the Case-marking of the wh-phrases and the type of adverbs, it is not likely that the slowdown that we observed reflects factors other than our own manipulations. Thus, we interpret this slowdown as a spill-over effect of a Filled Gap Effect created by the dative NP in region 6, the critical region.

The embedded verb position (the region 8) also yielded a significant main effect, but only in the participants analysis (F1 (1, 59) = 4.35, P < 0.05; F2 (1, 19)= 2.45, P = 0.15). The interaction of Case-marking type and Adverb type was not significant (all Fs < 1). Pairwise comparisons showed that within the Adverb condition the embedded verbs were read slower in the Dative condition than in the Nominative condition. The effect was significant in the participants analysis, and in the items analysis, the effect was marginally significant (F1 (1, 59)=5.0, P < 0.05; F2 (1, 19)=2.33, P =0.07). Again the same logic applies to this region. There is no reason other than the manipulation we made in the target items that may cause the slowdown. One might be concerned with the fact that the region contains the embedded verb. Remembering, however, that all the conditions have conditional

verbs at the embedded verb position, it is unlikely that the processing of the conditional verb was responsible for any of the differences in reading times. Thus, again we interpret this effect as a spill-over effect that carried over from region 6, the critical region.

Given that it is likely that the reading time slowdown at these two regions reflects a Filled Gap Effect, we may see a clearer effect if we can combine these two regions. The combining of region 7 and the region 8 showed a significant main effect of the Case-marking type ($F1\ (1, 59) = 10.2$, $P < 0.01$; $F2\ (1, 19)=5.13$, $P < 0.05$). Pairwise comparisons revealed that within the Adverb conditions the combined region was read more slowly in the Dative Condition than in the Nominative Condition ($F1(1, 59)=10.05$, $p< 0.01$; $F2(1, 19)=4.17$, $p < 0.05$). Because all the items in these regions were the same across-the-board, it was not problematic to combine these two regions. Therefore we take this effect as indicating that a Filled Gap Effect took place at the critical region (region 6). Note that the combination of the same regions in the comparison in the Mosi conditions did not yield a significant difference in reading times (all $Fs < 1$). This result indicates that in the *Mosi* conditions, the second dative NP did not create a surprise effect.

Other subsequent regions did not yield any significant effects (all $Fs < 1$).

Discussion

There were two findings in this experiment. First, the degree adverb associated with the embedded subject basically did not affect the longer dependency formation bias. Even if the embedded subject NP is modified by an adjective with a

degree adverb, the fronted dative wh-NP is preferentially associated with the embedded clause. Thus, we basically replicated Aoshima an colleagues results. Although we obtained slightly weaker result than Aoshima et al.'s original study, our study supports the view that when a construction contains a fronted dative wh-phrase, it is preferentially associated with the most deeply embedded clause.

Second, the presence of *Mosi* blocks the LDB. If *Mosi* is present at the beginning of the embedded clause, the fronted phrase is not preferentially interpreted in the embedded clause (the conditional clause). This result suggests the following two points. First, the results support the hypothesis that the longer dependency is formed solely because of the parser's desire to establish a Wh-Q association. In other words, the parser does not create a gap in the embedded clause if it cannot support the Q-particle on the embedded verb. Second, the parser appears to foresee that the embedded verb of the conditional clause cannot host the Q-particle by means of the conditional adverb *Mosi*. In other words, upon encountering *Mosi* the parser predicts the structure of the upcoming conditional clause, and as a result it can recognize that the embedded verb bearing the conditional form cannot host the Q-particle. If this account is correct, it implies that the parser is equipped with a powerful predictive mechanism that allows it to pre-process the structural skeleton of an upcoming conditional clause because it predicts the morphology of the embedded verb which reflects a complex structural relation among V, I and C in the conditional clause structure (see the discussion in Chapter 4 for details), not just the phrase structure.

4.3.5. Experiment 6: Sentence Fragment Completion Task

An additional offline sentence fragment completion task was conducted with three purposes. The first purpose was to examine whether the presence of *Mosi* regulates the bias for the generation of a conditional clause structure even if the sentence initial wh-phrases are provided in the fragments.

The second purpose was to provide an additional test of Japanese speakers' preference not to produce a Q-particle in the embedded clauses if *Mosi* is provided, even in the cases where a wh-phrase is provided in the fragments.

The third and last purpose of this study is to provide another assessment for testing whether fronted wh-phrases are preferentially associated with the embedded clauses if *Mosi* is provided in the fragments.

By testing these three points, we can provide further support for the findings in the previous experiment.

Participants

52 native speakers of Japanese participated in the experiment. All of them were students at Hiroshima University. They were paid $5.00 or its equivalent for their participation in the experiment, which lasted 30 minutes.

Materials and design

The experimental materials consisted of 12 sets of 4 conditions each. The Four conditions independently manipulated the type of Case-marking for wh-phrases

(Dative vs. Nominative) and the type of Adverb (Mosi vs. Degree adverb). A sample set of experimental conditions for this study is summarized in (21)

(21) a. *Dative/Mosi*

Dono-buka-ni sono buchoo-wa mosi
which-subordinate-dat that department-chief-top mosi
hijoosiki-na sinnyuu-shain-ga …
thoughtless new-recruit-nom

b. *Dative/Adverb*

Dono-buka-ni sono buchoo-wa
which-subordinate-dat that department-chief-top
shooshoo hijoosiki-na sinnyuu-shain-ga …
little thoughtless new-recruit-nom

c. *Nominative/Mosi*

Dono-buka-ga sono buchoo-ni mosi
which-subordinate-nom that department-chief-top mosi
hijoosiki-na sinnyuu-shain-ga …
thoughtless new-recruit-nom

d. *Nominative/Adv*

Dono-buka-ga sono buchoo-ni
which-subordinate-nom that department-chief-top
shooshoo hijoosiki-na sinnyuu-shain-ga …
little thoughtless new-recruit-nom

All four conditions contained a wh-phrase at the beginning of each fragment. The Dative condition began with a dative-marked wh-phrase, and the Nominative condition with a nominative-marked wh-phrase.

In the Dative conditions, the dative-marked wh-phrase was followed by a topic-marked NP. The topic-marked NP was followed by either Mosi or by a degree adverb that modified the adjective that immediately followed it. On the other hand, in the Nominative conditions the nominative-marked wh-phrase was followed by a dative-marked NP, and the remaining material was the same as in the Dative conditions. In all four conditions, each fragment ended with a nominative-marked NP. Participants were asked to complete the remaining part of each fragment.

The phrases used in the fragments were taken from the target items used in Experiment 5. The 12 sets of items were distributed among four lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with thirty-six unrelated items in a random order. The filler items were designed in such a way that we could prevent speakers from adopting a strategy of using the same structure in all completions.

Results

This study yielded a total of 476 codable sentence fragment completions. The completions were classified according to the number of conditional clauses used to complete the sentence fragment, the position of the question particles that indicated the scope of the wh-phrase, and the argument structure of the embedded clause verb. Responses were classified as conditional clauses if a conditional-marked verb was

provided. Responses were classified as embedded questions if exactly one question marker on the embedded verb was provided. A Q-particle could also be placed on the main verb or on both the main and embedded verbs. Responses were also coded based on the argument structure of the verbs provided in the completions, in order to provide an additional criterion for the interpretation of the dative-marked wh-phrase. If the wh-dative was interpreted in the embedded clause, for example, there should be a verb that can take the dative NP as its argument.

In the analysis of embedded clause type, 61.3% of fragments (292 trials) were completed as embedded conditional clauses with a conditional morpheme on an embedded clause.

There were large differences across conditions in the number of conditional clause responses. The proportion of trials in which a conditional verb was provided in the embedded clause for the Dative/Mosi condition was 86.1% (124/144 trials), for the Dative/Adv condition was 16.6% (24/144), for the Nominative/Mosi condition was 87.5% (126/144), and for the Nominative/Adverb condition was 12.5% (18/144). Other completions in the *Mosi* conditions contained either blank or ungrammatical completions in which no conditional marker was provided. A χ^2 test showed that the proportion of embedded conditional completions was significantly different between the *Mosi* conditions and Adverb conditions ($\chi^2(1)= 300.502$, $p<0.01$). χ^2 tests showed that there were significant difference between the Dative/Mosi condition and the Dative/Adverb condition ($\chi^2(1)=138.996$, $P<0.01$), between the Dative/Mosi condition and the Nominative/Adverb condition ($\chi^2(1)=156.086$, $P<0.01$), between the Nominative/Mosi condition and the Dative/Adverb condition ($\chi^2(1)= 144.751$,

P<0.01), and between the Nominative/Mosi condition and the Nominative/Adverb condition ($\chi^2(1)= 162.0$ P<0.01). There were no other reliable differences in the other comparisons.

The proportion of conditional clauses provided in the completions in each condition is summarized in Table 13.

Table 13

Conditions	Embedded Clause-Type			
	Conditional		Others	
	N	%	N	%
Dative/Mosi	124	86.1	20	13.8
Dative/Adverb	24	16.6	120	83.3
Nominative/Mosi	126	87.5	18	12.5
Nominative/Adverb	18	12.5	126	87.5
Total	292	50.6	284	49.3

In the analysis of question type, 9.7% of fragments (46/476 trials) were completed as indirect questions, with a question marker on an embedded verb only. There were differences across conditions in the number of indirect question completions. The proportion of trials in which a question particle was provided on the embedded verb was 1.56 % for the Dative/Mosi condition (2/128 trials), 27.0% for the Dative/Adverb condition (26/96 trials), 2.27% for the Nominative/Mosi condition (3/132 trials), and 13.3% for the Nominative/Adverb condition (15/112 trials). χ^2 tests

showed that the proportion of embedded question particle completions was significantly different between the Dative/Mosi condition and Dative/Adverb condition ($\chi^2(1)=39.304$, $P<0.01$), and between the Dative/Adverb condition and the Nominative/Mosi condition ($\chi^2(1)=37.407$, $p<.01$). The other comparisons did not show significant differences. Results of the question-type analysis are summarized in Table 14.

Table 14

Conditions	Question Type					
	Embedded		Main		Both	
	N	*%*	*N*	*%*	*N*	*%*
Dative/Mosi	2	1.56	112	87.5	0	0
Dative/Adverb	26	27.0	62	64.5	19	13.0
Nominative/Mosi	3	2.27	120	90.9	0	0
Nominative/Adverb	15	13.3	102	91.0	12	8.33
Total	46	9.82	396	78.8	31	5.38

In order to check where the Dative Wh-phrase is interpreted, especially in the *Mosi* conditions where Q-particles are not used in the embedded clause, we analyzed verb argument structure. In this analysis, we included those trials in which a Q-particle was provided on the matrix verb only. This reflects the fact that in the *Mosi* conditions if the Q-particle was not provided in the matrix clause, such trials were

likely to be ungrammatical because the Wh-phrase must be licensed by the Q-particle, and the possibility that a Q-particle was provided in the embedded clause is quite low.

In the analysis of verb argument structure, 39.7% of completions with conditional morphemes contained a dative argument taking embedded verb (116/292). This proportion was higher in the Dative conditions (96/292) than in the Nominative conditions (20/292), a difference that was significant ($\chi^2(1)=79.208$, $P<0.01$). χ^2 tests showed that there were significant differences in the proportion of the dative-taking conditional verbs between the Dative/Mosi condition and the Nominative/Mosi condition ($\chi^2(1)=40.648$, $P<0.01$) and between the Dative/Mosi condition and the Nominative/Adverb condition ($\chi^2(1)=13.820$, $p<.01$). The other comparisons did not show significant differences. The significant difference between the Dative/Mosi condition and the Nominative/Mosi condition is informative for us. This strongly indicates that the wh-dative NP in the *Mosi* conditions is preferentially associated with an embedded clause. Totals for the analyses of verb argument structure are shown in Table 15.

Table 15

Counts and percentages of verbs that select a dative-marked argument in an embedded conditional

Condition	Embedded Only		Main Only	
	N	%	N	%
Dative/Mosi	84	67.7	24	22%
Dative/Adverb	12	50.0	4	30%
Nominative/Mosi	16	12.6	44	89%
Nominative/Adverb	4	22.2	14	53%
Total	116		86	

Discussion

There were three main findings in this study. First, the effect of *mosi* in biasing the sentence generation toward conditional clauses was again observed. This is a replication of experiment 3. Thus, we can confidently conclude that *mosi* creates a strong bias toward upcoming conditional clauses.

Second, it is not so a strong measure, but the Question type analysis suggests that Wh-datives tended to be associated with the embedded clause. Specifically the comparison between the Dative/Adverb condition and the Nominative/Adverb condition is the most informative in this respect. The analysis showed that more Q-particles were produced in the completions in the Dative/Adverb condition than in the Nominative/Adverb condition. This suggests that the dative-marked wh-phrase was more likely to be interpreted in the embedded clause compared to a condition that does not have such possibilities.

The analysis of Q-particle also supports a relation between the conditional morphology and the Q-particle. The fact that the *Mosi* conditions contained very few embedded Q-particles suggests that *Mosi* worked as a reliable cue for the upcoming conditional clause, and that Japanese speakers do not try to embed an additional clause inside conditional clause in order to host a Q-particle associated with the fronted wh-phrase either, supporting our earlier speculation.

These two findings help us to fulfill two of the purposes of this study. First, we could confirm that *Mosi* indeed creates a strong bias toward conditional clauses. Second, Fronted wh-phrases seem to be preferentially associated with an embedded clause. Note, however, that the conclusion that we can draw from the analysis of Q-

particle positions was not strong because we could not find a significant difference between the Dative/Adverb condition and the Nominative/Adverb condition, the comparison we used in the online experiments.

Let us think about the third purpose of the study. Based on the results of our Filled Gap study, we expected that the offline experiment would also show the same tendency, i.e., the fronted dative-marked Wh-phrase should not be preferentially associated with the embedded clause if *Mosi* provides an early warning for an upcoming conditional clause structure because it is not compatible with the motivation for the LDB. However, the analysis of argument structure revealed that this was not the case in the offline sentence completion task. In the Dative/Mosi condition there were significantly more dative taking verbs in the completed conditional structures. This effect cannot be related to the presence of *Mosi* alone, because it exists in the Nominative/Mosi conditions also but they are not completed with embedded verbs that a take dative argument. Rather, this effect should be interpreted as being created by the fronted wh-dative phrases. This result directly contradicts the finding in the Filled-Gap experiment.

Remember, however, that we have emphasized that Japanese conditional clauses are not islands for long-distance scrambling based on native speakers' acceptability judgments. As long as native speakers' intuitions are reliable, we have a way to interpret a seemingly problematic result in our offline experiment. First of all, if conditional clauses are not islands as the acceptability judgments suggest, it is not surprising if even a long-distance scrambled phrase is interpreted in the conditional

clauses. Rather, the situation in online experiment is surprising from this point of view. Now what account can we give about this online offline contrast?

Although informal, we can give the following reasoning for why a longer dependency might not be blocked in the offline setting. There seems to be a crucial difference between online sentence processing and offline judgments or sentence completions. That is, the difference in resources that readers can access. Typically, in online sentence processing, the parser can access only limited resources. Because in an online self-paced reading context readers have to read sentence from left-to-right, they cannot re-access the left-context once each phrase is read. Thus, if a sentence contains dependent elements such as displaced elements or elements that must be associated with another elements, readers have to store them in memory during the processing of the remaining part of the sentence, until they are licensed (Gibson 2000; Lewis 2001; among many others).

In this online situation, it is plausible if the parser tries to establish a dependency involving these elements as soon as possible, so as to release these elements from working memory. If the licensing relation is easy for the parser to infer, due to, for example, an overt morphological marking, the parser can establish a licensing structure and release the element from working memory. For example, we have seen that the thematic requirement on the fronted dative phrase seems to be established easily by the parser. What we have discussed was that the thematic role of the dative phrase can be read off from the overt Case-marking and basic word order in Japanese sentences. Thus, the parser does not need to associate the fronted NP with the lexical verb in order to confirm the actual thematic role that it receives, as long as

the thematic requirement is concerned. The findings from our first two experiments also suggest the same point.

On the other hand, if the licensing structure is not obvious from the current input items alone, the situation becomes different. Wh-Q-particle association is just such a case. As we have seen, the scope of a wh-phrase cannot be determined based on the position of the wh-phrase. Rather it is fixed by the Q-particle that is affixed to the verbs. Because of this, the word order, or the Case-marker does not help the parser to establish the licensing structure for a wh-phrase. The presence of the lexical verb is crucially necessary, and the embedded verb is always the first verb to be encountered. Thus, the parser that has a desire to satisfy grammatical features of each lexical item tries to associate the wh-phrase with the embedded verb in order to confirm its scope at the earliest point. By doing so, the parser can release the wh-phrase from working memory as early as possible.

The intuition behind our conclusion in the previous subsection was if *Mosi* is provided in a sentence, then a wh-Q-particle dependency should not be established in the embedded clause because of the conditional morphology. In this case the most deeply embedded verb is not the first position for Q-particle, tather the verb that immediately follows the embedded verb is the first possible host for Q-particle.

Going back to the offline situation, during the offline judgment or sentence completion task, native speakers can access much richer resources than in the online situation. They have more time for reading each sentence or they can even repeat the sentence and examine the structure in detail. In this situation, native speakers do not need to establish grammatical relations in such a hurry. There is no strong driving

force urging them to establish sentence structure. If this is the case, we lose the reason to block longer dependency formation even when *Mosi* is provided.

If such requirement for earlier dependency formation is motivated by resource limitations, then in the offline context, the dependency blocking effect is also not motivated. Thus, it is not surprising that the longer dependency is not to be blocked.

However, this reasoning does not provide an indication of why longer dependency seems in fact to be preferred in the Mosi condition. This is clearly a puzzling situation, and we do not have any clear answer at this point.

5. General Discussion: The Silhouette of the Predictive Mechanism

5.1. The Nature of the Predictive Mechanism (or the Things that the Parser should be Able to do)

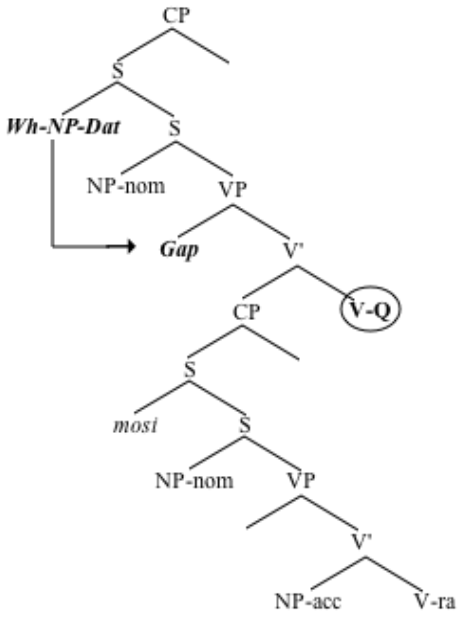
At the outset of this chapter, we laid oud the aims of this study. They were: [i] to investigate the real motivation behind the LDB; and [ii] to investigate the nature of the predictive mechanism underlying the human sentence processor. The second aim can be rephrased as the following two questions. (i) What counts as a potential trigger for the predictive mechanism (the trigger problem)? (ii) What is the representation that the predictive mechanism generates (the representation problem). There is another important problem: What is the algorithm employed by the predictive mechanism (the algorithm problem)? We will discuss the algorithm problem in Chapter 5 in detail. So this problem will be put aside temporarily. In this section, we will pay attention to the first three problems: the motivation for the LDB; the possible triggers for the predictive mechanism; and the representation that the predictive mechanism accesses.

First let us talk about the motivation for the LDB. The results from three online experiments suggest that the LDB is solely motivated by the wh-feature on a fronted dative-marked wh-phrase. Experiments 1 and 2 revealed that if the wh-feature is stripped from the fronted dative NP, a longer dependency was not formed. Experiment 5 further strengthened this view.

In experiment 5, we created a situation where the wh-feature on the fronted materials cannot be satisfied by the most deeply embedded verb bearing the

conditional morpheme. We manipulated the parser's prediction for upcoming structures by means of the conditional adverb *Mosi*. If *Mosi* is present, it provides an early warning for an upcoming conditional clause, and as a consequence of this early warning, the LDB is blocked. Under the assumption that the parser avoids extra embedding whenever possible, and if the most deeply embedded verb cannot bear the Q-particle, the earliest possible position where the verb bearing the Q-particle can appear is the position immediately following the most deeply embedded verb. The schematic structure illustrated in (22) clearly shows this point.

(22)



In a structure like (22), upon encountering *mosi* the parser notices that the embedded verb cannot bear a Q-particle, and that the earliest position where the Q-particle can appear is the matrix verb position. Additionally, the grammar allows wh-Q-particle association at the first gap position, thus there is no reason that the parser should

further stretch the dependency to go into the embedded clause. If we do not take this account, it is not clear why the fronted wh-phrase should not be associated with the embedded clause even though conditional clauses are not islands.

One might argue that conditional clauses show a mild island effect, and that this mild island effect is sufficiently strong to block the LDB. However, this account has a shortcoming. In the offline experiment, Experiment 6, we have found some evidence that the fronted wh-phrase is preferentially associated with the embedded conditional clause. The mild island account thus cannot explain this offline result as well as native speakers' judgments.

Taken together, this study shows that the LDB is solely motivated by the wh-feature conveyed by the fronted dative Wh-phrase.

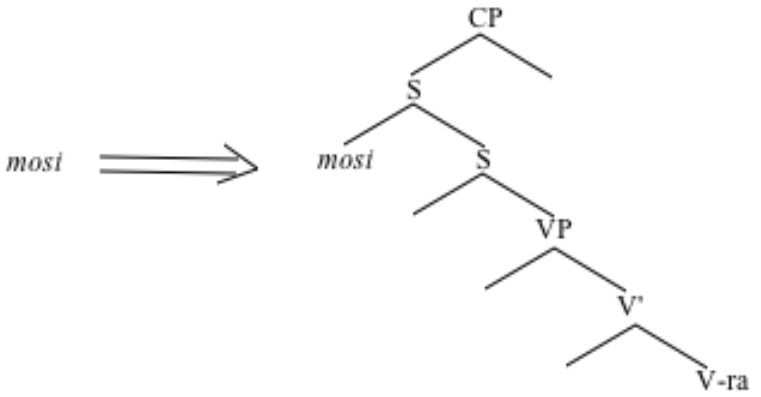
In the previous discussion, we informally proposed that *mosi* creates a prediction for an upcoming conditional clause structure. However, we left two important questions unresolved: how the predictive mechanism is triggered and what representation is built by the predictive mechanism? These are the questions we will try to answer here. Let us discuss "the trigger problem" first. This is the question of what kind of information is used to create a prediction for an upcoming structure? In the previous chapter, we discussed that the parser can use a relatively indirect cue to anticipate the upcoming structure, such as the Classifier Mismatch. In this study on conditional clauses, we have found that a dependent element such as a conditional adverb like *mosi* can be used to foresee an upcoming conditional clause. This prediction process is quite straightforward compared to the case of Classifier Mismatch.

As is extensively discussed in Chapter 4, syntactically *Mosi* has a strong relation with the conditional clause. As we have discussed in various places, *Mosi* can only be licensed in a conditional clause. Therefore, if *Mosi* is present, the conditional clause must be present too. Thus *mosi* can signal an upcoming conditional structure. We will see more formal and detailed discussion on this point in the later chapter, Chapter 5.

We have understood that *mosi* can create a prediction of some sort for an upcoming conditional structure. Now the question is what is actually predicted, i.e., the representation problem. A series of experiments that we have described have made it obvious that the predicted structure must be sufficiently informative that it can block LDB. However, it is not clear what kind of structural representation counts as a sufficiently informative representation. The key to revealing this point lies in the motivation for the LDB. As we have seen, to block the LDB, it must be clear that the verb cannot host a Q-particle. This, in turn, means that when the LDB is blocked, the parser calculates that a Q-particle cannot be hosted by the embedded clause. This information is morphological information. Assuming that *Mosi* is structurally licensed by the conditional morpheme (see Chapter 4 for a detailed discussion), the structure built by the predictive mechanism should represent the phrase structure of the conditional clause, and the morphology of the conditional verb that licenses *Mosi* should not host the Q-particle. If this is true, the predictive mechanism should be able to build the full phrase structure of a conditional clause, going all the way down to the head of VP where the conditional morpheme suffixes and all the way up to CP where the conditional morpheme *-ra*, a complementizer, ultimately sits. Both the VP

structure and the CP structure seem to be necessary. Because of its affixal nature, the conditional morpheme requires a verb as a host. A CP is also necessary because the conditional suffix *-ra* or its variants is understood as a head of CP. Thus if *-ra* is predicted, the CP projection should also be predicted.

(23)



Finally a question arises, namely what kind of sentence processing algorithm allows such prediction of the upcoming structure? We will explore this question in Chapter 5

5.2. Reconsidering the Processing of Relative Clause Islands (or the Things that the Parser should be Able to Foresee)

So far, we have seen what the parser should be able to do when an adverb like mosi is provided. In the discussion above, we have seen that the prediction of verbal morphology is a necessary component of the blocking of the LDB. Assuming that this morphological prediction is true, it raises an interesting question about the apparent online island effect created by RCs. In the previous chapter, we concluded that LDB

was blocked in the environment of predicted RCs because RCs are islands in Japanese. However, the above discussion on conditional clauses casts doubt on this view of the online processing of RC islands.

Looking closely at the morphological property of RC verbs, it becomes clear that RC verbs also cannot host a Q-particle. As discussed in Chapter 4, Japanese RC verbs must bear the so-called Predicate Adnominal form (Hiraiwa, 2000; Hiraiwa, 2001), and the Predicate Adnominal form is not compatible with the Q-particle. The incompatibility of the Adnominal form and the Q-particle can be understood in a way that the empty C and Q-particle are competing the same slot, exactly in the same way as the conditional morpheme and Q-particle do.

In Japanese, if a CP is an immediate constituent of an NP, the embedded verb inside the CP must bear the Adnominal form. Based on this observation, we can interpret the online island effect of RCs from a different perspective, i.e., because the Adnominal form is predicted, the LDB is blocked. We have seen that a Genitive Numeral Classifier (GNC) requires a host NP. Therefore, the GNC can potentially create a prediction for an upcoming NP. On the other hand, a Nominative NP can create a prediction for a clausal projection (Miyamoto 2002). Presumably S/IP and CP nodes are predicted by means of a nominative NP due to the fact that a nominative NP is typically licensed by a finite *Infl* (Chomsky 1981, 1986a, 1986b; Takezawa 1987), and the fact that the finite IP is dominated by a CP (Chomsky 1986a, 1986b). If a Classifier Mismatch takes place, the two predicted nodes, the NP and the CP are to be connected to each other in some way, under models of sentence processing incorporating a fully incremental structure building mechanism (Aoshima et al. 2004;

Bader and Lasser 1994; Crocker 1996; Schneider 1999; Stabler 1994; Sturt and Lombardo 2005 among others). If so, the predicted CP must be incorporated into the structure as a complement or an adjunct (an immediate constituent) of the predicted NP. In either way, if it becomes clear that the CP must be attached to the NP, the grammar of Japanese tells the parser that the embedded verb inside the CP must bear the Adnominal form, and thus the CP must be headed by an empty C. Assuming that the parser can calculate the consequences of the prediction, i.e., that the Q-particle cannot appear in the predicted CP because it has to host an empty C, the parser has sufficient information to stop creating the longer dependency. This is so because the wh-feature cannot be satisfied at the embedded verb position because the Q-particle cannot appear there. The reading time difference at the embedded verb position that we have observed in the second experiment in Chapter 2, indeed supports this view. Because the verbal morphology is predicted, a reading time difference appeared at the embedded verb position. Under this account, again the parser has to predict all the way down to the VP projection and all the way up to the CP projection upon encountering the classifier mismatch. However, if the parser incorporates a predictive mechanism that can predict the verbal morphology in an RC, we can derive the online RC island effects. Under this account, the blocking of the LDB in RCs and conditional clauses can be understood in the same way.

Note, however, that the same question arises as we asked in the previous section, namely whether there is an algorithm that allow the prediction that we have illustrated so far or not. In the next chapter, Chapter 5, an algorithm will be proposed that can appropriately capture the blocking of the LDB.

CHAPTER 5. ON THE NATURE OF PREDICTIVE MECHANISMS

1. Introduction

In this chapter, we are concerned with the remaining questions involving the nature of the predictive mechanism, namely the Algorithm Problem. So far we have seen some evidence that the parser is equipped with a powerful predictive mechanism that can predict detailed upcoming grammatical structures. Our intuition is that the parser projects a detailed syntactic structural skeleton based on bottom-up information. What we have seen was that the parser is able to use indirect or direct cues to build the upcoming structure. Furthermore, by looking through the cases where the Longer Dependency Bias (LDB) is blocked, we have investigated the representation that the predictive mechanism generates. The remaining problem for us is what is the actual mechanism of sentence processing that can derive these predictions. In this chapter, a parsing algorithm will be proposed that can explain the predictive mechanism of sentence processing as well as various well-known properties of the human sentence processor.

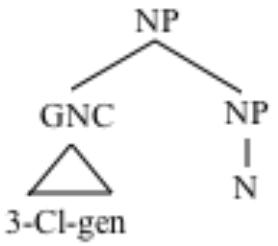
2. The Requirements for the Predictive Mechanism

In Chapters 2 through 4 we have seen aspects of the predictive mechanism underlying the sentence processor through detailed investigations on Japanese Relative Clauses (RCs) and Conditional Clauses. In the course of the discussion we have paid special attention to two aspects of the predictive mechanism. One is the trigger of the predictive mechanism. What kind of bottom-up information is used to

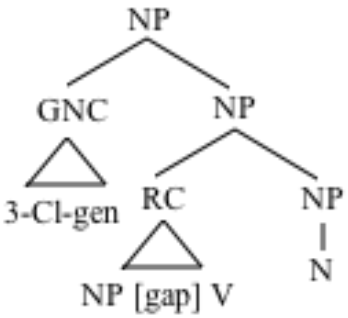
initiate a prediction for upcoming structures. We have seen two types of triggers: indirect cues and direct cues.

When the parser predicts upcoming head-final RCs in Japanese a relatively indirect cue is used. We have confirmed in the earlier chapter that a Classifier Mismatch that does not have any direct grammatical relation to a specific construction like RCs can indeed allow the parser to predict an upcoming RC structure. Let us first clarify the specifics of the classifier mismatch cue. As we have seen, a Genitive Numeral Classifier (GNC) must be associated with its host NP. We assumed the structural relation between the GNC and its host NP shown in (24). Because of the semantic requirement of the classifier, the head of the host NP must be a noun that is semantically compatible with the classifier. Furthermore, we assumed that GNC and the host NP must be structurally, not linearly, adjacent to each other. Between the GNC and NP, an RC can intervene as in (25), but it does not affect these requirements of the GNC, and also the presence of the GNC does not affect the relation between the RC and the NP that it modifies. The semantic compatibility between the classifier and the head of the host NP and the structural requirement that is illustrated in (24) are two requirements that the GNC has to fulfill. As should be clear, there is no specific relation between the GNC and the RC. Both are optional modifiers for the NP, and neither requires the presence of the other.

(24)



(25)



Only in the type of configurations in (25), however, the classifier mismatch can come about. In this configuration, the classifier and its linearly adjacent NP, for example the subject of the RC, can semantically mismatch, and the semantic mismatch is only possible in this type of complex NP configuration. And when the parser recognizes this mismatch, the parser notices that the mismatch can only be legitimate in the complex NP configuration such as (25), and this information leads to the prediction of the RC structure. NEither the classifier nor the NP that is accidentally adjacent to the classifier bears a direct signal for the RC itself. Rather, the semantic mismatch is the crucial signal that allows the parser to foresee the upcoming RC structure. For this reason, the cue from the classifier mismatch is indirect, i.e., it has no direct relation to the RC structure.

In the study of conditional clauses, we have discovered another type of trigger, i.e., the conditional adverb *Mosi*. This is a more direct cue for the predictive mechanism. We have seen that grammatically *Mosi* stands in a tight one-to-one

relation with the conditional clause. If *Mosi* is there, a conditional clause must be present because *Mosi* can only be licensed in the conditional clause. By means of this tight grammatical relation, *Mosi* allows the parser to predict an upcoming conditional clause structure.

Now let us turn to the representation built by the predictive mechanism. We have discussed that both in the case of RCs and conditional clauses, the representation created by the predictive mechanism must be something that can block longer dependency bias. In the case of RC prediction, this seemed to be straightforward at first glance. Because RCs are islands in Japanese a longer dependency that enters an island domain is not legitimate. Thus, a possible approach was that a representation that can encode the islandhood of RCs is built by the parser, e.g., the representation of subjacency. However, the studies on conditional clauses provided us with another way that can lead us to the online island effect.

The studies on conditional clauses revealed that the LDB is blocked during online sentence processing because *Mosi* sets a prediction of the upcoming conditional morphology, which is not compatible with the Q-particle. A series of experiments revealed that the real motivation for the LDB is wh-Q-particle association. Thus, if the conditional verb cannot host the Q-particle, there is no reason to pursue the longer dependency. Thus, prediction of the embedded verbal morphology is crucial for blocking the LDB.

Within the approach where we put an import on the morphological aspect of the predicted structure, there is a possibility that the online island effect of the RCs can also be treated in the same way. A detailed syntax of RCs has revealed that the

embedded verb in RCs also bears special verbal morphology, the so-called Predicate Adnominal form (Hiraiwa 2000, 2001 among others). What we have confirmed was that the Adnominal form is also not compatible with the Q-particle. Thus, if we think that the Classifier Mismatch leads to the prediction of the Adnominal morphology, the fact that the LDB is not seen in the predicted RC structure may be sufficiently explained.

In the following discussion, I will summarize both subjacency approaches and morphology based approaches.

2.1. Subjacency Approaches

In this subsection, let us summarize advantages and disadvantages of subjacency-based approaches.

First we have to note that it does not seem to be the case that all the complex NPs are islands in Japanese. For example the following contrast suggests that a complement clause of a noun is not an island or at least not as strong an island as RCs.

(26) a. Noun Complement Clauses

Dono-spy-ni tantee-wa [$_{NP}$[$_{CP}$ hannin-ga joohoo-o
which-spy-dat detective-top suspect-nom information-acc
nagasita toiu] shuchoo]-o tenkai-sita-no?
leaked comp claim-acc made-Q?

Lit. "Which spy did the detective make a claim that the suspect leaked the information to?"

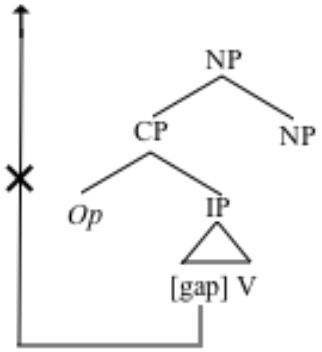
b. Relative Clauses

*Dono-spy-ni tantee-wa [$_{NP}$[$_{CP}$hannin-ga nagasita]joohoo]-o
which-spy-dat detectie-top suspect-nom leaked info-acc
nyuushu-sita-no?
obtained-Q?

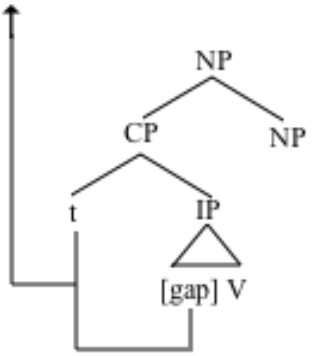
Lit."Which spy$_1$ did the detective obtain the information$_2$ that the suspect leaked t_2 to t_1?"

This contrast suggests that structure specific to RCs is crucial to the islandhood of RCs. The fact that a domain is a complex NP is not sufficient to derive its islandhood. The crucial difference between RCs and noun complement clauses is whether operator movement is involved or not (for the details about movement approaches to Japanese RCs, see Chapter 2). In a regular head-external RC in Japanese, empty operator movement is crucially involved in its derivation. Thus, in head-external RCs there is a gap that corresponds to the relative head. On the other hand, noun complement clauses do not involve operator movement of the type involved in RCs. Given this difference, we can derive the islandhood of the RCs from a bounding theory incorporating the classical subjacency condition (Chomsky 1973, 1981; Rizzi 1982 among many others), in which NP and IP are the bounding categories and movement crossing two of these nodes in one step creates a subjacency violation. Specifically, assuming that the specifier of CP in head-external RCs is occupied by an empty operator corresponding to the argument gap inside the RC, A'-movement out of an RC cannot pass through the CP-spec position and it has to move across IP and NP in one step. On the other hand, assuming that in the noun complement clauses, the CP-spec position is available for A'-movement, and thus the subjacency violation can be side-stepped.

(27) a. Relative Clauses



b. Noun Complement Clauses



With this discussion in mind, let us think about subjacency approaches to online RC island effects. In order for subjacency approaches to derive the online island effect, the presence of the empty operator in the CP-spec position is crucial. Thus, if the existence of two bounding nodes and the presence of the empty operator in the CP-spec are predicted, then the islandhood of the upcoming structure is also predicted. This is equivalent to assuming that the full structural skeleton of the RC structure, including the empty operator, is projected when the upcoming RC structure is predicted.

Although incorporating such information is possible, there are some difficulties with this type of approach from the perspective of online structure building. The most crucial challenge is the difficulty in recognizing the location and the identity of the gap during parsing In other words, until the verb's argument structure or the relative head becomes available, sometimes the parser cannot determine whether the gap is contained in the RC or which argument slot corresponds

to the gap. To understand this problem we have to understand the typology of RCs and also the ways in which an RC can be processed.

There are various types of RCs in Japanese. Above all, one of the most eccentric types of RCs is the so-called Gapless RCs and their variants. In Japanese, certain RCs do not contain an argument gap position (Hoshi 1995, 2004; Inoue 1976; Kamio 1983; Kuno 1973; Kuroda 1974, 1975/76, 1976/77, 1999; Matsuda 1993; Murasugi 1991 among many others). Let us see some examples.[31] The important property of the examples in (28a) and (29a) is that the strings constituting the embedded clauses can be stand-alone matrix clauses if they are not incorporated into the NP environments.Put differently, no argument seems to be missing in the embedded clauses, and thus they do not contain explicit gap positions.

(28) a. [$_{NP}$ San-satu-no [$_{NP}$[$_{RC}$ yuumee-na dezainaa-ga
3-cl-gen famous designer-nom
Rosa(-no-tame)-ni cover-o dezain-sita]hon]].
R(-for)-dat cover-acc designed book
"Three books which a famous designer designed its cover to for Rosa."

b. Yuumee-na dezainaa-ga Rosa(-no-tame)-ni
famous designer-nom R(-for)-dat
cover-o dezain-sita.
cover-acc designed.
"A famous designer designed the cover (for something) for Rosa."

[31] In this study, we will loosely classify gapless RCs as RCs that superficially do not contain a gap that corresponds to the relative head. Thus, we include some RCs that are not analyzed as gapless RCs in Japanese syntax.

(29) a. [$_{NP}$ San-nin-no [$_{NP}$[$_{RC}$ Sidney-ga Rosa-ni
three-cl-gen S-nom R-dat
otooto-o shookai-sita] yuujin]].
younger-brother-acc introduced friend
Lit. "Three friends whose brother Sidney introduce to Rosa."

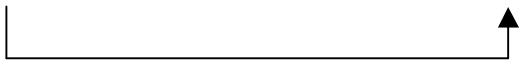
b. Sidney-ga Rosa-ni otooto-o shookai-sita.
S-nom R-dat younger-brothers-acc introduced
"Sidney introduced his younger brothers to Rosa."

The possibility of gapless RCs creates potential problem for incremental structure building under subjacency-based approaches. When these RCs are read from left-to-right, it is not clear where the gap is located or which argument could correspond to the gap in the RCs until the parser encounters the embedded verb or the head of the relative clauses because there is always a possibility that the predicted structure turns out to be a gapless RC like (28) or (29) (see Yamashita et al. 1993; Yamashita 1995 for related discussion). Even though the classifier mismatch can indicate that some NP is located outside of the embedded clause, it does not help the parser to recognize the gap in these cases because, as we have seen, even in these cases the gapless RC can be there.

Let us examine the point above more in detail under the subjacency-based account. Under the subjacency-based approach, what the parser has to do when it encounters the classifier mismatch is to insert an empty operator in the embedded CP-spec position. If the empty operator is inserted, the parser can project the structure that can derive the subjacency effect. Because the CP-spec position is occupied, the

displaced elements at the beginning of the sentence cannot be associated with the embedded clause without violating the subjacency condition. Thus, as long as online structure building respects the mental grammar, the LDB should be blocked.

(30) Wh-NP-Dat … [$_{NP}$ GNC [$_{NP}$[$_{CP}$ *Op* [$_{IP}$ Subject …]] *NP$_{(host)}$*]



This approach, however, involves an important assumption. That is, the operator indicates the presence of a gap position corresponding to an argument inside the embedded clause.[32] Therefore, under this approach the parser's initial choice upon encountering the classifier mismatch is a gapped RC analysis rather than a gapless RC analysis. From this it follows that if the RC turns out to be a gapless RC, it is predicted that there will be a surprise effect associated with properties related to gapless RCs, e.g., Filled Gap Effect because all of the argument slots are filled by overt lexical elements. Although it has not been tested whether this prediction is correct or not, it does not seem that there is an independent reason that the gapless RC analysis should not be an initial choice for the parser. Specifically given that examples like (29) are quite natural (they are read as natural as gapped RCs), and do not seem to induce any difficulty at least at the level of intuitive judgment, it is not clear why the gapless RCs should create any problems. Of course, we have to

[32] Note that the semantic information from the classifier also suggests that there is an argument gap. For example the classifier for *books*, can tell the parser that the relative head NP is headed by a noun *book*. However, this information is not so decisive because the type of the RCs as in (28) exists in Japanese.

consider various factors such as the frequency of such gapless RCs or acceptability differences between gapless RCs and gapped RCs, and careful conclusion must be drawn. Although we do not have any experimental results that can tell us whether gapless RCs are problematic or not, there is no compelling argument for the view that the parser's initial choice should be gapped RC analysis, and we do not have any strong evidence that supports it either. If so, it is plausible to assume that the gapless RCs do not create any troubles. If it is true, we can say that those two options are equally plausible as first choices for the parser. Under this hypothesis, what is the representation that the predictive mechanism should create?

It is possible to Insert an empty operator corresponding to the argument gap upon encountering the classifier mismatch. However, in this case the potential error in choosing a gapped RC analysis arises. Given this the safest strategy is to insert the operator after encountering an explicit bottom-up information that can tell the parser which argument is missing in the embedded clause. Thus under this approach, it is possible that the parser initially cannot project the representation that can satisfactorily represent the islandhood of the RC structure. The islandhood of RC becomes clear only after the verb or other input confirms which argument slot can be the gap position. Thus, this approach predicts that until the embedded verb position or the relative head position a sentence containing a classifier mismatch could be treated as non-island. This approach, in turn, predicts further that even in the classifier mismatch condition, the Filled Gap effect could take place, contrary to our findings.

To maintain a subjacency-based approach, we have to assume that an empty operator is inserted in anyways, without considering the possibility of gapless RCs. This is not totally desirable option but not at all implausible option.

2.2. Morphology-Based Approaches

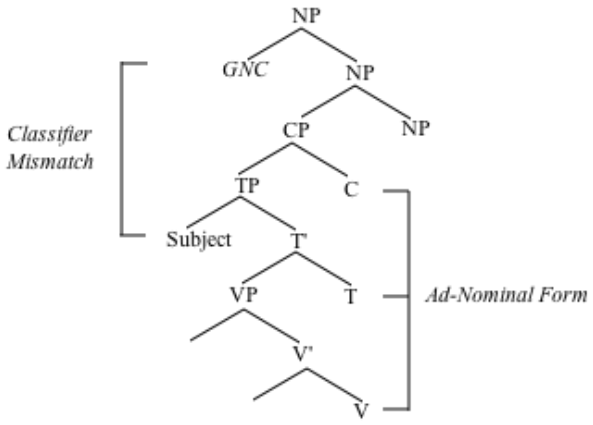
In this subsection, we will turn to Morphology-Based Approaches and discuss their advantages and disadvantages.

Let us turn to the other alternative, the morphology-based approach. Under the morphology-based approach, the LDB is blocked if the predicted embedded verb is not compatible with the Q-particle, which can confirm the scope of the fronted wh-phrase. Thus, this approach does not appeal to a subjacency-based account. In other words, even if the predicted upcoming structure does not fully represent the structure of the RC, the blocking effect of the LDB is expected. The question for this approach is what is necessary for the predictive mechanism to predict the upcoming verbal morphology.

As we have seen in Chapter 2, the Predicate Adnominal form requires an empty C that is created by the morphological amalgamation of V, T and C (Hiraiwa 2000, 2001). Thus, it is sufficient for the parser to recognize that the Adnominal form requires the empty C in order to block the LDB. If the Complementizer is occupied by the empty C, it follows that the Q-particle, which is another instance of Complementizer, cannot appear in the same slot. Additionally, they are only legitimate if the constituent containing the Adnominal-Inflected verb is an immediate constituent of an NP (when they are structurally adjacent to an NP in the structure of adjunction or the complement of an NP). For these reasons, the information that the

predicted structure is incorporated into the structure of NP in some way is crucial for the Adnominal morphology to be predicted. To fulfill these requirements, what the parser has to project upon encountering the classifier mismatch is the full-fledged phrase structure of CP headed by the empty C that is incorporated into the structure of NP either as an adjunct or as a complement.

(31)



The result of Experiment 2 in Chapter 3 indeed supports this claim. In the online reading experiment in Experiment 2, we obtained the result that the embedded RC verb was read more easily when a classifier mismatch is provided compared to the control condition. If the upcoming verbal morphology was not predicted, we would not expect such a result at the embedded verb position. In other words, because the morphological form of the verb is predicted, the embedded verb was easier to read.

There are several advantages of this morphology-based approach. First of all, this approach does not appeal to the presence or absence of an empty operator to explain the online islandhood of the predicted RCs. Thus, the potential problems associated with the empty operator do not arise. Second, this approach predicts that if the morphology of the embedded verb that is not compatible with the Q-particle is

predicted by the parser, the longer dependency bias is blocked. This means that we can explain the blocking of the LDB in the predicted RCs and Conditionals in the same fashion. In the case of conditional clauses, the prediction of the verbal morphology was crucial too. Also, this approach predicts that the longer dependency is blocked in Noun Complement clauses too, if they are predicted in some way because they have the same Adnominal morphology on the embedded verb. Third, as long as the parser has a mechanism that can predict the upcoming verbal morphology upon encountering the classifier mismatch, this approach does not sacrifice incrementality. This, however, does not naturally follow fron anything. Thus we need to call for an explicit predictive mechanism. This problem leads us to the Algorithm problem.

2.3. On the Requirements for the Predictive Mechanism: Summary

In the discussion so far, we have compared two potential approaches to the online island effect of the predicted RCs. We have seen several advantages of the morphology-based approach over the subjacency-based approach.

3. Perspectives on Online Sentence Processing: A Quest for Incrementality

In the previous section, we have raised the question of what type of algorithm allows the parser to predict the structures of RCs or Conditional clauses that satisfy the requirements that we have illustrated above. To answer this question, we will first review a general property of human sentence processing, incrementality, and discuss parsing algorithms that have been proposed in the literature. The main purpose of this discussion is to specify what is a psychologically plausible design for the human sentence processor that can accommodate the predictive mechanism that we have discussed.

3.1. Incrementality

So far we our discussion loosely presupposed the notion of incrementality. In this subsection, let us briefly summarize the notion of incrementality of human sentence processing.

Current evidence from psycholinguistic research suggests that human sentence processing is largely incremental, in the sense that syntactic structures are built and interpretations become available on a word-by-word basis without delay (Marslen-Wilson 1973). So-called garden path phenomena (Bever 1970; Pritchett 1988 among many others) have often been understood as evidence for incrementality. Garden path phenomena tell us that when the parser faces an ambiguous input, it does not delay its decision until crucial disambiguating information becomes available. Rather the parser makes a commitment to an initial analysis that may turn out to be an incorrect analysis. When the parser is garden-pathed, consciously or experimentally detectable

surprise effects arise at the point where the initial misanalysis is resolved. Let us take an example from English. The string before the main verb *sank* in (1) can be parsed as a simple independent clause. When the main verb is encountered, a surprise effect arises, and readers experience difficulty in reading. If the parser waits until the element that can confirm the structure of the sentence, this garden path effect is not expected. The fact that readers experience the surprise effect suggests that the parser commits itself to the initial main clause analysis before the main verb is encountered, and the verb *sank* forced the parser to reanalyze the structure from the main clause structure to the reduced relative clause structure, resulting in the garden path effect.

(32) The boat floated down the river sank.

To account for the garden path phenomena, many researchers emphasized the importance of the licensing head. For example, the theory proposed by Abney (1987) or Pritchett (1992), the so-called head-driven parsing, explained the garden path phenomena in a way that the sentence processing is driven to satisfy the requirements of the licensing heads, typically verbs. Under this approach, an example like (1) creates garden path effect because the materials before the verb *sank* can satisfy all the requirements that the verb *floated* conveys, such as thematic requirements.

Head-driven strategies imply that in a head-final language, e.g., Japanese, the processor waits until the final word of a phrase before building the phrase structure and making a commitment to an analysis. In other words, head-driven theories predict that the input materials are not integrated into a syntactic structure before the

licensing head becomes available in the input (Pritchett 1991b, 1992a, 1992b among others). However, there are various lines of evidence against this particular view either from head-final languages like Dutch (Frazier 1987) or Japanese (Aoshima et al. 2004; Inoue 1991; Inoue and Fodor 1995; Mazuka and Itoh 1995; Miyamoto 2002) or even from a head-initial language like English (Sturt and Lombardo 2005). Let us take a look at some examples from Japanese.

Inoue (1991) cites the following example from Japanese.

(33) Sidney-ga Rosa-ni ringo-o tabeta inu-o ageta.
S-nom R-dat apple-acc ate dog-acc gave
"Sideny gave Rosa the dog which ate the apple."

Japanese native speakers experience some difficulty dealing with this sentence when they reach the verb *tabeta* "ate". This difficulty in reading is not expected if the parser delay the structure building until it encounters the verb. Because, when the embedded verb "ate" becomes available, its argument structure information also becomes available. Based on the argument structure information, the parser can notice that the previously encountered three NPs are not compatible with the verb's argument slots, and conclude that the string should be analyzed as a relative clause rather than a simple clause. Thus, the difficulty indicates that the parser commits an analysis before it encounters the embedded verb. Based on the case-particles that each NP bears, the three NPs are initially postulated as coarguments of the same clause. When the

embedded verb is encountered, it becomes clear that the initial analysis is failed, resulting in the garden path effect.

Inoue's argument is based on native speakers' intuitive judgment, but some experimental studies support the same point too. Miyamoto (2002) found that the upcoming relative clause structure can be facilitated when a sentence contains a sequence of NPs with the same Case Markers. He examined the following pair of sentences.

(34) a. Ofisu-de shokuin-ga kakaricho-o[$_{RC}$ t_1 ocha-o
office-at employeee-nom manager-acc tea-acc
dasita] **josei-ni** teineini shookai-sita.
served woman-dat politely introduced
"At the office, the employee politely introduced the manager to the woman who served the tea."

b. Ofisu-de shokuin-ga kakaricho-ni[$_{RC}$ t_1 ocha-o
office-at employeee-nom manager-dat tea-acc
dasita] **josei-o** teineini shookai-sita.
served woman-acc politely introduced
"At the office, the employee politely introduced the woman who served the tea to the manager."

Miyamoto makes two important observations. First, when there are two accusative NPs in a sentence as in (34a), the second accusative NP creates a disruption. Second, in (34a), the relative head noun was read more easily than in (34b). According to

Miyamoto these two observations are not explained by head-driven approaches. Under the head-driven approaches, the disruptive effect of the second accusative NP is not predicted because these NPs come into the input before the verbs and thus they should not be processed. Furthermore, head-driven approaches do not predict the facilitation effect at the relative head position in(34a). In both of the conditions in (34), everything is the same except for the accusative NPs. Thus, the facilitation effect cannot be attributed to any other factors than the presence of the two accusative NPs in (34a).

Japanese has the so-called Double Accusative Constraint (Harada 1973a, 1974; Hiraiwa 2002; Kuroda 1965 among many others) that excludes a clause containing two accusative NPs. In (34a) the parser has to insert a clause boundary between two accusative NPs in order to analyze the string grammatically in satisfying the Double Accusative Constraint. Once a clause boundary is inserted, the embedded clause structure is created and it becomes clear that the subject position in the embedded clause is empty. Thus, the parser can predict the upcoming relative clause structure, resulting in the facilitation of the relative head position.

According to this account, the accusative NPs before the embedded verb must be processed in advance of the embedded verb. Thus this finding is not compatible with head-driven approaches, and supports the incremental structure building.

Let us see another piece of evidence for incrementality that is directly relevant to our study. As we have seen several times already, when a sentence contains a fronted dative marked wh-phrase, it is associated with the most deeply embedded clause. To probe for this LDB, we used Aoshima et al.'s Japanese Filled Gap

paradigm. In Aoshima et al.'s original experiments and also in our own experiments, the Filled Gap effect was observed at the embedded dative position before the verb is encountered. This preverbal Filled Gap effect cannot be explained under the head-driven approaches. If the sentence structure is not built before the verb is encountered, we do not expect the Filled Gap effect to take place before the verb position. However, we observed a slowdown at the second occurrence dative position which is before the verb. Thus Filled Gap effect also supports the incremental model of sentence processing, and reject the head-driven approaches.

We can understand the incrementality as a general property of the human sentence processing. Now the question is what kind of mechanism can derive the incrementality. From now our discussion turns to psychologically plausible sentence processing strategies.

3.2. Mechanisms for Sentence Processing

In the following subsections, we will review some of the previously proposed sentence processing mechanisms informally. Our focus here is which type of algorithm can appropriately capture the incrementality of human sentence processing without problems. To investigate this point, we will discuss the advantages and problems of three strategies: Purely bottom-up strategies; purely top-down strategies; and mixed strategies.

3.2.1. Purely Bottom-Up Strategy and Their Problems

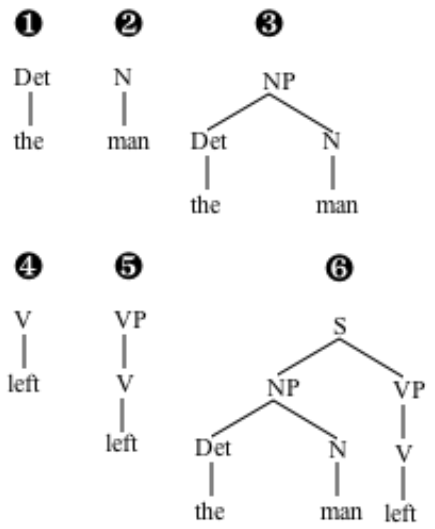
Bottom-up parsers are driven by the words in the sentence, using the rules of the grammar to combine the words into constituent structure in a "bottom-up" fashion. Let us see how a purely bottom-up parser processes a sentence "the man left". First, let us assume simple phrase structure grammar in (35).

(35)

S	-- >	NP VP	Det	-- >	the, a, every …
NP	-- >	N	N	-- >	John, Mary, man, woman, book …
NP	-- >	Det N	P	-- >	on, with, about …
NP	-- >	NP PP	V	-- >	run, say, put, open, read, give
PP	-- >	P NP			
VP	-- >	V			
VP	-- >	V NP			
VP	-- >	VNP PP			

(36) illustrates how the sentence "the man left" is processed.

(36)



At step 1, a determiner is found, so the Det node can be projected. At the step 2, a noun "man" is found, the rule of the grammar allows the projection of N node. If there is a rule in which these categories appear on the right-hand-side, the algorithm can combine them to the category in the left-hand-side of the same rule. At this point we have two categories, det and N that are compatible with the right-hand-side of the second NP rule. Thus at step 3, the algorithm combines these categories into one NP. That is, they are combined using the rule NP--> Det N rule. The same process continues until all the words are processed. At the end of the process, the algorithm projects the S node, and we are left with a complete sentence.[33]

It is important to note that step 2 and step 3 in (36) can be applied at any point. Within each step, there can always be alternatives if a word has an ambiguity.

[33] This is quite a simple and informal illustration, but for our purposes this informal illustration is sufficient. Basically this illustration, and the following illustrations of the top-down and left-corner algorithm are adopted from Crocker (1999). For the details, readers are recommended to see Crocker (1999) and reference therein.

Thus if an ambiguous word like *bank* which can be either N or V is in the input, projecting the N or projecting V are the alternatives for this word. Also there can be alternatives if more than one phrase structure rule has a right-hand-side that matches the categories.

The advantage of this algorithm is that the structure is built using the bottom up information of each lexical item in the input, and there is no necessity of combining the category, the algorithm does not make severe errors. The categories can be combined when the item is found in the input that can specify which rule should be applied to project the category.

Although it has an advantage, there are two problems with this algorithm. First, within this algorithm the sentence is not processed incrementally. This problem becomes obvious when we look at a sentence with a right-branching structure like (37).

(37) $[_{IP}$ The dog $[_{VP}$ saw $[_{NP}$ the cat $[_{CP}$ that $[_{IP}[_{VP}$ chased $[_{NP}$ the mouse]]]]]].

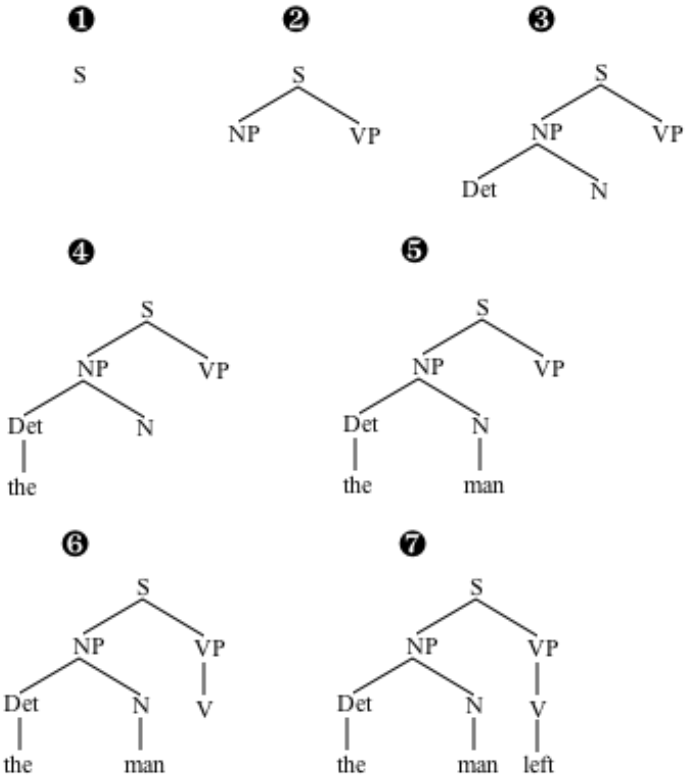
Within the bottom-up algorithm, the CP node cannot be attached to the NP node until the word "mouse" at the end of the sentence is processed. This is so because $[_{CP}$ that …] cannot be projected before the $[_{IP}[_{VP}$ chased $[_{NP}$ the mouse]]] is projected, the $[_{IP}[_{VP}$ chased …]] cannot be projected before the $[_{VP}$ chased $[_{NP}$ the mouse]] is assembled, and the $[_{VP}$ chased …] cannot be projected before the $[_{NP}$ the mouse] is assembled. Thus in the right-branching structure, the algorithm has to delay the process until the very end of the sentence.

The second problem is related to the first problem. Because of the first problem illustrated above, the algorithm has to store the completed categories uncombined until the end of the sentence in the right-branching structure. Thus, this algorithm predicts the difficulty with the purely right-branching structure, which normally does not create difficulty (Babyonyshev and Gibson 1995; Chomsky and Miller 1963 Gibson 1991; Yngve 1960 among others). For this reason, the purely bottom-up algorithm is not a reasonable algorithm for human sentence processing.

3.2.2. Purely Top-Down Strategies and Their Problems

Another possible strategy is a purely top-down strategy. In contrast with the bottom-up algorithm, the top-down algorithm starts processing from postulating the top node, S. Thus, the process proceeds from top to bottom. (38) illustrates how the sentence “the man left” is processed using the grammar in (35).

(38)



At step 1, the algorithm initiates the process by postulating the S node. The algorithm can find the rule in which the left-hand side corresponds to the postulated category. So at step 2, the algorithm applies the S -- > NP VP rule, and expands the tree. In this way, the algorithm repeats applying the rule for the first daughter of each node until the bottom of the tree is reached. When the bottom of the tree is reached, the algorithm looks back up the tree, and finds another node that can be expanded in the same way. The parser repeats this process until all the nodes are expanded and reach the bottom of the tree.

An obvious advantage of this algorithm is its incrementality. Because the algorithm proceeds strictly from top-to-bottom and left-to-right, the sentence is

processed incrementally. When the parser reaches the terminal "the" in the subject of the sentence in (38), all the necessary nodes are processed and integrated into the structure.

Although incrementality is preserved, there are some problems with this strategy. First this strategy predicts difficulty with the left-branching structure like (39). Because the process proceeds from the first daughter of S to right, and the algorithm does not look back up the tree before it reaches the terminal of the first daughter, the parser has to store many incomplete NP node before it reaches the terminal *My* in (39). Thus, the more left-recursion occurs in the sentence, the more NP nodes are there for the parser to keep track with, and the left-recursion structure is predicted to be difficult to process contrary to fact (Abney and Johnson 1991; Babyonyshev and Gibson 1995; Gibson 1991 among many others).

(39) $[_{NP}[_{NP}[_{NP}[_{NP}$ My cousin's] aunt's] dog's] tail] fell off.

On the other hand, this strategy does not have problems processing the right-branching structure as in (40). In the right-branching structure, the terminal of each node "comes soon" without many intervening incomplete categories.

(40) $[_{IP}[_{NP}$ the dog]$[_{VP}$ saw $[_{NP}$ the cat $[_{CP}$ that $[_{IP}[_{VP}$ chased $[_{NP}$ the mouse]]]]]]].

The algorithm incorporating the purely top-down strategy, therefore, predicts that the left-branching structure is more difficult than the right-branching structure if the

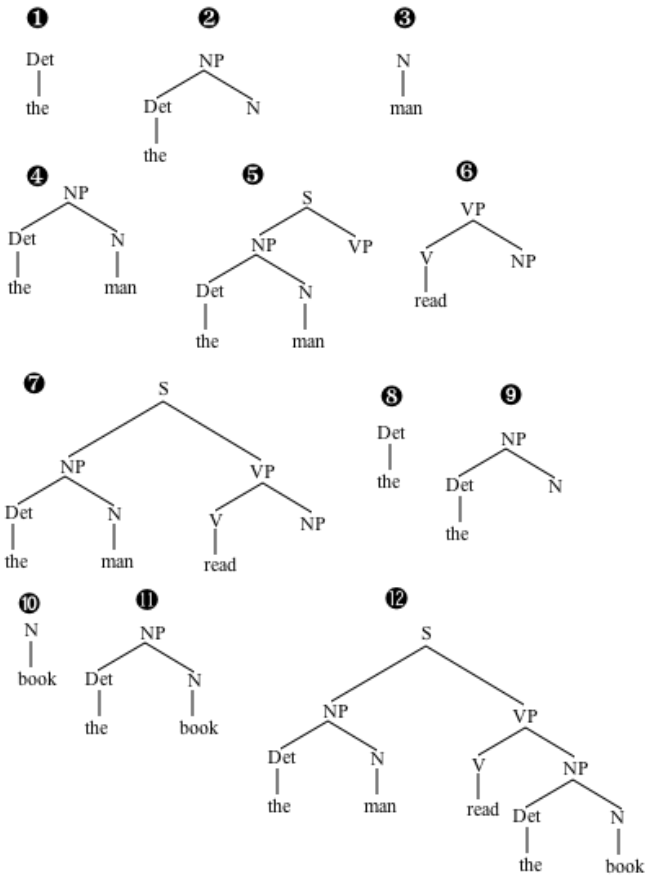
number of embedded is the same. Furthermore, the algorithm predicts that predominantly left-branching languages like Japanese are difficult to process. However, there is empirical evidence that left-embedding structures are not difficult to process for humans, and they are certainly easier than, for example, center-embedding constructions (Abney and Johnson 1991; Babyonyshev and Gibson 1995; Gibson 1991, 1998; Mazuka et al. 1989; Miller and Chomsky 1963 among many others). Thus, a purely top-down strategy is not a reasonable strategy for human parser.

3.2.3. The Left-Corner Strategy

Because of the problems with purely bottom-up and purely top-down parsing algorithm, it has been proposed that human parser operates in partially top-down and partially bottom-up fashion. One such mixed algorithm is a left-corner algorithm or left-corner strategy (Abney and Johnson 1991; Aho and Ullman 1972; Babyonyshev and Gibson 1995; Gibson 1991; Johnson-Laird 1983; Kimball 1973, 1975; Resnik 1992; Stabler 1994).

The left-corner algorithm processes a sentence in a way in which it processes the leftmost daughter of the right-hand side of a rule of the grammar in a bottom-up fashion, and applies the rest of the rule from the top-down fashion. Let us see how a left-corner parser processes the sentence “the man read the book.”

(41) The man read the book.



In (41), when the algorithm finds the word *the*, it allows a *Det* node to be built. *Det* is the left-most category at the right-hand side of the second NP rule in (35). The second NP rule tells that it can be expanded as *Det* and *N* (NP -- > Det N). Thus the sibling of the NP, the N node is projected. The word *man* comes into the input, and it allows the N node to be built, which can be attached as the right-hand member of the NP, and NP is completed by attaching the newly projected N into the predicted N node. When the NP is completed, the *S* rule in (35) tells that NP is the left-most category of the right-hand side of *S*. Thus, the NP allows the S node and its sibling, VP-node to be

built. The word *read* comes into the input, which can project up to V and VP (the step 6). VP is the right most member of the currently constructed S node. So the newly projected VP is attached to the predicted VP (the step 7). The word *the* comes into the input and as we have seen it can project an NP node and an N node, the sibling of the NP can be predicted (the step 8 and 9). The word *book* comes into the input, projecting the N node. The newly projected N node is incorporated into the predicted N node (step 10 and 11). The newly projected NP is the right-most category of the current VP node, thus the newly projected NP can be incorporated into the predicted NP node (step 12) and the sentence structure is completed.

An advantage of this strategy is that it allows for more incremental parsing than purely bottom-up parsing because it allows some constituents to be integrated into the structure before the constituent has been completed.

Furthermore, because initial rule application is driven by bottom-up information, a left-branching structure does not create a problem.

Additionally, a left-corner parser can resolve a problem with a bottom-up algorithm such as the categories are stored unattached until the end of the sentence in the right-branching structure. Because a left-corner algorithm allows a certain prediction, and categories are integrated into the structure as soon as it becomes clear that the predicted node and the newly projected node are the same categories, the parser does not need to store the categories unattached until the end of the sentence even in the right-branching structures.[34]

[34] This part is treated differently depending on the type of left-corner algorithm. For example, the so-called Arc Eager algorithm allows for the parser to integrate even incomplete categories into the

Because it can resolve the problems with purely bottom-up or purely top-down strategies, left-corner strategies are understood as one of the most psychologically plausible parsing algorithms. However, this algorithm is not totally free from problems. Schneider (1999) points out one of the problems with the standard left-corner algorithm. The problem is concerned with the processing of head-final languages. As we have seen earlier, a head-final language like Japanese shows massive temporary ambiguity because of its head-finality and the lack in the left-edge markers of constituents (Hirose 1999). Because of the left-edge ambiguity, in languages like Japanese, the same string of words can be continued in a various ways as illustrated in (42).

(42) Sideny-ga Rosa-ni ringo-o…
S-nom R-dat apple-acc

a. … ageta. (*Simple Clause*)
… gave
"Sideny gave the apple to Rosa."

b. … ageta-to itta. (*Complement Clause*)
… gave-comp said.
"Sideny said that he gave the apple to Rosa."

---

structure. On the other hand, the so-called Arc Standard algorithm does not allow the incomplete categories to be integrated into the structure and the same type of problem arises as the bottom-up algorithm. In this dissertation, we will adopt the arc eager algorithm, and not talk about the other type of algorithms. See Abney and Johnson (1991), Stabler (1994) and Schneider (1999) for detailed related discussion.

c. … ageta-node yorokonda. (*Adjunct Clause*)

… gave-because be-pleased.

"Sidney is pleased because he/someone gave the apple to Rosa."

d. … ageta hito-o shookai-sita. (*Relative Clause*)

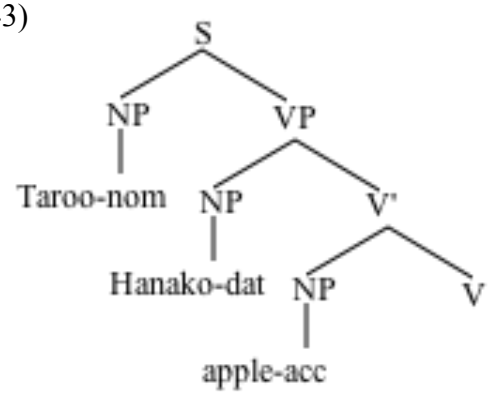
… gave person-acc introduced.

"Sideny introduced to Rosa the person who he gave the apple to."

Recall that even in a head-final language like Japanese, the string in (42) is interpreted before a licensing head like a verb is encountered. This, in turn, suggests that the parser commits itself to a particular analysis before the verb is encountered. However, if it is true, the left-corner parser predicts that (42b, c, and d) create difficulties because any of these continuation requires some reanalysis.

When the string in (42) is encountered, the left-corner parser projects the structure in (43). If there is no other bottom-up information, the algorithm does not project the other structure. However, if the parser commits to this particular structure, when the verb with the complementizer, or the relative head is encountered, the structure must be reanalyzed from a simple clause structure in (43) to other structures such as a complement clause structure, an adjunct clause structure or a relative clause structure. If such reanalysis creates difficulty (Pritchett 1992b; Sturt 1997), the continuations in (42b, c, and d) should be difficult to read. However, none of these continuations are particularly difficult.

(43)



Given that all the continuations in (42) are read easily, how can we derive this fact without sacrificing the advantages of Left-Corner Algorithm?

3.2.4. SPARSE (Schneider 1999): Incrementality and Flexibility

Facing the problem of the left-edge ambiguity discussed above, Schneider (1999) proposes a variant of left-corner parsers that has sufficient flexibility in predicting the upcoming structure. This parsing algorithm is called SPARSE. As we will see shortly, SPARSE resolves all the major problems of top-down and bottom-up parsing in the same way as the left-corner algorithm, and it further resolves the problem of left-edge ambiguity without sacrificing incrementality. Thus we will adopt SPARSE as the syntactic structure building algorithm.

One of the key features of Sparse is that it is not forced to predict a specific syntactic category. Rather, SPARSE make a prediction of the upcoming structure on the basis of features. This is made possible by adopting Bare-Phrase Structure (Chomsky 1995). In other words, in SPARSE, the minimal building block accessible to the parser is the syntactic feature rather than syntactic category. SPARSE predicts

the upcoming structure in a particular manner. The parser is allowed to predict structure only if a fully connected tree cannot be built otherwise. Thus the predicted features are required to license the currently processed word into the structure.

Let us see how a sentence like (44) is processed with SPARSE.[35]

(44) Dorothy will see munchkins.

When the parser encounters the word Dorothy, the parser does not build any structure. However, the features conveyed by the word are activated.

(45) *Dorothy*: [Cat: Noun], [Case: Nom, Acc], [Num: Sg], [Person: 3]

This feature bundle can be read as "the word Dorothy conveys the categorical feature 'Noun', the Case feature either Nominative or Accusative, Number feature 'singular' and the person feature 'third person'". As we can see clearly, each lexical item is understood as the bundle of features in SPARSE following the formalism of Bare Phrase Structure.

The next word, *will*, is encountered. The lexical entry of *will* contains at least the ones in (46).

[35] The discussion here basically follows Schneider (1999)'s Chapter 2.

(46) Will

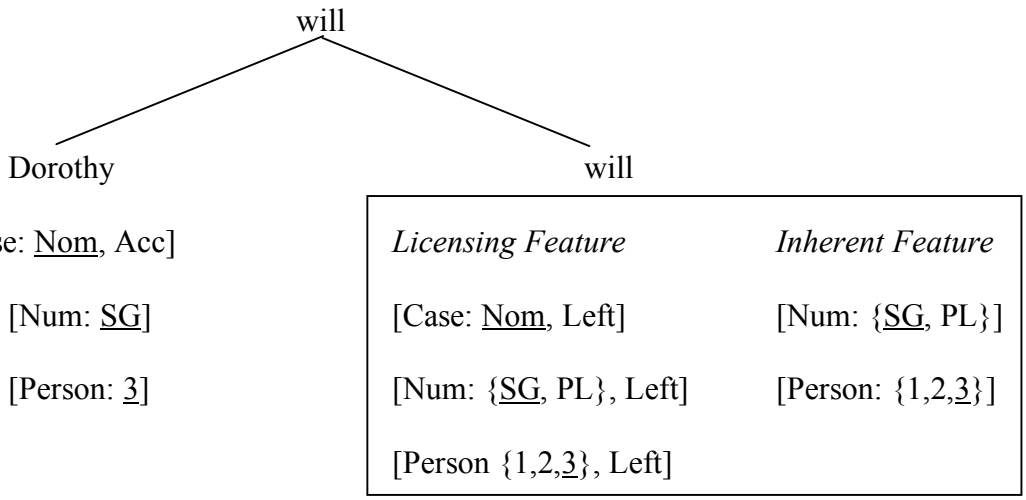
Inherent Features	Licensing Features
[Cat: V]	[Case: Nom, Left] [Num: {Sg, Pl}, Left] [Person: {1,2,3}, Left] [Cat: V, Right] [VForm: Infin, Right]

The feature structure in (46) illustrates another important feature of SPARSE. SPARSE distinguishes two types of features that can be stored in the lexical entry for each word: the inherent features and the licensing features. The inherent features include a type of the feature such as Case or Category, and the value of the feature, e.g., Nominative, Accusative or Dative for Case, and so on. Inherent features specify the grammatical category of a word, its morphological features and so on. Licensing features, on the other hand, define the relations such as Case assignment, selectional relations, or theta role assignment. Because directionality of these relations is different from language to language (according to, for example, the head directionality parameter (Travis 1989)), the licensing features include such information of directionality also in a way as [Case: Nom, Left].

When the parser encounters a new word, its licensing features are looked up in order to determine whether they can license the previously encountered word, *Dorothy*. If they can, *Dorothy* is attached to *will*. In this case, *will* can license the

[Cat: Nom] feature of *Dorothy* by the licensing feature [Cat: Nom, Left]. Now the structure in (47) is built.

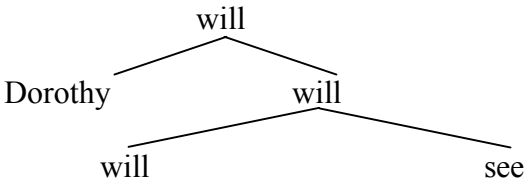
(47)



In SPARSE, such relations as complement or specifier are expressed by the non-null intersection of the values of the features that are shared by two combined elements. For example, in (47), the value 'singular' or the value 'third person' is chosen from the set of number features or person features for *will* as a result of an unambiguous Number or Person feature on the noun *Dorothy*. (47) shows features that are evaluate at this point. The underscored values in the feature sets in (47) are checked and specified by intersecting the set of the features of the combined heads.

The parser encounters the next word, *see*. When *see* is encountered, the parser examines its features to determine whether it can be attached to the previous word, *will*. The inherent features of *see* are [Cat: verb] and [VForm: Infin]. It can be licensed by *will*, and combined as a complement of it. This step is illustrated in (48).

(48)



[Category: verb, Right] [Cat: verb]

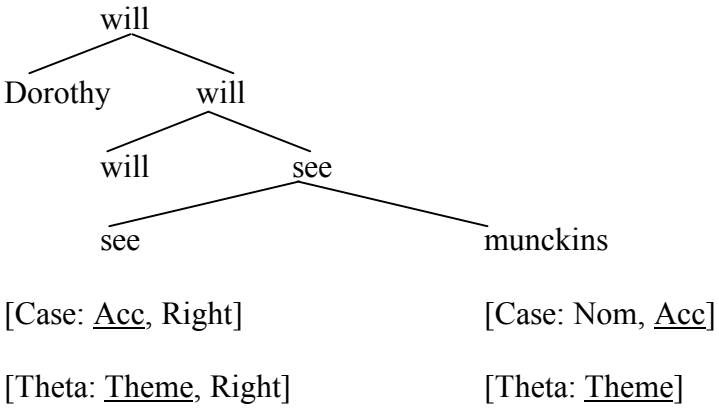
[VForm: Infin, Right] [VForm: Infin]

Finally the *munchkins* comes into the input. Other than the number feature, it has the same features *Dorothy* in (45).

(49) *munchkins*, [Case: Nom, Acc], [Theta: Theme]

The inherent feature of *munchkins*, [Case: Nom, Acc] is intersected with the licensing feature of the previous verb see, [Case: Acc, Right]. So *munchkins* is integrated into the structure as the complement of *see*. The resulted structure is in (50)

(50)



So far we have seen how the structure of simple sentence can be parsed by SPARSE. Now let us look at how SPARSE handles left-edge ambiguity. The important point in the following discussion is when and how the structure is predicted. Schneider proposes that heads are predicted only when they are needed to allow for an incremental parse. In other words, predicted heads are posited only when the structure cannot be built incrementally if they are not posited. First let us see how the predicted heads are posited using relatively a simple example. Schneider cites (51) from German.

(51) NP-nom NP-acc Vtrans

… dass er den Hund sah

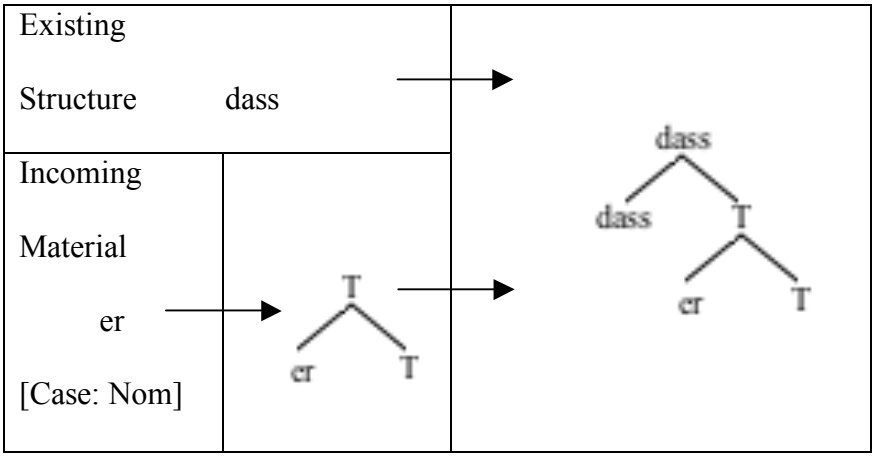
… that he-nom the dog-acc saw

"… that he saw the dog" (e.g., I know that he saw the dog.)

When the nominative pronoun *er* is encountered, it cannot be attached into the main tree, because the previously encountered complementizer, *dass* does not assign nominative case. Because *er* cannot be attached into the existing structure, the

distinguished feature for *er*, nominative case, is used as the basis for a licensing head. Now the lexicon is searched for heads that can assign nominative case, and T head is the only head that assigns nominative case. Therefore, a T head is posted as the licenser for the nominative pronoun, and the entire pronoun-tense complex is attached as the complement of the complementizer *dass*.

(52)

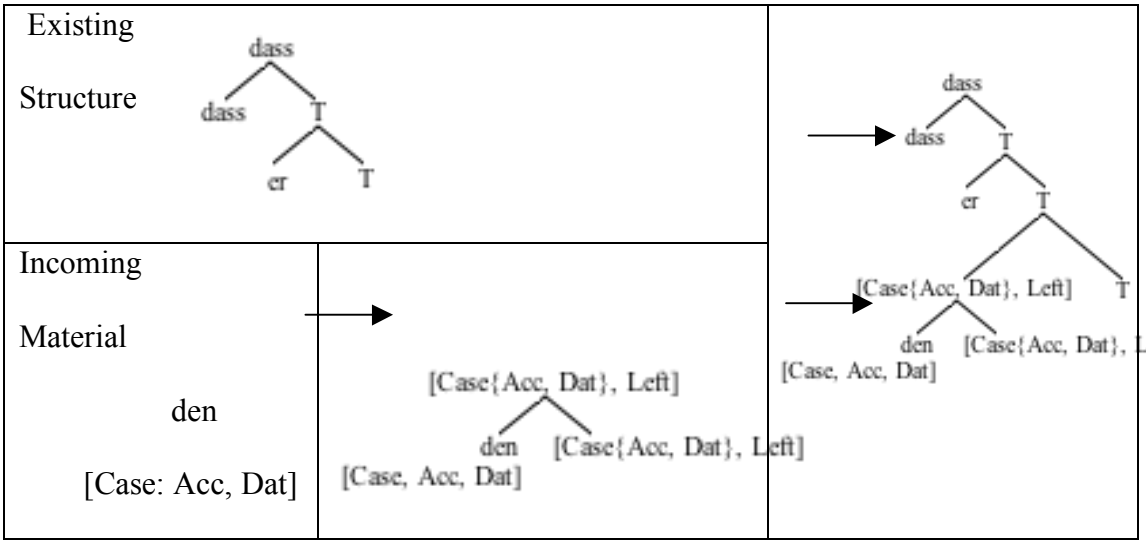


In this case, only one type of head can license the nominative pronoun. Thus it can be guaranteed that a T head will license *er* in the tree.

When the determiner *den* is encountered, again this word cannot be directly attached to the existing tree. Therefore the search for the licensing head is initiated. In this case however, there are multiple possible heads that can license the determiner because, the inherent features are the [Case, Acc, Dat]. The head that can license this determiner is either a verb or a postposition. The only shared properties of these heads are that they assign case. Thus SPARSE posits an underspecified head that

contains only the licensing feature [Case: {Acc, Dat}, Left]. and the determiner-predicted head complex is incorporated into the existing structure.

(53)



In this case, the predicted head is underspecified head with only the licensing features, therefore, either the verb or postposition is compatible with the predicted head and the desired flexibility is obtained. I will skip the remaining parts here.

Schneider also cite an example from Japanese and shows how a head-final relative clause structure in (54) could be parsed.

(54) Hirosi-ga [$_{NP}$[$_{CP}$[$_{IP}$ ø Masao-o mita]] otoko]-o …

H-nom M-acc saw man-acc

This type of structure can be parsed without conscious difficulty (see Mazuka and Itoh 1995 for related discussion and observation).

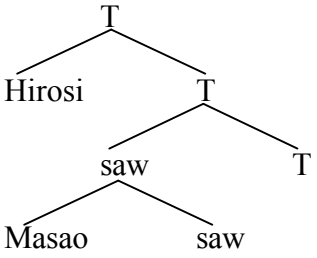
At the point immediately before the NP *otoko-o* "man-acc" is encountered as in (55), it has the structure of a simple SOV clause as in (56).

(55) `Hirosi-ga Masao-o mita`

`H-nom M-acc saw`

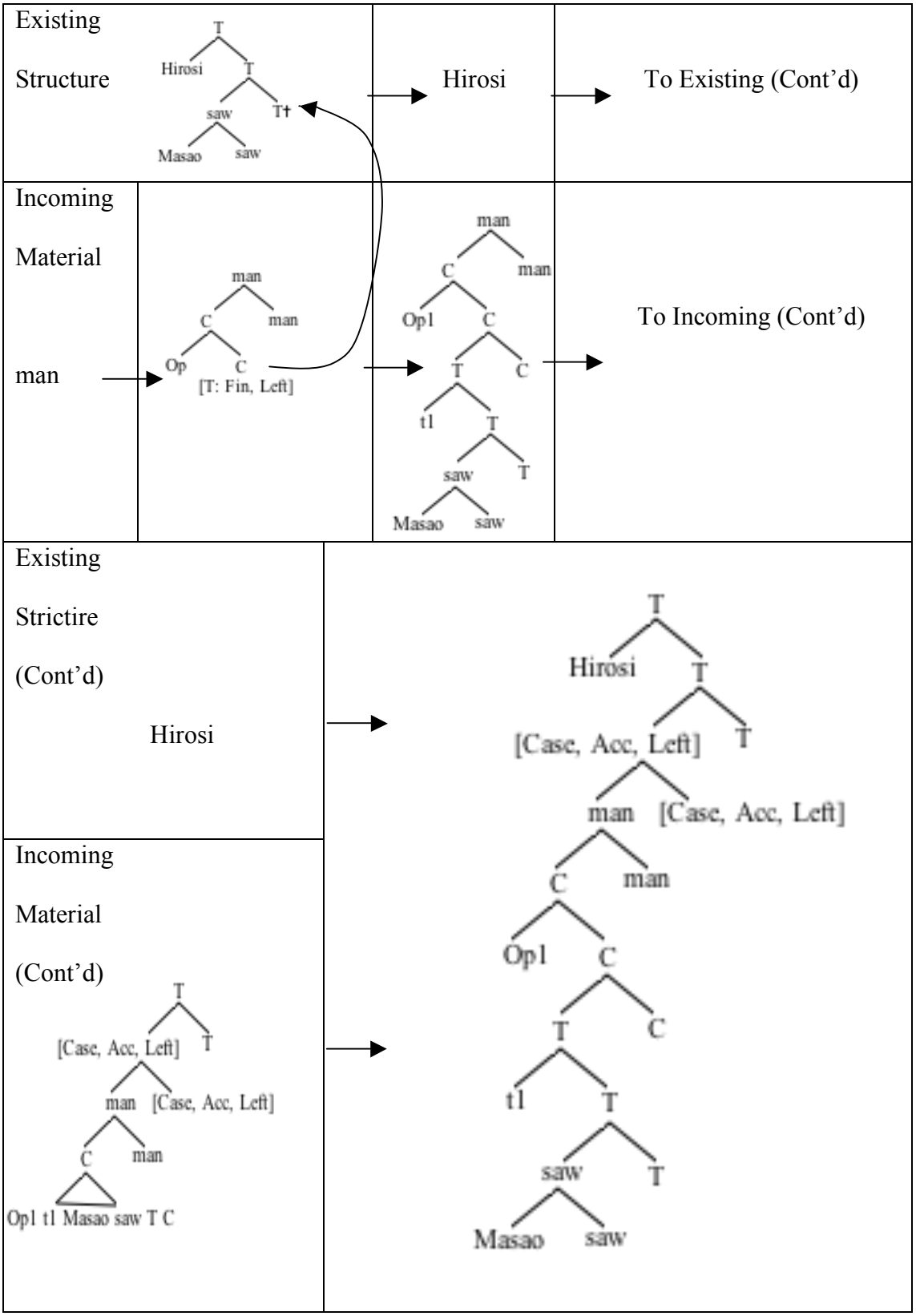
"Hirosi saw Masao."

(56)



When the NP *otoko-o* is encountered, it can only be interpreted as the head of a RC. To build an RC, the entire VP-T constituent must be reanalyzed from a position inside the matrix clause to a position inside an RC. The subject NP *Hirosi* is left unaffected. When the NP *otoko-o* is encountered, it cannot be attached directly to the existing tree. The search for predicted heads is initiated. A null C head for an RC along with the associated operator is found, that adjoins to the left of an NP. Schneider speculate that the C head of an RC selects for an NP, so it will be returned by the search for elements that select for NPs. Once this predicted structure is built, the VP-T constituent from the existing tree can be stripped off from the simple clause structure and can be taken as the complement of the CP. As a consequence of positing an operator, a trace is inserted as the subject of *saw*.

(57)


Existing
Structure
T
Hirosi
T
saw
T†
Masao
saw
Hirosi
To Existing (Cont'd)
Incoming
Material
man
man
C
man
Op
C
[T: Fin, Left]
man
C
man
Op1
C
T
C
t1
T
saw
T
Masao
saw
To Incoming (Cont'd)
Existing
Strictire
(Cont'd)
Hirosi
Incoming
Material
(Cont'd)
T
[Case, Acc, Left]
T
man
[Case, Acc, Left]
C
man
Op1 t1 Masao saw T C
T
Hirosi
T
[Case, Acc, Left]
T
man
[Case, Acc, Left]
C
man
Op1
C
T
C
t1
T
saw
T
Masao
saw

Let us summarize the discussion so far. The most important aspect of SPARSE is that it preserves incrementality as well as the flexibility of the prediction. Because of its flexibility, SPARSE can successfully deal with the left-edge ambiguity without problems. In this sense, SPARSE is one of the most reasonable parsing algorithms for head-final languages. Therefore, in this work, we adopt SPARSE as the basic structure building algorithm.

4. How to Block the Longer Dependency?

From now, we will investigate how SPARSE can derive a prediction for an upcoming RC structure or conditional clause structure. In the course of the discussion it will be made clear that we have to modify SPARSE to capture the predictions of these structures.

In this section we will concentrate on showing how the non-interrogative complementizers can be predicted. In the course of the discussion, we will see a specific way to enhance the predictive component of the parsing algorithm. For the sake of simplicity, I will keep most of the discussion here informal.

4.1. Dealing with the Indirect Cue

First let us discuss the case of RC prediction. As we have discussed, to capture the fact that the longer dependency bias is blocked in a predicted RC, it is sufficient to predict the Adnominal form of the embedded verb because it is not compatible with the motivation for the LDB, the wh-feature. In the discussion on the syntax of

RC, it has made clear that the Adnominal morphology is formed by V-$T_{(finite)}$-$C_{(empty)}$. Thus, the parser has to predict this verbal complex by means of the classifier mismatch. With this discussion in mind, let us see whether SPARSE can predict the Adnominal morphology.

First let us examine an example of simple NP that does not contain a classifier mismatch as in (58).

(58) san-nin-no gakusee-ga …

three-$cl_{(person)}$-gen student-nom

"Three students …"

First the genitive marked classifier is encountered and its feature structure is activated.

(59) san-nin-no[36]

[Cat: Class], [Case: Gen], [Sem: Person]

Then the second word comes into the input. It has the following feature structure.

[36] In order to keep the discussion as simple as possible, we will not be concerned with the matter of how each head are attached to the case-particle, or how each classifier is attached to numeral or genitive case particle.

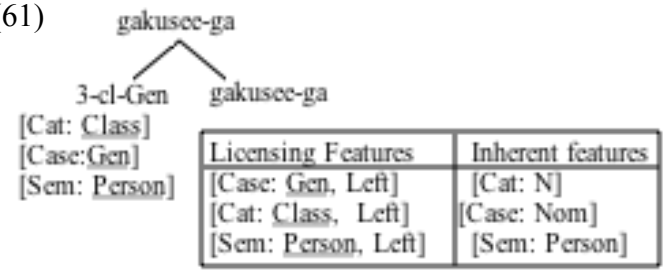
(60) gakusee-ga:

Inherent Features	Licensing Features
[Cat: N] [Case: Nom] [Sem: Person]	[Case: Gen, Left] [Cat: Class, Left] [Sem: Person, Left]

Some notes are in order. We have to specify the licensing feature so that we can capture the relation between the classifier and its host noun. Basically classifiers are licensed by a certain head noun, and they must have specific semantic relations to the head noun. We can capture such relations by postulating a licensing feature [class: person, left] for example. The feature for classifiers must specify the semantic features, and this feature and noun's inherent semantic feature must be the same. N heads also license genitive case. Thus in the same fashion, we can postulate a licensing feature for the genitive case like [Case: Gen, Left].

Now the parser inspects the licensing feature of *gakusee-ga* to determine whether it can be combined with the previously encountered classifier. The classifier is attached to *gakusee-ga* as a result of licensing the [Case: Gen] feature and [Sem: Person].

(61)



Let us turn to an example of an RC containing classifier mismatch in (62).

(62) san-satu-no [$_{NP}$[$_{CP}$[$_{IP}$gakusee-ga [$_{VP}$gap yon]-da]ø] hon]
three-cl$_{(book)}$-gen student-nom read-past-ø book
"Three books that the student read."

What we have seen is that when the semantic incompatibility between the classifier *san-satu-no* 'three-classifier-gen' and its adjacent nominative NP, *gakusee-ga* 'student-nom' is encountered, the parser can predict the morphology of the embedded verb.

When the parser encounters the genitive-marked numeral classifier, its feature structure is activated. Because classifiers have narrowly specified semantic information, we can add semantic features to its inherent features.

(63) san-satu-no
[Cat: Class], [Case: Gen], [Sem: Book (or book like object)]

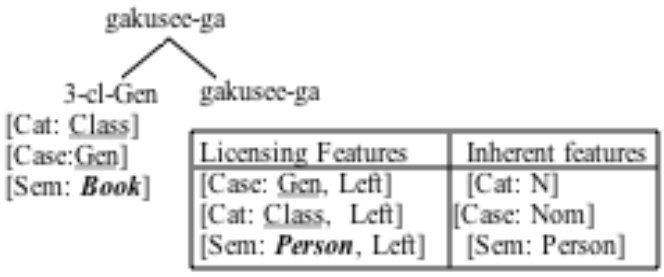
The next NP *gakusee-ga* is then encountered. It has the features as we have seen above, repeated here as (64).

(64) gakusee-ga:

Inherent Features	Licensing Features
[Cat: N] [Case: Nom] [Sem: Person]	[Case: Gen, Left] [Cat: Class, Left] [Sem: Person, Left]

Now the parser inspects the licensing features of *gakusee-ga* in order to determine whether it can license the previously encountered classifier. Because, the [Case: Gen] and [Cat: Class] on the classifier can be checked, the parser attempts to combine these two. However, when they are combined, it becomes clear that [Sem: book] cannot be checked by [Sem: Person, Left]. This semantic mismatch initiates the parser's attempt to reanalyze the structure. A necessary assumption here is the parser tries to combine these two elements because there can be non-null intersection of the values in the two sets of features. Under this view, the classifier mismatch effect is a reflection of the parser's attempt to combine these two elements.

(65)

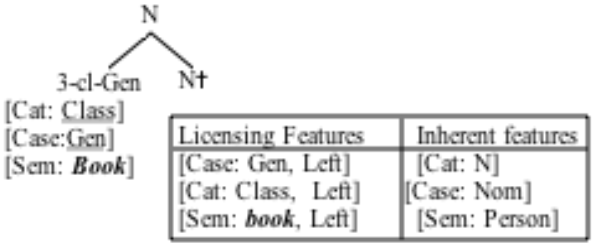


Now the reanalysis takes place, and initially combined elements are separated. Because of this reanalysis, these two elements cannot be attached into the structure, and the search for predicted heads is initiated at this point.

First, the lexicon is searched for heads that can license the classifier. This search returns only one type of head. Because the classifier bears genitive case the head that can license genitive case is searched for.[37] It is a Noun. Furthermore, the classifier is semantically specified, in this case it is specified for *books* or 'book like objects'. For the sake of exposition, here let us assume that the classifier is unambiguously specified for *books*. Thus, the noun with the semantic value of *book* is the one that can license this classifier. The classifier, then projects the structure in (66).

[37] Note that this prediction involves the prediction of adjunction structure because it is the structural licensing relation between the classifier and the host NP as we have seen in Chapter 3.

(66)

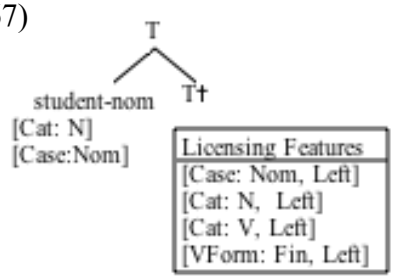


Next, the nominative-marked NP cannot be incorporated into the structure either. Thus the predicted head is searched in the lexicon. Assuming that nominative Case in Japanese is licensed uniquely by finite T head (Takezawa 1987), this search also returns a single head, the finite T. Thus, the following predicted structure can be built.[38]

[38] Both the prediction of the finite T for the nominative NP and N head for genitive classifier are made based on overtly marked case information. In Japanese, if an NP is not explicitly case marked, it is not possible to tell which Case it bears, without considering such information as verb's argument structure. Thus it is plausible to assume that in a situation where NPs do not bear Case markers, these specific predictions cannot be made. Rather the underspecified heads should be predicted in such a case.

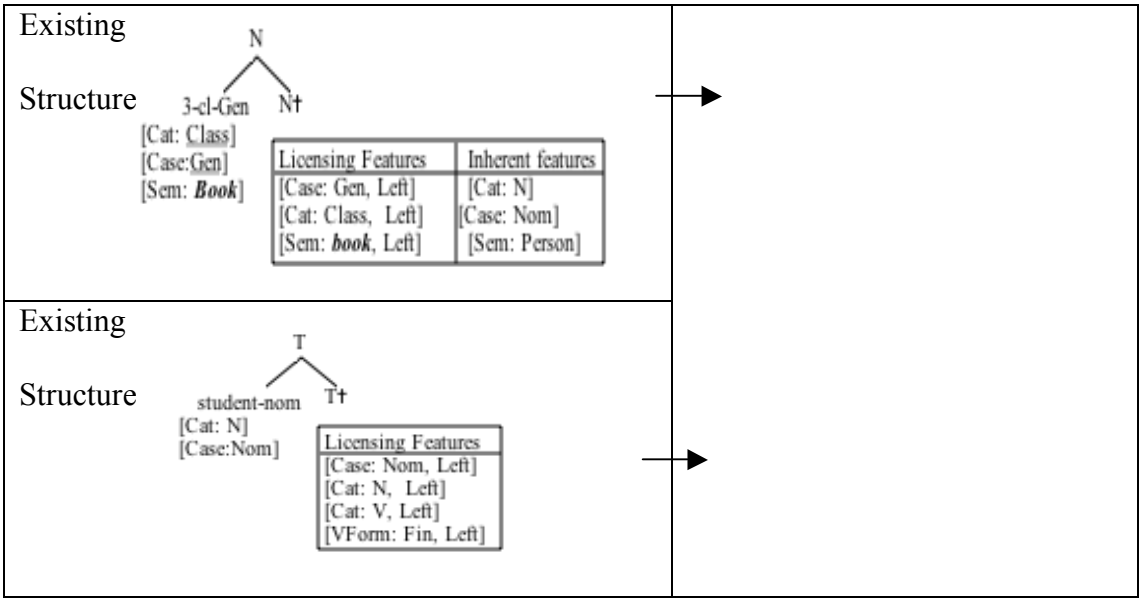
A note, however, is in order. Although these assumptions are indeed plausible, it is an interesting question to ask what is predicted if NPs do not bear Case. In order to investigate this question, Japanese is not a good language. In Japanese Case dropping seems to be well-constrained, although its nature (see Kageyama 1993; Yoshida and Yoshida 1993 among many others). Well-know generalizations are the following. The nominative Case on the subject NP and Genitive Case are normally not omitted. Given these properties of Japanese Case particles, it is difficult to test the situation where Case particles are dropped. In this regard, a language that allows Case dropping or that does not employ explicit case morphology (such as Chinese) is interesting language to test. The same question is addressed in Hsu (2006).

(67)



Now, we have two predicted structures. The parser then tries to incorporate these two predicted structures. However, N head does not license the category T. Thus these two structures cannot be combined, and the process is doomed to breakdown.

(68)



For these two structures to be combined grammatically, "the structure of an RC" must be predicted, and it means that the CP on the top of the predicted IP must be predicted, and it must be attached to the predicted NP as an adjunct.

4.2. Enhancing the Predictive Component

It shall be clear now that the predictive power of SPARSE is too limited to accommodate the prediction of the RC structure. The question here, then, is whether we can modify the algorithm so that we can accommodate the RC prediction or the algorithm should be discarded altogether and we should seek a new algorithm.

The key points are how can the CP structure be predicted in addition to the predicted IP, and how it can be incorporated into the structure as an adjunct of the predicted NP. In addition to these key points, we have to sufficiently restrict the

power of the predictive mechanism so that the predictive mechanism does not become too powerful.

Let us see how we can modify SPARSE to allow for a prediction of CP on top of the predicted IP. The modification that we propose is to allow the parser to access the information of the predicted head. As we have seen, Nominative NPs in Japanese are licensed by the finite T. Thus, it is plausible to assume that if an NP bearing the Nominative Case is processed, SPARSE can predict the finite T rather than an underspecified licensing head. If a finite T head is predicted, its intrinsic features should also be predicted. Assuming that a finite T must be selected and thus be dominated by a C or CP (Chomsky 1986b), we can incorporate this information as finite T head's intrinsic feature, i.e., if a finite T is predicted, it is also predicted that there must be a C head that selected the finite T. This, in turn, implies that if the predicted T is not a finite T, it does not have such an intrinsic feature. So for example, a non-finite T can be selected by C (in control construction), V (in ECM construction) or P (in ECM construction) (Chomsky 1986a). Thus, to preserve flexible structure building and avoid potential errors, the parser should predict an underspecified head, if non-finite T is specifically predicted. Consequently, the environment in which C is predicted is quite limited. A question arises with respect to the recursive prediction. We have seen that the predicted finite T allows for a further prediction of C. Now, the question is whether the predicted C head allows for a further prediction of the structure or not. Considering the selectional requirements of C, the answer is no. A C or CP is selected by a wide variety of categories. For example, C is selected by a V such as *think*, by an N such as *claim* and so on. Furthermore, it can be a root category

also. Thus, as in the case of non-finite T, a prediction set by the predicted C should be an underspecified head or given the possibility of root-C, it should not predict anything. Thus, we can plausibly assume that the predicted C head can predict an underspecified head that licenses C. Given that the predicted C can only predict the underspecified head, the recursive prediction should be stopped at this point. It is not plausible to assume that the predicted underspecified head can predict some category. Thus, when the underspecified head is predicted, the recursive prediction is terminated. In this way, we can allow the parser to predict the C head, and we can still sufficiently restrict the predictive power of the parser.

The prediction of a finite T allows also for a prediction of the V head. Either finite T or infinite T requires V. We can assume that the licensing feature for V is specified in T's lexical information. Thus if T is predicted, V is also predicted.

Now let us turn to the second part, i.e., the question of how the parser attaches the predicted CP as an adjunct of the predicted N. There are at least two plausible ways to consider. The first possibility is to allow the parser to just concatenate the predicted C and the predicted N. The other possibility is to make use of the information of the predicted head.

Simple Merger Approach

Let us discuss the first possibility. Because SPARSE incorporates Bare Phrase Structure (Chomsky 1995) as a theory of phrase structure, it can simply merge any two existing elements. Thus, the predicted C head and the predicted N head can be simply merged.

If we adopt this approach, what is the course of structure building that the parser should follow? At the point when these two heads are merged, however, the projection or the label of the resulted complex cannot be determined. In other words, this parser does not specify whether the C is attached as an adjunct or a complement at the point of merge. The projection can be determined when the lexical N head comes into the input, and its licensing features are accessed by the parser. If the N head takes a complement clause, the C is analyzed as a complement, but if the N head does not take any complement, the C is analyzed as an adjunct. This possibility, therefore, appeals to the benefit of the flexible structure building to the utmost extent.[39]

Note however that in this approach we have to assume that there is not a preference of the parser for the complement clause analysis over the adjunct clause analysis. If there is such a preference, and if we incorporate such a preference in our mechanism in a way that attaches C to N as a complement whenever possible, we predict that the building of a RC structure always involves a reanalysis from the preferred complement clause analysis to the less preferred adjunct analysis. Therefore, if there is no complement clause preference, this approach allows for a right prediction.

There is, however, a problem with this approach. Under this approach, the structure of the RC can be successfully built. However, the morphology of the embedded verb, key information for blocking the longer dependency bias, is not

[39] Discussion with Amy Weinberg and Jeff Lidz led me to this possibility. I would like to express my gratitude to them for letting me realize this simpler approach.

predicted. From the simple merger of C and N, the morphology of the embedded verb does not follow. Thus, if we do not find a way to predict the predicate adnominal form of the embedded verb, this approach is not appealing.

Another problem of this approach is if the parser allows simple merger without taking into consideration the features of predicted heads, it allows for the merger of the predicted T and N without projecting the C head. The resulting structure is obviously problematic. Assuming that the structure of RCs involves adjunction of C to N, this approach allows for a grammatically non-legitimate structure.

Thus, to maintain simple merger approach, we have to resolve these problems, which at this point, we do not have any way.

Accessing the Lexical Contents of the Predicted Head

Another possibility is to allow the parser to make use of the information of the predicted head. Based on selectional and licensing properties of classifiers, it is possible that when the classifier is processed, a specific N head is predicted as we have seen. For example, if there is a classifier for books, *satu*, in a sentence, then there must be an NP headed by the noun *book*. Thus, the classifier creates the prediction for a specific N head. Because a specific noun like *book* is predicted, its semantic information and licensing features should be accessible by the parser. A noun like *book* does not take a complement clause. Thus, it does not have a licensing feature for the C. Because of this information, the option of attaching C as a complement of N is excluded. This, in turn, means that if the predicted N can take a

complement clause, the predicted C can be attached as a complement. On the other hand, *book* or any noun can normally take CP as an RC. RCs or any clausal element are optional elements for a noun. A way to incorporate this information in the system is to assume that each N has licensing feature for C as an optional element. Thus, no matter what the lexical semantic property of the predicted N head might be, if an N is predicted, it can also be predicted that C is licensed by an optional element.[40] Here, let us stipulate that optional licensing features of each head are always accessible but need not to be checked. This is a plausible way to accommodate the optionality of elements such as modifiers, complements of NPs and so on. Under this view, if the category N can be predicted, the optional features of the predicted N head are also predicted. Based on the optional licensing feature for C, the predicted C and the predicted N can be combined.

A further modification is required for this approach. As we discussed earlier, for us the morphology of the embedded verb is crucial. Thus, we have to modify the parser so that it can predict the morphology of the embedded verb. The morphological requirement that we have to capture is that an RC must have predicate adnominal morphology on its embedded verb (or any predicate that is incorporated into an NP must have adnominal morphology). This means that a N and the form of the predicate that is associated with the N has a one-to-one relation. Given this one-to-one relation, we can plausibly assume that each N head has a specification for adnominal form. We

[40] The parser adopted by Gibson (1991) and Babyonyshev (1996) has the same type of mechanism. Their parser allows for the prediction of the optional element that can be 'ignored' if the lexical item compatible with the predicted optional node does not come into input.

can capture this requirement by the licensing features, i.e., N has an optional feature that licenses predicate adnominal form. In the case of RCs, that the embedded verb has the predicate adnominal form means that the complementizer is empty. Therefore, we can assume that each N head has an optional licensing feature for empty C.

Modified SPARSER

Let us see how these modifications work in predicting and incrementally structuring the upcoming RC in the environment of the classifier mismatch. We take an example of (62), repeated here as (69).

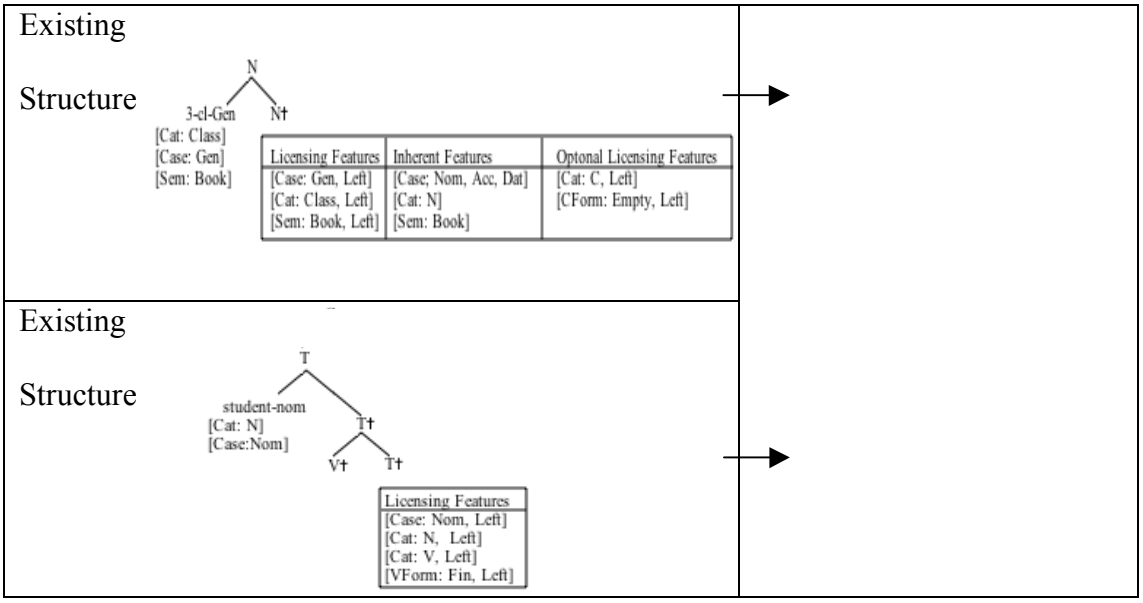
(69) san-satu-no [$_{NP}$[$_{CP}$[$_{IP}$ gakusee-ga [$_{VP}$ gap yon]-da]ø] hon]

three-cl$_{(book)}$-gen student-nom read-past-ø book

"Three books that the student read."

The problem of this construction arises after the parser commits to reanalysis. The crucial step is when the parser reaches (70).

(70)

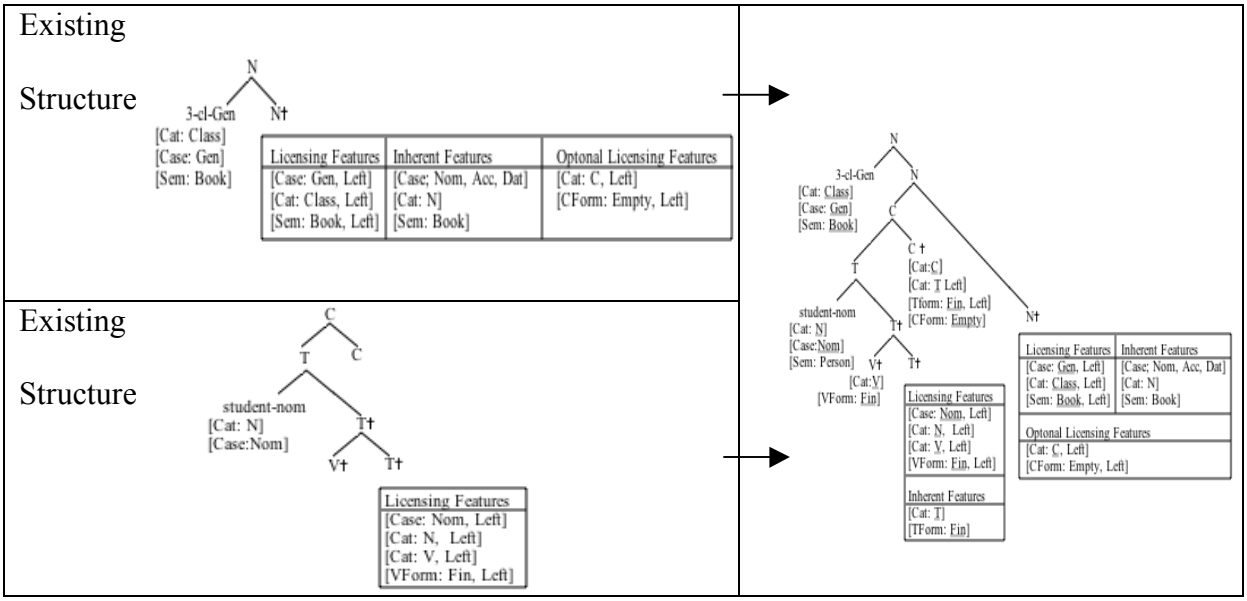


At (70), the N head that can license the classifier is predicted, i.e., the N head with the semantic feature of *book*. Additionally, the finite T head is predicted by means of the nominative-marked NP. The modified SPARSE can access the licensing features of the predicted N, including its optional licensing features. But no licensing features can license the predicted T head. Therefore, these two existing structures cannot be combined. Because these two heads cannot be combined, the further prediction must be set in which a head that can license T is searched for in the lexicon. As we have discussed, finite T must be licensed by C. Therefore, the lexicon returns a C head. The C is projected and combined with T. At the same point, the licensing feature of the finite T allows for the top-down prediction of the V head. Note that the category N, as the complement of V, cannot be predicted because the V's subcategorization information is not clear if a lexical V is not inserted. Thus, even though the verb can

be predicted, the complement of the verb cannot. Put differently, the predicted V in this case is not specific.

Now, the modified SPARSE can combine the predicted C and N. The optional licensing features of N can license the existing C head. Thus, the predicted C head can be combined with the predicted N head. The result is (71).

(71)



For the sake of explication, let us summarize the process from (70) to (71) step-by-step.

(72) Step 1: Search for features that can license [Case: Genitive].

Step 2: Project N head

Step 3: Search for features that can license [Case: Nom]

Step 4: project T head.

Step 5: Search for features that can license [T: finite]

Step 6: Project C head

Step 7: Check optional licensing feature [C: left] on the existing N head against existing C head

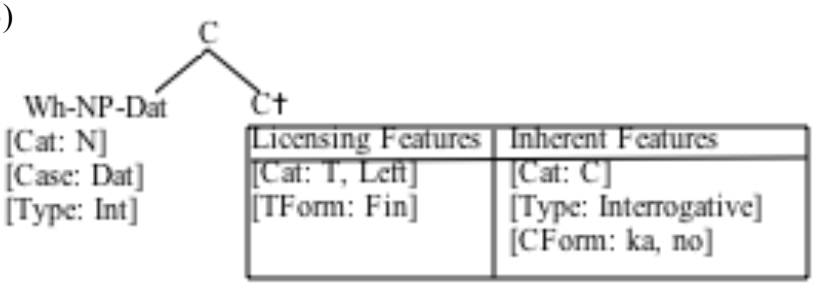
Step 8: Combine N head and C head.

An obvious prediction of this analysis is if the lexical verb is encountered and if it hosts an overt complementizer, the parser has to reanalyze the structure, resulting in a surprise effect. However, if a verb without a complementizer is encountered, the parser incorporates it into the structure without any problem. This is exactly what we saw in the case of classifier mismatch, i.e., we observed the effect at the embedded verb position.

We can ask if this enhancement is too powerful or not. If the predictive component is too powerful, it makes a wrong prediction. This enhancement is, however, well-constrained. Even though it allows for recursive predictions, it is not always possible. As is explicitly stated, these extra predictions are possible only if a specific head is predicted, and an attempt to combine the predicted structures is sure to failed.

The question we have to ask now is whether the predicted representation is sufficient to block LDB. Presumably the fronted wh-phrase can predict the following structure.

(73)



Other details aside, the C head predicted by means of fronted Wh and the C head predicted by the classifier mismatch does not match, specifically in the feature [CForm]. The interrogative C must have a Q-particle, and it should be specified in the inherent features. Because of this mismatch, the C head predicted by the classifier mismatch cannot check the features in C. Thus, the structure in (73) and (71) cannot be combined. As a consequence, LDB is blocked.

4.3. Dealing with the Direct Cue

Let us turn to how the SPARSE can manage a prediction of the upcoming conditional clauses. What we want to capture is how the conditional adverb *mosi* sets a prediction of the upcoming conditional verbal morphology. There are two key points. One is that *Mosi* has a direct relation to C head. The second key point is that *Mosi* is also licensed by T head. As it has been extensively explored in the syntax of conditionals, *Mosi* requires T head. Thus if we can sufficiently specify these relations

in terms of features, it will lead us to understand how *Mosi* predicts the conditional clauses and how the predicted conditional structure can block the LDB.

With the above discussion in mind, let us first see how the adverbs are treated in SPARSE. Schneider (1999) adopts the assumption from Categorial Grammar (Steedman 1996, 2000) in which an adjunct selects its host. This information can be represented by specifying the licensing features of each adverb. For example, a temporal adverb like *tomorrow* has the following licensing feature.

(74) tomorrow: *[Cat: V, Left]

(74) can be read as that *tomorrow* is licensed by the verb on its left. The asterisk on the licensing feature indicates that the lexical entry does not project. This feature specification is, however, not sufficient to capture the distribution of an adverb like *tomorrow*. As the example (75) shows, it can occur at the sentence final position. By the time when *tomorrow* is encountered, the parser has built the structure in (76).

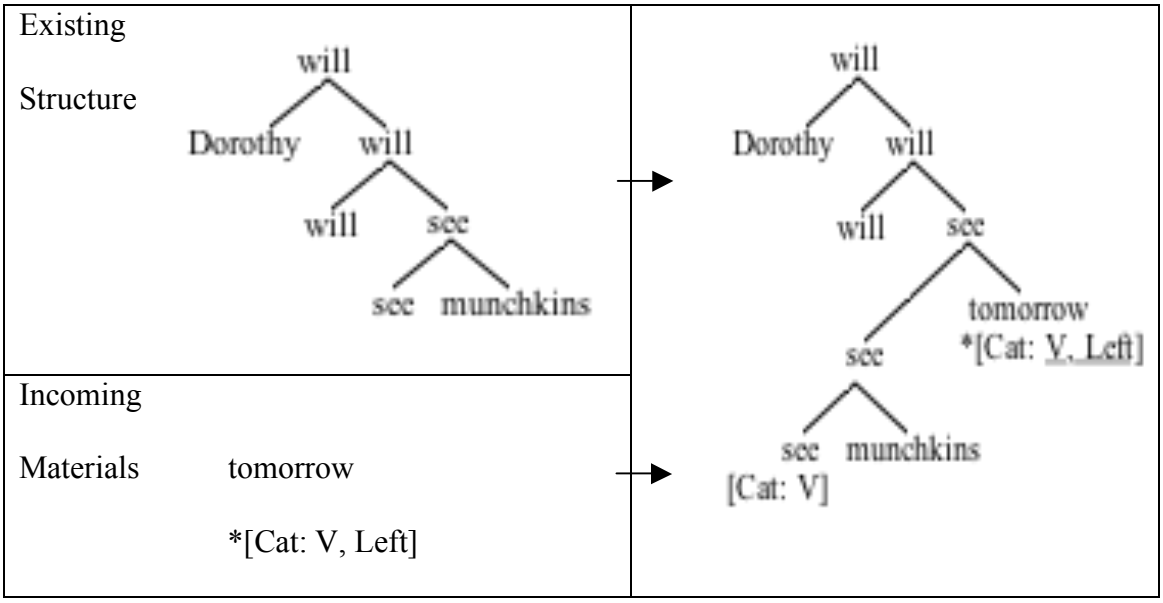
(75) Dorothy will see munchkins tomorrow.

(76)

Because of its licensing features, *tomorrow* cannot be attached to the most recently processed word, *munchkins*, which is an N. When a new word cannot be attached to the structure based on its licensing features, the parser searches the tree upwards along the right edge of the tree for a possible host for the word.[41] The first constituent that is encountered by the search, is the projection of *see*. Upon encountering *see*, the parser attempts to compare the features of *tomorrow* against the features on *see*. The licensing feature *[Cat: Verb, Left] of *tomorrow* allows the parser to combine it with *see*, and as a result, *see* project the new constituent. This step is illustrated in (77).

[41] For the details of the search process, see Schneider (1999). We will not go into details of this part because our discussion is no directly related to the issue of how a sentence final materials are incorporated into the structure.

(77)



This treatment of adverbs, however, has a problem with sentence initial adverbs as in (78), as Kazanina (2005) points out.

(78) Tomorrow Carmen will dance flamenco.

To accommodate sentence initial adverbs, Kazanina proposes that adverbs have the following features.

(79) Tomorrow: *[Cat: {[V, Left], [T, Right]}]

By revising the licensing features of adverbs in this way, Kazanina proposes that when the sentence initial adverb is encountered, it projects the T-head by means of

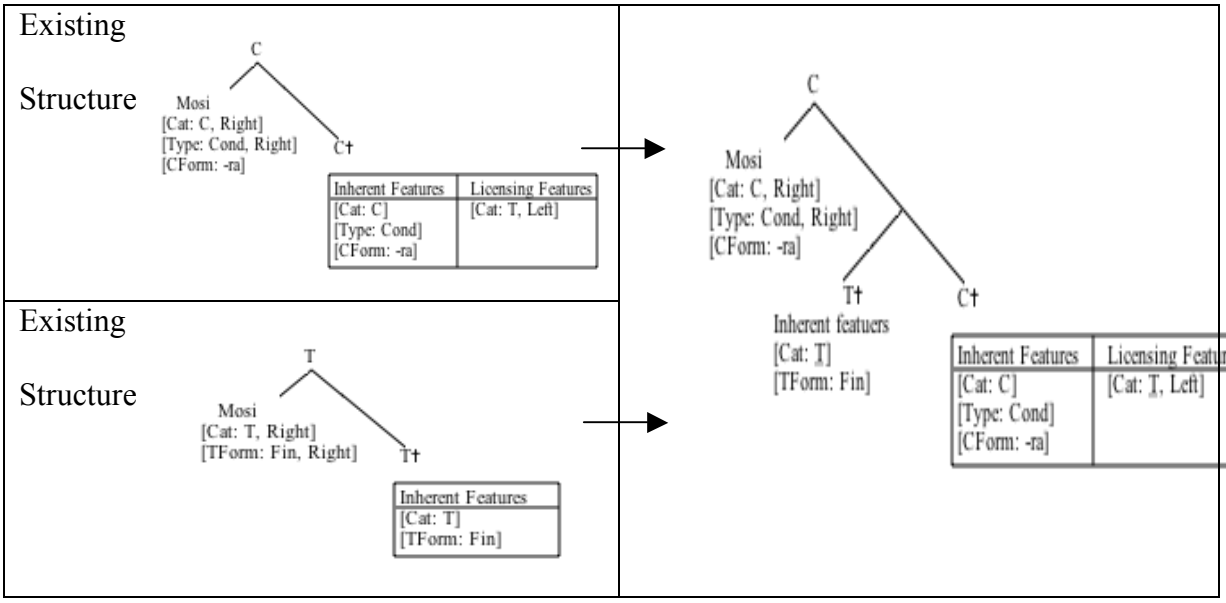
*[Cat: T, Right] feature, this is because there is no projection of V in the sentence when *tomorrow* is encountered in (78).

Because of the head-final property, Japanese adverbs always precede its host constituent. Thus, the situation is similar to the sentence initial adverbs in English. Following Kazanina's proposal, we can capture the licensing relation of *Mosi* and its hosts by specifying the licensing features. A possible way to do is as in (80). In (80), *Mosi*'s licensing features specify two types of licensing relations, i.e., the relation with specific type of C and the relation with the Finite Clause. Again we can specify a feature related to semantics. Here we adopt "clause type" as such semantic features. Let us call it "Type". Other than that the features we use are the same type of features we have used in the previous discussion.

(80) *Mosi*: [Cat: C, Right], [Type: Cond, Right] [CForm: -ra, Right],
[Cat: T, Right] [CForm: Fin, Right]

A potential problem of this solution is that it allows for the parser to predict multiple heads at the same time when the parser encounters *Mosi*. This parallel predicted heads can be illustrated as in (81).

(81)



The problem of this approach is that we need to constrain when the parallel prediction is possible and when it is not. However, we do not have any well-motivated constraints. Furthermore, allowing for this type of parallel prediction is almost equivalent to allow for the parallel structure building. Once we allow for parallel structure building in this case, we do not have any reason not to allow for that in the other cases. If the parallels structure building is allowed in general, it is possible to capture the left-edge ambiguity by the parallel structure building. If we can capture the left-edge ambiguity in this way, we do not need to postulate the flexible structure building mechanism like SPARSE in the first place. However, if we allow for the parallels structure building mechanism for ambiguous cases, we have to have a theory of which predicted structure is preferentially picked up by the parser, and which is not. This approach opens up to an important and interesting issue of the architecture of sentence processing mechanism. However, this approach may add unnecessary

complication to the overall architecture of the parser. So if possible it is better not going for this type of approach.

If we do not appeal to the parallel prediction, how can we capture the licensing relation between *Mosi* and two heads, T and C? Looking closely at the syntax of conditionals, it seems to be true that conditional C can only associate with a finite T. If so we can employ the same type of mechanism that we proposed in the case of indirect cue, i.e., using the licensing feature of the predicted head.

Because *Mosi* has really a tight connection to the conditional C, the possible range of the forms of conditional Cs is quite limited. Thus, we can assume that when *Mosi* predict the upcoming C head, the predicted C head is sufficiently specified. If so, the parser call for the licensing features of predicted C, exactly in the same way as what N head that is predicted by the classifier mismatch did in the prediction of RC structure. In order to accommodate this type of prediction, we have to just specify the licensing feature of *Mosi* and conditional C as in (82).

(82) a. *Mosi*: Licensing features

*[Cat: C, Right], [Type: Cond, Right], [CForm: -ra, right]

b. C: Inherent features,

[Cat: C] [Type: Cond], [CForm: -ra]

Licensing features

[Cat: T, Left], [TForm: Fin, Left]

The hidden assumption here is that C has at least three inherent features, the categorical feature, the semantic feature (the clause type), and specific form. If it is the case that just C is predicted and these features are not specified, the feature representation should look like C: [Cat: C], [Type: Declarative, Interrogative, Conditional …], [CForm: ø, -to, -ka …]. Because of the number possible clause types and CForms are restricted, it is not problematic to list all of these values in the lexicon.

Let us take an example in (83) and see how the parsing proceeds with these features.

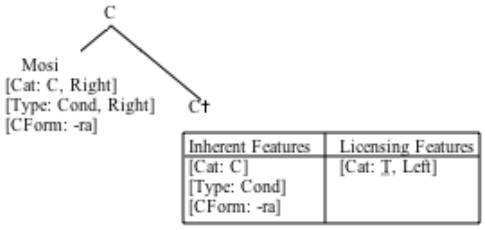
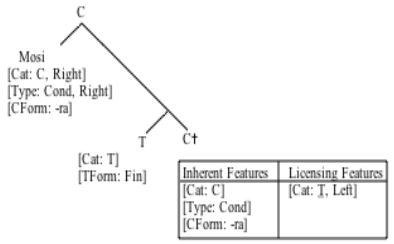
(83) `Mosi gakusee-ga  kaetta-ra`

`mosi student-nom leav-cond`

"If the student leaves …"

First, the parser encounters *Mosi*. The features of *Mosi* in (82a) are activated. At this point the C head is projected based on the licensing features of *Mosi*. Next, because the predicted C is sufficiently specified, its licensing features can be accessed by the parser. Based on C's licensing features, the T can be projected as the complement of C.

(84)

Mosi *[Cat: C, Right] [Type:Cond, Right] [CForm: -ra, right]		

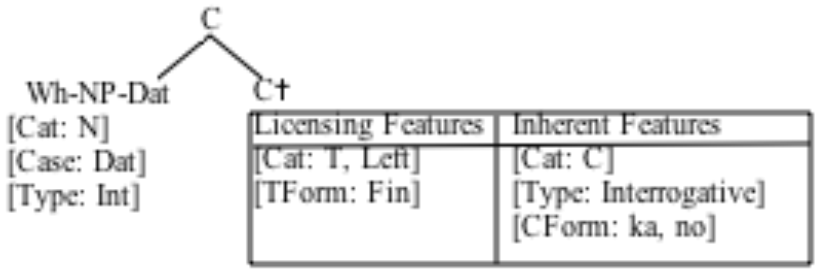
Based on the information on *Mosi* the parser can build this rich structure before the Subject is encountered.

In this system, again the specific verbal morphology is predicted. Thus, if the newly incoming lexical V does not have the predicted form, the parser has to initiate the reanalysis, resulting in the surprise effect. In the experimental studies on Conditional Clauses, we have seen that the effect took place at the position of the conditional verb. This system correctly predicts such effect at the conditional verb position.

With the discussion so far in mind, let us examine whether this structure can represent sufficient information to block the Longer Dependency or not.

Let us assume that the fronted wh-phrase can project the following structure.

(85)



Looking at the features of the predicted C, it is clear that the Inherent Features are not compatible with the predicted Conditional C. Specifically, the clause type feature and CForm features are not matched. If so, the C head predicted by wh-NP-dat cannot be associated with the C head predicted by *Mosi*. As a result, the C head predicted by Mosi and C head predicted by Wh-phrase cannot be combined. Thus it is possible to block the LDB using this incompatibility of the predicted C heads.

5. Conclusions

In this chapter we have see how we can modify SPARSE in order to predict sufficiently rich structure to block the LDB. We have revised SPARSE so that the parser can operate powerful recursive predictions. By this modification we made it possible to predict the specific verbal morphology that is not compatible with Q-particle. We have seen that this prediction of the verbal morphology in principle allows the parser to block the LDB in the predicted RCs or Conditional Clauses by means of the incompatibility of predicted Cs.

The remaining issue is how the LDB is actually blocked. To explore this issue we have to understand how the wh-feature drives the longer dependency. Aoshima's (2003) system is designed so that all the fronted dative NPs create LDB. However as we have seen, only the wh-phrases create longer dependency. How can we capture this distinction? This question is beyond the scope of this study. So we would like to leave the issue open here.[42]

[42] This issue is explored in Aoshima et al. (2006). Readers are recommended to refer to Aoshima et al.'s work for further details.

CHAPTER 6. GENERAL CONCLUSIONS

The preceding chapters investigated issues related to the strong incrementality of human sentence processing. I have argued that the strong incrementality is supported by a powerful predictive mechanism. Throughout this thesis, we have asked the following three questions. What are potential triggers for the predictive mechanism? What are the representations that the predictive mechanism generates? What is the appropriate algorithm that is responsible for the predictive mechanism? To answer these questions, we have investigated the interactions between wh-dependency formation and two specific constructions in Japanese, Relative Clauses and Conditional Clauses from a theoretical viewpoint (Chapter 2), a perspective of real-time sentence processing (Chapter 3 & 4), and a computational modeling (Chapter 5).

We have claimed that the parser of wh-dependencies is driven by the satisfaction of syntactic constraints. Specifically, we have claimed that wh-feature satisfaction is the driving force behind wh-dependency formation in real-time sentence processing. Based on this finding, we have investigated how wh-dependency formation interacts with the predictive mechanism. Through a series of experimental studies, we have discovered that when the upcoming RC structure and conditional clause structure are predicted, the long-distance wh-gap dependency is not created. On the basis of these findings, we tried to figure out how the human sentence processing can be strongly incremental.

In chapter 2, we have seen examined detailed syntactic properties of RCs and conditionals in Japanese. In this chapter we have revealed the following properties of

RCs and conditionals. We have argued that the derivation of RCs contains the movement of the empty operator on the one hand, and the structure of RC involves CP rather than just IP on the other. The first point, the movement of the empty operator is supported by various observations on the so-called connectivity effects. The second point, the CP analysis, is motivated by the formation of the so-called Predicate Adnominal morphology.

In this chapter, the detailed internal and external syntax conditional clauses have also been revealed. There were three important discoveries in this chapter. First, Japanese conditional clauses are not islands in Japanese, even though they are adjunct clauses. Second, conditional adverb *Mosi* is uniquely licensed by a finite clause and conditional morpheme. Finally, by examining the complex conditional morphology, we have revealed that conditional clause has multi-layered CP structure.

In chapter 3, we have seen the first piece of evidence for the strong incrementality. We have seen that the semantic incompatibility between the genitive-marked classifier and its adjacent NP can set a prediction of the upcoming RC structure. Based on the discovery that the wh-gap dependency does not penetrate into the predicted RC structure, we have argued that the predicted structure must be rich enough to represent the islandhood of RCs.

In chapter 4, we have discovered the second piece of evidence for the strong incrementality. We have seen that the upcoming conditional clause is predicted when the conditional adverb *Mosi* is provided. Based on the observation that the LDB is blocked also in the predicted conditional clauses, we have argued that *Mosi* sets a

prediction of the upcoming conditional morphology that is not compatible with the motivation for the LDB.

In chapter 5, we have discussed what type of algorithm can derive the strong incrementality without sacrificing flexible structure building. We have argued that a type of Left-Corner algorithm is psychologically more plausible algorithm for human sentence processing. We adopted a particular parsing mechanism called SPARSE. We argued that SPARSE can incrementally build the structure without sacrificing the necessary flexibility. However, even SPARSE cannot predict the full structure of the upcoming RCs or Conditional Clauses. In order to capture the prediction of the upcoming RCs and Conditionals, we have modified the predictive component of SPARSE. Our modification was as follows. A top-down prediction based on the predicted head is possible only if the parser can predict a specific head with specific semantic content. With this modification, the parser can set recursive prediction only in limited environment.

This work has focused on the processing of relative clauses and conditional clauses, and we have seen that sentence processing is strongly incremental. We have further argued that the strong incrementality is supported by a powerful predictive mechanism. However, there are some remaining issues in this work. We have not made clear how the long-distance wh-gap dependency can be stopped in the predicted RCs and Conditionals. We have shown how these structures are predicted but we have not shown how these predicted structures interact with the LDB. To resolve this issue, we have to understand how the longer dependency formation is created. To

achieve the understanding of the LDB it is needed to discover the mechanism of the longer dependency formation.

Additionally, to examine empirical validity of the proposed parsing mechanism, we have to experimentally investigate the environment where the unambiguous structure is predicted before the head is encountered. In this respect, more studies on head-final languages as well as head-initial languages are necessary.

Bibliography

Abney, Steven P, and Johnson, Mark. 1991. Memory Requirements and Local Ambiguities of Parsing Strategies. Journal of Psycholinguistic Research 20:233-250.

Abney, Steven. 1989. A Computational Model of Human Parsing. Journal of Psycholinguistic Research 18:129-144.

Aho, A, and Ullman, J. 1972. The Theory of Parsing, Translation, and Compiling (Vol. 1). Englewoods Cliffs, NJ: Prentice-Hall.

Aoshima, Sachiko, Phillips, Colin, and Weinberg, Amy. 2004. Processing Filler-Gap Dependencies in a Head-Final Language. Journal of Memory and Language 51:23.

Aoyagi, Hiroshi. 1998. On the Nature of Particles in Japanese and Its Theoretical Implications, University of Southern Calrifornia: Ph.D. dissertation.

Auster, Paul. 1985. *City of Glass*. Los Angels, CA: Sun and Moon Press (Reprinted in Auster, Paul. 1987. *The New York Trilogy*)

Auster, Paul. 1986. *Ghosts.* Los Angels CA: Sun and Moon Press (Reprinted in Auster, Paul. 1987. *The New York Trilogy*)

Ausster, Paul. 1987. *The New York Trilogy*. New York, NY: Penguin Books

Auster, Paul. 1989. *Moon Palace*. New York, NY: The Viking Press.

Auster, Paul. 1994. *Mr. Vertigo*. New York, NY: The Viking Press.

Auster, Paul. 2004. *Oracle Night*. New York, NY: Henry Holt & Company.

Babyonyshev, Maria, and Gibson, Edward. 1995. Processing Overload in Japanese. MIT Working Papers in Linguistics 26:1-35.

Bader, Markus, and Lasser, Ingeborg. 1994. German Verb-Final Clause and Sentence Processing: Evidence for Immediate Attachment. In Perspectives on Sentence Processing, eds. Charles Clifton Jr., Lyn Frazier and Keith Reiner, 225-242: Lawrence Erlbaum.

Beck, Sigrid. 1996. Quantified Structures as Barriers for Lf Movement. Natural Language Semantics 4:1-56.

Berwick, Robert C. 1991. Principles of principle-based parsing. In Principle-based parsing: Computation and psycholinguistics, ed. R.C. Berwick, 1-37: Kluwer Academic Publishers.

Bever, Thomas G. 1970. The Cognitive Basis for Linguistic Structures. In Cognition and the Development of Language, ed. J. R. Hayes, 279-360. New York: Wiley.

Bhatt, Rajesh, and Pancheva, Roumyana. 2001. Conditionals. Ms. University of Texas, Austin, and USC.

Bianchi, Valentina. 2000. The Raising Analysis of Relative Clauses: A Reply to Borsley. Linguistics Inquiry 31:123-140.

Boeckx, Cedric, and Niinuma, Fumikazu. 2003. Conditions on agreement in Japanese. Natural Language and Linguistic Theory 22:453 - 480.

Boland, Julie E., and Boehm-Jernigan, Heather. 1998. Lexical constraints and prepositional phrase attachment. Journal of Memory and Language 39:684.

Boland, Julie E., Tanenhaus, Michael K., Garnsey, Susan M., and Carlson, Greg N. 1995. Verb Argument Structure in Parsing and Interpretation: Evidence from Wh-Questions. Journal of Memory and Language 34:774.

Boskovic, Zeljko, and Takahashi, Daiko. 1998. Scrambling and Last Resort. Lingustic Inquiry 29:347-366.

Browning, Marguerite. 1987. Null Operator Constructions. Ph.D. Dissertation, Massachusetts Institute of Technology.

Chomsky, Noam. 1972. Studies on Semantics in Generative Grammar. The Hague: Mouton.

Chomsky, Noam. 1973. Conditions on Transformations. In A Festschrift for Morris Halle, eds. Stephen Anderson and Paul Kiparsky, 232-286. New York: Holt Rinehart & Winston.

Chomsky, Noam. 1977. On Wh-Movement. In Formal Syntax, eds. Peter Culicover, Thomas Wasow and Adrian Akmajian, 71-132. New York: Academic Press.

Chomsky, Noam. 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, Mass.: MIT Press.

Chomsky, Noam. 1995. The Minimalist Program. Cambridge, Mass.: The MIT Press.

Cinque, Guglielmo. 1984. A-Bar Bound Pro Vs. Variable. Ms. University of Venice.

Cinque, Guglielmo. 1990. Types of A'-Dependencies: Linguistic Inquiry Monographs; 17. Cambridge, Mass.: MIT Press.

Citko, Barbara. 2005. On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge. Linguistc Inquiry 36:475-497.Collins, Chris. 1998. A Note on Extraction from Conditionals. In Cornell Working Papers in Linguistics, eds. N. Adisasmito-Smith and T. Doeleman, 57-66. Ithaca, NY: CLC, Cornell University.

Collins, Chris. 1998. A Note on Extraction from Conditionals. In Cornell Working Papers in Linguistics, eds. N. Adisasmito-Smith and T. Doeleman, 57-66. Ithaca, NY: CLC, Cornell University.

Crocker, Matthew W. 1994. On the Nature of Principle-Based Sentence Processor. In Perspectives on Sentence Processing, eds. Charles Clifton Jr., Lyn Frazier and Keith Rayner, 245-266. Hillsdale, New Jersey: Lawrence Earlbaum Associates, Publishers.

Crocker, Matthew W. 1996. Computational Psycholinguistics: An Interdisciplinary Approach to the Study of Language: Studies in Theoretical Psycholinguistics; V. 20. Dordrecht; Boston: Kluwer Academic Publishers.

Frank, Robert. 2002. Phrase Structure Composition and Syntactic Dependencies. Cambridge, Mass.: MIT Press.

Frazier, Lyn, and Rayner, Keith. 1987. Resolution of Syntactic Category Ambiguities: Eye Movements in Parsing Lexically Ambiguous Sentences*1. Journal of Memory and Language 26:505.

Frazier, Lyn. 1987. Syntactic Processing: Evidence from Dutch. Natural Language and Linguistic Theory 5:519-559.

Fukui, Naoki, and Takano, Yuji. 2000. Nominal Structure: An Extension of the Symmetry Principle. In The Derivation of Vo and Ov, ed. Peter Svenonius, 210-254. Amsterdam: John Benjamins.

Fukui, Naoki. 1995. Theory of Projection in Syntax: Studies in Japanese Linguistics. Stanford, Calif.: CSLI Publications; Tokyo: Kurosio Publishers.

Gibson, Edward. 1991. A Computational Theory of Human Linguistic Processing: Memory Limiations and Processing Breakdown, Carnegie Mellon University.

Gibson, Edward. 1998. Linguistic Complexity: Locality of Syntactic Dependencies. Cognition 68:1-76.

Gorrell, Paul. 1995. Syntax and Parsing: Cambridge University Press.

Grimshaw, Jane. 1991. Extended Projection, Brandeis University: Ms.

Guimarães, Maximiliano. 2004. Derivation and Representation of Syntactic Amalgams. Ph.D. Dissertation, University of Maryland.

Hagstrom, Paul. 1998. Decomposing questions. Ph.D. Dissertation, Massachusetts Institute of Technology.

Han, Chung-Hye, and Romero, Maribel. 2004. The Syntax of Whether/Q.Or Questions: Ellipsis Combined with Movement. Natual Language and Linguistic Theory 22:527-564.

Harada, Kazuko Ito. 1972. Constraints on Wh-Q Binding. In Studies in Descriptive and Applied Linguistics 5, 180-206. Tokyo: Division of Language, International Christian University.

Harada, Shin-Ichi. 1971. Ga-No Conversion and Idiolectal Variation in Japanese. Gengokenkyu 60:25-38.

Harada, Shin-Ichi. 1973. Counter Equi Np Deletion. In Annual Bulletin, Research Institute of Logopedics and Phoniatrics, 133-143: University of Tokyo.

Harada, Shin-Ichi. 1973a. Constraints on Binding. Studies in English Literature:41-72.

Harada, Shin-Ichi. 1973b. Counter Equi NP Deletion. In Annual Bulletin, Research Institute of Logopedics and Phoniatrics, 133-143: University of Tokyo.

Harada, Shin-Ichi. 1974. Remarks on relativization. Annual Bulletin, Reserach Institute of Logopedics and Phoniatrics 7:113-147.

Harada, Shin-Ichi. 1976. Ga-No Conversion Revisited. Gengokenkyu 70:23-38.

Harada, Shin-Ichi. 1976. Honorifics. In Syntax and semantics 5: Japanese generative grammar, ed. Masayoshi Shibatani, 499-561. New York: Academic Press.

Hendriks, Petra. 2003. "Either" as a Focus Particle. Ms. Center fo Language and Cognition, University of Groningen.

Hinds, John. 1973a. Some Remarks on Soo Su-. Papers in Japanese Linguistics 2:18-30.

Hinds, John. 1973b. Some Remarks on Soo Su-. Papers in Japanese Linguistics 2:18-30.

Hiraiwa, Ken. 2000. On Nominative-Genitive Conversion. MIT Working Papers in Linguistics 39.

Hiraiwa, Ken. 2001. Nominative-Genitive Conversion Revisited, MIT: ms.

Hiraiwa, Ken. 2002. Facets of Case: On the Nature of the Double-O Constraint, MIT: ms.

Hirose, Yuki. 1999. Resolving Reanalysis Ambiguity in Japanese Relative Clauses, Graduate Faculty in Linguistics, City University of New York.

Hoji, Hajime, Miyagawa, Shigeru, and Tada, Hiroaki. 1989. Np-Movement in Japanese. Ms. University of Southern California, Los Angels, Ohio State University, Columbus, and Massachusetts Institute of Technology, Cambridge.

Hoji, Hajime. 1985. Logical Form Constraints and Configurational Structure in Japanese, University of Washington: Doctoral Dissertation.

Honcoop, Martin. 1998. Excursions in Dynamic Binding. Ph.D. Dissertation, Leiden University.

Hornstein, Norbert, and Uriagereka, Juan. 2002. Reprojections. In Derivation and Explanation in the Minimalist Program, eds. David Samuel Epstein and T. Daniel Seely, 106-132. Oxford, UK

Hoshi, Koji. 1995. Structural and Interpretive Aspects of Head-Internal and Head-External Relative Clauses, University of Rochester: Ph.D. dissertation.

Hoshi, Koji. 2004. Parametrization of the External D-System in Relativization. Language, Culture and Communication 33:1-50.

Huang, Cheng-Teh James. 1982. Logical Relations in Chinese and the Theory of Grammar, Massachusetts Institute of Technology: PhD.

Inoue, Atsu, and Fodor, Janet. D. 1995. Information-Paced Parsing of Japanese. In Japanese Sentence Processing, eds. R. Mazuka and N. Nagai, 9-36. Hillsdale, NJ: Lawrence Erlbaum Associates.

Inoue, Atsu. 1991. A Comparative Study of Parsing in English and Japanese, Linguistics, The University of Connecticut.

Inoue, Kazuko. 1976. Henkee Bunpoo to Nihongo [Transformational Grammar and Japanese Language]. Tokyo: Taishuukan Shoten.

Ishii, Toru. 1997. An Asymmetry in the Composition of Phrase Structure and Its Consequences. Ph.D. Dissertation, University of California, Irvine.

Ishii, Yasuo. 1991. Operators and Empty Categories in Japanese, University of Connecticut: Ph.D. dissertation.

Jackendoff, Ray. 1972. Semantic Interpretation in Generative Grammar: Studies in Linguistics Series, 2. Cambridge, Mass.: MIT Press.

Jayaseelan, K. A. 1990. Incomplete Vp Deletion and Gapping. Linguistic Analysis 20:64-81.

Johnson-Laird, P. N. 1983. Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness: Cognitive Science Series; 6. Cambridge, Mass.: Harvard University Press.

Just, Marcel Adam, Carpenter, Patricia A., and Woolley, Jacqueline D. 1982. Paradigms and Processes in Reading Comprehension. Journal of Experimental Psychology 111:228-238.

Kamide, Yuki, and Mitchell, Don C. 1999. Incremental Pre-Head Attachment in Japanese Parsing. Language and Cognitive Process 14:631-662.

Kamio, Akio. 1983. Meisiku-No Koozoo [the Structure of Noun Phrases]. In Nihongo-No Kihonkoozoo [the Basic Structure of Japanese], ed. Kazuko Inoue. Tokyo: Sanseidoo.

Kaplan, Ronald M., and Zaenen, Annie E. 1989. Long-Distance Dependencies, Constituent Structure, and Functional Uncertainty. In Alternative Concepts of Phrase Structure, eds. Mark Baltin and Anthony S. Kroch, 17-42. Chicago: University of Chicago Press.

Katada, Fusa. 1991. The Lf Representation of Anaphors. Linguistic Inquiry 22:287-313.

Kato, Takaomi. 2005a. Neg-Raising, Last Resort, and the Nature of CSC. Harvard Working Papers in Linguistics 11:79-102.

Kato, Takaomi. 2005b. A Case against the Representational Approach to the Coordinate Structure Constraint. In Proceedings of the Thirty-Fifth Annual Meeting of the North East Linguistic Society (NELS 35), eds. Leah Bateman and Cherlon Ussery. Amherst: GLSA.

Kato, Yasuhiko. 1985. Negative Sentences in Japanese. Sophia Linguistica 19.

Kayne, Richard S. 1994. The Antisymmetry of Syntax. Cambridge: Mass: MIT Press.

Kikuchi, Akira, Oishi, Masayuki, and Yusa, Noriaki. 1994. Scrambling and Relativized L-Relatedness. In Mit Working Papers in Linguistics 24: Formal Approaches to Japanese Linguistics 1, 141-158. Cambridge: Mass: MIT Press.

Kim, Shin-Sook. 2002. Intervention Effects Are Focus Effects. In Japanese and Korean Linguistics 10, eds. Noriko Akatsuka and Susan Strauss, 615-628. Stanford: CSLI Publications.

Kimball, John P. 1973. Seven Principles of Surface Structure Parsing in Natural Language. Cognition 2:15-47.

Kimball, John P. 1975. Predictive Analysis and over-the-Top Parsing. In Syntax and Semantics, ed. John P. Kimball. New York, NY.: Academic Press.

Kitagawa, Yoshihisa (2006) Naze. Ms. Inidiana University.

Kroch, Anthony S. 1989. Amount Quantification, Referentiality and Long Wh-Movement. Ms. University of Pennsylvania.

Kroch, Anthony S., and Joshi, Aravind K. 1985. The Linguistic Relevance of Tree Adjoining Grammar. University of Pennsylvania Department of Computer and Information Science Technical Report MS-CIS-85-16.

Kuno, Susumu. 1973. The Structure of the Japanese Language. Cambridge, Mass.: MIT Press.

Kuroda, S-Y. 1988. Whether We Agree or Not: A Comparative Syntax of English and Japanese. In Papers from the Second International Workshop on Japanese Syntax, ed. W. Poster, 103-143: CSLI.

Kuroda, Shige-Yuki. 1965. Generative Grammatical Studies in the Japanese Language, MIT: Ph.D. dissertation.

Kuroda, Shige-Yuki. 1974. Pivot-independent relative clauses in Japanese i. Papers in Japanese Linguistics 3:59-93.

Kuroda, Shige-Yuki. 1975/76. Pivot-Independent Relative Clauses in Japanese Ii. Papers in Japanese Linguistics 4:85-96.

Kuroda, Shige-Yuki. 1976/77. Pivot-independent relative clauses in Japanese, iii. Papers in Japanese Linguistics 5.

Kuroda, Shige-Yuki. 1986a. What Happened Agter the Movement of Nps in La Jolla? In Working Papers from the First Sdf Workshop in Japanese Syntax, ed. Shige-Yuki Kuroda, 71-96. La Jolla: University of California San Diego.

Kuroda, Shige-Yuki. 1986b. Movement of Noun Phrases in Japanese. In Issues in Japanese Linguistics, eds. Takashi Imai and Mamoru Saito, 229-271. Dordrecht: Foris.

Kuroda, Shige-Yuki. 1999. Syuyoobunaizaigata Kankeisetu [Head Internal Relative Clauses]. In Kotoba No Kaku to Shuuen [Core and Periphery of Language], eds. Shige-Yuki Kuroda and Masaru Nakamura. Tokyo: Kurosio Publisher.

Lakoff, George, and Ross, John Robert. 1976. Why Can't Do So into the Sink? In Syntax and Semantics, Volume 7: Notes from the Linguistic Underground, ed. James D. McCawley, 101-111. New York: Academic Press.

Larson, Richard K. 1985. On the Syntax of Disjunction Scope. Natural Lanaguage and Linguistic Theory 3:217-264.

Lasnik, Howard, and Saito, Mamoru. 1992. Move A: Conditions on Its Application and Output. Cambridge: The MIT Press.

Lasnik, Howard, Depiante, Marcela A., and Stepanov, Arthur. 2000. Syntactic Structures Revisited: Contemporary Lectures on Classic Transformational Theory. Cambridge, Mass.: MIT Press.

Lasnik, Howard. 1981. Restricting the theory of transformations: A case study. In Explanation in linguistics, ed. Norbert & David Lightfoot Hornstein, 152-173: Longman.

Lasnik, Howard. 1981a. Restricting the Theory of Transformations: A Case Study. In Explanation in Linguistics, ed. Norbert & David Lightfoot Hornstein, 152-173: Longman.

Lasnik, Howard. 1981b. Restricting the Theory of Transformation: A Case Study. In Expanations in Linguistics, eds. Norbert Hornstein and David Lightfoot, 152-173. London: Longmans.

Lewis, Richard L. 1996. Interference in Short-Term Memory: The Magical Number Two (or Three) in Sentence Processing. Journal of Psycholinguistics Research 25:93-115.

Linebarger, Marcia C. 1987. Negative Polarity and Grammatical Representation. Linguistics and Philosophy 10:325-387.

Lombardo, Vincenzo, and Sturt, Patrick. 2002. Incrementality and Lexicalism; a Treebank Study. In The Lexical Basis of Sentence Processing: Formal, Computational and Experimental Issues, eds. Paola Merlo and Suzanne Stevenson. Amsterdam/ Philadelphia: John Benjamins.

Macdonald, Maryellen C, Pearlmutter, Neal, and Seidenberg, Mark S. 1994a. Syntactic Ambiguity Resolution as Lexical Ambiguity Resolution. In Perspectives on Sentence Processing, eds. Charles Clifton Jr., Lyn Frazier and Keith Rayner, 123-154. Hillsdale, New Jersey: Lawrence Earlbaum Associates, Publishers.

MacDonald, Maryellen C., Pearlmutter, Neal J., and Seidenberg, Mark S. 1994b. Lexical nature of syntactic ambiguity resolution. Psychological Review 101:676-703.

Marslen-Wilson, William. 1973. Linguistic Structure and Speech Shadowing at Very Short Latencies. Nature 244:522-533.

Masuoka, Takashi, and Takubo, Yukinori. 1992. Kiso Nihongo Bunpoo [the Basic Japanese Grammar]. Tokyo: Kuroshio Publishers.

Masuoka, Takashi, and Takubo, Yukinori. 1992. Kiso Nihongo Bunpoo [the Basic Japanese Grammar]. Tokyo: Kuroshio Publishers.

Matsuda, Yuki. 1993. On the So-Called Japanese Headless Relative Clauses: Complementation Analysis. Ms. University of Southern California.

Mazuka, Reiko, and Itoh, Kenji. 1995. Can Japanese Speakers Be Led Down the Garden Path? In Japanese Sentence Processing, eds. Rekio Mazuka and Noriko Nagai: Lawrence Erlbaum Associates.

Mazuka, Reiko, and Lust, Barbara. 1988. Why Is Japanese Not Difficult to Process? A Proposal to Integrate Parameter Setting in Universal Grammar and Parsing. Paper presented at NELS.

Mazuka, Reiko, and Lust, Barbara. 1990. On Parameter Setting and Parsing: Predictions for Cross-Linguistic Differences in Adult and Child Processing. In Language Processing and Language Acquisition, ed. Lyn Frazier and Jill de Villiers, 163-205. Dordrecht: Kluwer Academic Publishers.

Mazuka, Reiko, Itoh, Kenji, Kiritani, Shigeru, Niwa, Shin-Ichi, Ikejiri, Kayo, and Naitoh, Kazuo. 1989. Processing of Japanese Garden Path, Center-Embedded, and Multiply-Left-Embedded Sentences: Reading Time Data from an Eye Movement Study. Annual Bulletin, Research Institute of Logopedics and Phoniatrics 23:187-212.

Mazuka, Reiko. 1991. Processing of Empty Categories in Japanese [MAY]. Journal of Psycholinguistic Research 20:215-232.

McCawley, James D. 1982. Parentheticals and Discontinuous Constituent Structure. Linguistic Inquiry 13:91-106.

Mihara, Kenichi. 1994. Nihongo-No Toogo Koozoo: Seesee Bunpoo Riron to Sono Ooyoo [Syntactic Structures of Japanese: The Theory of Generative Grammar and Its Application]. Tokyo: Shohakusha.

Mikami, Akira. 1960. Zoo Wa Hana Ga Nagai. Tokyo: Kurosi Syuppan.

Miller, George A, and Chomsky, Noam. 1963. Finitary Models of Language Users. In Handbook of Mathematical Psychology Volume 2, ed. R Luce, R. Bush & E. Galanter, 419-491: John Wiley.

Miyagawa, Shigeru. 1989. Structure and Case Marking in Japanese: Syntax and Semantics; V. 22. San Diego: Academic Press.

Miyagawa, Shigeru. 1993. Lf Case-Checking and Minimal Link Condition. In Papers on Case and Agreement. Mit Working Papers in Linguistics, ed. Colin Phillips. Cambridge, MA: MIT Working Papers in Linguistics.

Miyagawa, Shigeru. 1998. On Islands. Ms. Massachusetts Institute of Technology.

Miyagawa, Shigeru. 2002. The Nature of Weak Islands. Ms. Massachusetts Institute of Technology.

Miyamoto, Edson, and Takahashi, Shoichi. 2002. The Processing of Wh-Phrases and Interrogative Complementizers in Japanese. Paper presented at Proceedings of the 10th Japanese/Korean Linguistic Conference.

Miyamoto, Edson. T. 2002. Case markers as clause boundary inducers in Japanese. Journal of Psycholinguistic Research 31:307-347.

Miyamoto, Edson. T. 2003. Reanalysis of clause boundaries in Japanese as a constraint-driven process. Language and Speech 46:23-52.

Moltmann, F. 1992. Coordination and Comparatives. Ph.D. Dissertation, Massachusetts Institute of Technology.

Munn, Alan B. 1993. Topics in the Syntax and Semantics of Coordinate Structures. Ph.D. Dissertation, University of Maryland.

Muraoka, Satoru, and Sakamoto, Tsutomu. 2003. The influence of lexical information on the clause boundary processing in japanese. [nihongo-ni okeru setsu-kyoukai kaiseki-ji-no goi-joohoo-no eekyoo]. Paper presented at Paper presented at Japanese Cognitive Science Society, University of Electro-Communications.

Murasugi, Keiko. 1991. Noun Phrases in Japanese and English: A Study in Syntax, Learnability and Acquisition, University of Connecticut, Storrs: Ph.D. dissertation.

Nakamura, Michiko. 1999/2000. Processing of Relative Clauses: A Psycholinguistic Study of Adult Native Speakers of Japanese. In University of Hawaii at Manoa Working Papers in Linguistics, 59-71.

Nakamura, Michiko. 2003. Processing of Multiple Filler-Gap Dependencies, The Graduate Division, The University of Hawaii.

Nakau, Minoru. 1973. Sentential Complementation in Japanese. Tokyo: Kaitakusha.

Nemoto, Naoko. 1993. Chains and Case Positions: A Study from Scrambling in Japanese, University of Connecticut: Ph.D. dissertation.

Nishigauchi, Taisuke. 1990. Quantification in the Theory of Grammar. Dordrecht: Kluwer.

Ochi, Masao. 2001. Move F and Ga/No Conversion in Japanese.

Pesetsky, David Michael. 2000. Phrasal Movement and Its Kin: Linguistic Inquiry Monographs; 37. Cambridge, Mass.: MIT Press.

Pesetsky, David. 1987. Wh in Situ: Movement and Unselective Binding. In The Representation of (in)Definiteness, eds. Eric J. Rueland and Alice G. B. ter Meulen. Cambridge, Mass: MIT Press.

Pollard, Carl Jesse, and Sag, Ivan A. 1994. Head Driven Phrase Structure Grammar. Chicago: University of Chicago Press.

Pritchett, Bradley L. 1991a. Subjacency in a Principle-Based Parser. In Principle-Based Parsing: Computation and Psycholinguistics, eds. Robert C. Berwick, Steven. P. Abney and Carol Lee Tenny, 301-345. Dordrecht: Kluwer Academic Publisher.

Pritchett, Bradley L. 1991b. Head Position and Parsing Ambiguity. Journal of Psycholinguistic Research 20:251-270.

Pritchett, Bradley L. 1992a. Grammatical Competence and Parsing Performance. Chicago, IL: The University of Chicago Press.

Pritchett, Bradley L. 1992a. Parsing with Grammar: Islands, Heads, and Garden Paths. In Island Constraints, ed. Helen Goodluck, & Michael Rochemont, 321-349: Kluwer Academic Publishers.

Pritchett, Bradley. 1988. Garden Path Phenomena and the Grammatical Basis of Language Processing. Language 64:539-576.

Resnik, Philip. 1992. Left-Corner Parsing and Psychological Plausibility. In The Proceedings of the Fifteenth International Conference on Computational Lniguistics, Coling-92, 191-197.

Rizzi, Luigi. 1982. Violations of the wh island constrain and the subjacency condition. In Issues in italian syntax, ed. Luigi Rizzi. Dordrecht: Foris.

Rizzi, Luigi. 1990. Relativized Minimality: Linguistic Inquiry Monographs; 16. Cambridge, Mass.: MIT Press.

Rochemont, Michael S. 1986. Focus in Generative Grammar. Amsterdam; Philadelphia: J. Benjamins Pub. Co.

Rohde, Douglas. 2001-2003. Linger: A Flexible Platform for Language Processing Experiments: MIT.

Rooth, Matts, and Partee, Barbara Hall. 1982. Conjunction, Type Ambiguity and Wide Scope Or. In Proceedings of the First West Coast Conference on Formal Linguistics (Wccfl 1), eds. M Macken and N Wiegand.

Ross, John R. 1967. Constraints on Variables in Syntax, Linguistics, MIT: Ph.D.

Ross, John Robert. 1984. Inner Islands. In Proceedings of the Tenth Annucal Meeting of the Berkeley Linguistic Society, eds. C. Brugmann and M. Macaulay. Berkley, California: Berkely Linguistic Society.

Rullmann, Hotze. 1995. Maximality in the Semantics Os Wh-Constructions. Amherst, MA: GLSA (Graduate Linguistic Student Association), Department of Linguistics, University of Massachusetts.

Sag, Ivan A. 1976. Deletion and Logical Form, Ph.D. Dissertation, Massachusetts Institute of Technology.

Saito, Mamoru. 1982. Case Marking in Japanese: A Preliminary Studies, MIT: Generals paper.

Saito, Mamoru. 1983. Case and Government in Japanese. Proceedings of WCCFL 2:247-259.

Saito, Mamoru. 1985. Some Asymmetries in Japanese and Their Theoretical Implications, MIT: Ph.D. dissertation.

Saito, Mamoru. 1989. Scrambling as Semantically Vacuous a'-Movement. In Alternative Conceptions of Phrase Structure, eds. Mark Baltin and Anthony S. Kroch, 182-200. Chicago: The University of Chicago Press.

Sakai, Hiromu. 1994. Complex Np Constraint and Case-Conversions in Japanese. In Current Topics in English and Japanese, ed. Masaru Nakamura, 179-203. Tokyo: Hituzi Syobo.

Schneider, David Andrew. 1999. Parsing and Incrementality, University of Delaware: Ph.D. dissertation.

Schwarz, Bernhard. 1999. On the Syntax of Either.Or. Naturaal Language and Lingusitic Theory 17:339-370.

Shibatani, Masayoshi. 1978. Nihongo No Bunseki [the Analysis of Japanese]. Tokyo: Taishuukan.

Shimoyama, Junko. 2001. Wh-Constructions in Japanese, University of Masachusetts: Ph.D. dissertation.

Stabler, Edward P. 1994. The Finite Connectivity of Linguistic Structure. In Perspectives on Sentence Processing, eds. Charles Jr. Clifton, Lyn Frazier and Keith Rayner, 303-336. Hillsdale, New Jersey: Lawrence Earlbuam Associates, Publishers.

Steedman, Mark. 1996. Surface Structure and Interpretation: Linguistic Inquiry Monographs; 30. Cambridge, Mass.: MIT Press.

Steedman, Mark. 2000. The syntactic process. Cambridge, Massachusetts: The MIT Press.

Stepanov, Arthur V. 2001. Cyclic Domains in Syntactic Theory, University of Connecticut: Ph.D. dissertation.

Sturt, Patrick, and Crocker, Matthew W. 1995. Incrementality and Monotonicity in Syntactic Parsing. Edinburgh Working Papers in Cognitive Science 11:23-66.

Sturt, Patrick, and Crocker, Matthew W. 1996. Monotonic Syntactic Processing: A Cross-Linguistic Study of Attachment and Reanalysis. Language and cognitive processes 11:449-494.

Sturt, Patrick, and Lombardo, Vincenzo. 2005. Processing coordinate structures: Incrementality and connectedness. Cognitive Science:291-305.

Sturt, Patrick. 1997. Syntactic reanalysis in human language processing, University of Edinburgh: Ph. D. Dissertation.

Szabolcsi, Anna, and den Dikken, Marcel. 1999. Islands. GLOT International 4-6:3-9.

Takahashi, Daiko. 1990. Negative Polarity, Phrase Structure, and the ECP. English linguistics 7:129-146.

Takahashi, Daiko. 1993. Movement of Wh-Phrases in Japanese. Natural Language and Linguistic Theory 11:655-678.

Takano, Yuji. 2004. Coordination of Verbs and Two Types of Verbal Inflection. Linguistic Inquiry 35:168-171.

Takezawa, Koichi. 1987. A Configurational Approach to Case Marking in Japanese, University of Washington.

Tanaka, Hidekazu. 2003a. Remarks on Beck's Effect: Linearity in Syntax. Linguistic Inquiry 34:314-323.

Tanenhaus, Michael K., Boland, Julie, Garnsey, Susan M., and Carlson, Greg N. 1989. Lexical structure in parsing long-distance dependencies. Journal of Psycholinguistic Research 18:37-50.

Tateishi, Koichi. 1994. The Syntax of "Subjects": Studies in Japanese Linguistics; [3]. Stanford, Calif. Tokyo: Center for the Study of Language and Information;

Terada, Michiko. 1990. Incorporation and Argument Structure in Japanese. Doctoral. dissertation, University of Massachusetts.

Travis, Lisa. 1989. Parameters of Phrase Structure. In Alternative Conceptions of Phrase Structure, eds. Mark Baltin and Anthony S. Kroch, 263-279. Chicago: The University of Chicago Press.

Uehara, Keiko, and Bradley, Dianne C. 2002. Center-Embedding Problem and the Contribution of Nominative Case Repetition. In Sentence Processing in East Asian Languages, ed. Mineharu Nakayama, 257-287. Staonford, California: CSLI Publications.

Watanabe, Akira. 1992a. Wh-in-Situ, Subjacency and Chain Formation, MIT: MIT Occasional Papers in Linguistics.

Watanabe, Akira. 1992b. Subjacency and S-Structure Movement of Wh-in-Situ. Journal of East Asian Lingusitics 1:255-291.

Watanabe, Akira. 1994. A Cross-Linguistic Perspective on Japanese Nominative-Genitive Conversion and Its Implications for Japanese Syntax. In Current Topics in English and Japanese, ed. Masaru Nakamura, 341-369. Tokyo: Hitsuji Shobo.

Watanabe, Akira. 1994. A Cross-Linguistic Perspective on Japanese Nominative-Genitive Conversion and Its Implications for Japanese Syntax. In Current Topics in English and Japanese, ed. Masaru Nakamura, 341-369. Tokyo: Hitsuji Shobo.

Watanabe, Akira. 1996. Case Absorption and Wh-Agreement: Studies in Natural Language and Linguistic Theory; V. 37. Dordrecht; Boston: Kluwer Academic Publishers.

Watanabe, Akira. 1996. Nominative-Genitive Conversion and Agreement in Japanese: A Cross-Linguistic Perspective. Journal of East Asian linguistics 5:373-410.

Watanabe, Akira. 2002. Parametrization of Quantificational Determiners and Head-Internal Relatives, University of Tokyo: Ms.

Watanabe, Akira. 2003. Wh and Operator Constructions in Japanese. Lingua 113:519-558.

Wilder, Chris. 1999. Right Node Raising and the LCA. In The Proceedings of West Coast Conference on Formal Linguistics 18.

Yamashita, Hiroko, Stowe, Laurie, and Nakayama, Mineharu. 1993. Processing of Japanese Relative Clause Construction. Paper presented at Japanese Korean Linguistics.

Yamashita, Hiroko. 1995. Verb Argument Information Used in a Pro-drop Language: An Experimental Study in Japanese [Sept]. Journal of Psycholinguistic Research 24:333-347.