LINEARIZATION: A DERIVATIONAL APPROACH TO THE SYNTAX-PROSODY INTERFACE

by

KAYONO SHIOBARA

B.A., The University of Tokyo, 1996
M.A., The University of Tokyo, 1998

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY in THE FACULTY OF GRADUATE STUDIES (Linguistics)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 2004

© Kayono Shiobara, 2004
ABSTRACT

The major goal of this thesis is to account for a certain class of word order alternations in natural languages, in particular heavy NP shift in English and short-scrambling in Japanese. My central claim is that the properties of these alternations are best accounted for as PF interface phenomena constrained by correspondence conditions on the mapping from syntax to prosody. I develop a model of grammar in which linearization is distributed between core syntax and the prosody-syntax interface: using an incremental structure-building mechanism based on that of Phillips (1996, 2003), I provide a derivational model of the syntax-prosody mapping in which the unit of spell-out is defined by correspondence relations between syntactic objects and prosodic objects. This approach, which I refer to as the Prosodic Phase Hypothesis, provides a prosodically based account of the distinctive properties of heavy NP shift and short-scrambling, including not only clearly prosodic factors such as weight and sentence level stress, but also, indirectly, sensitivity to semantic/pragmatic factors such as focus.

The gist of the Prosodic Phase Hypothesis is that the general prosodic properties of a particular language constrain the linearization of verbal dependents in the language. In English, a clause corresponds to an Intonational Phrase in the default case, and mobile prosodic prominence, e.g. accenting and deaccenting, is used in order to encode semantic/pragmatic information. In Japanese, a syntactic phrase corresponds to an Intonational Phrase in the default case, and deletion or a word order alternation is used in order to encode semantic/pragmatic information. I attempt to reduce the difference between the two languages to the Lexical Pitch Parameter, which differentiates languages according to whether their lexically specified pitch features are distinctive (Japanese) or not (English).

The Prosodic Phase Hypothesis imposes a particular division of labor between the syntax and the prosody, and shifts the explanatory basis of certain types of linearization from pure syntax toward the PF interface. This is inspired by the guiding idea of the Minimalist Program (Chomsky 1995), which seeks to derive the conditions imposed on the language faculty from external constraints at the interfaces.
# TABLE OF CONTENTS

Abstract ii  
Table of Contents iii  
List of Abbreviations vi  
Acknowledgements viii  

## CHAPTER 1  Introduction: Modeling Linearization 1

1.1 Major Goals of the Thesis 1  
1.2 Empirical Domain of the Thesis: Heavy NP Shift in English and Short-Scrambling in Japanese 1  
1.2.1 Heavy NP Shift in English 1  
1.2.2 Short-Scrambling in Japanese 4  
1.2.3 Summary of Major Empirical Findings 5  
1.3 Theoretical Background and Assumptions 6  
1.3.1 The Model of the Grammar 6  
1.3.2 Linearization in the Prosodic Phase Hypothesis 7  
1.3.3 The Prosodic Phase Hypothesis as a Division of Labor between Grammatical Components 8  
1.4 Organization of the Thesis 9  

Appendix Linearization in Other Contemporary Grammatical Models 9

## CHAPTER II  The Prosodic Phase Hypothesis 15

2.1 The Architecture of the Language Faculty 15  
2.2 Building Syntactic Objects 17  
2.2.1 The Mechanism of Left-to-Right Structure-Building 17  
2.2.1.1 The Structure-Building Operation 18  
2.2.1.2 The Direction of Structure-Building 18  
2.2.1.3 The Course of Structure-Building 19  
2.2.2 Motivating Left-to-Right Structure-Building 23  
2.3 Building Prosodic Objects 24  
2.4 Pairing Syntactic Objects and Prosodic Objects 26  
2.4.1 Operational Definitions of Semantic/Pragmatic Notions 26  
2.4.1.1 The Feature [Focus] 26  
2.4.1.2 Broad-Focus 27  
2.4.1.3 Narrow-Focus 28  
2.4.1.4 Summary 29  
2.4.2 Broad-Focus Sentences with Default Sentence Stress 30  
2.4.2.1 General Mapping Algorithm 30  
2.4.2.2 English 34  
2.4.2.3 Japanese 37  
2.4.3 Narrow-Focus Sentences with Focal Stress 39  
2.4.3.1 Mapping Constraints for [Focus] 39  
2.4.3.2 English 42  
2.4.3.3 Japanese 44  
2.4.4 Summary: A Parallel Derivational Approach to Prosodic Structure 46  
2.5 Interface Levels 47  
2.5.1 PF Interface 47  
2.5.2 LF Interface 48
CHAPTER III  
Linearization of Verbal Dependents in English

3.1. Prosodic Properties of Linearization in the Verbal Domain

3.1.1. Overview

3.1.2. Broad-Focus Sentences

3.1.2.1. Default Sentence Stress

3.1.2.2. The Distinction between Canonical and Marked Orders

3.1.2.3. Derivations

3.1.3. Narrow-Focus Sentences

3.1.3.1. Focal Stress and the Prosodic Weight Effect

3.1.4. Residue: Unexpected Results of the Elicitation

3.1.4.1. Pronunciation Task

3.1.4.2. Judgment Task

3.2. More on the Distinction between Canonical and Marked Orders

3.2.1. The Prosodic Weight Condition and Processing Efficiency

3.2.2. The Prosodic Markedness Effect and Performance Frequency

3.3. Deriving the Properties of Marked Order

3.3.1. Arguments for the Prosodic Weight Condition

3.3.1.1. More on the Right-Edge Heaviness Effect in XP-Shift

3.3.1.1.1. CP/IP-Shift

3.3.1.1.2. Non-DP-Shift

3.3.1.1.3. DP-Shift with Preposition Stranding

3.3.1.2. Locality of XP-Shift

3.3.2. Arguments for Incremental Structure-Building in the Syntax

3.3.2.1. Syntactic Constituency Tests

3.3.2.1.1. Coordination Test

3.3.2.1.2. Ellipsis Test

3.3.2.2. Base-Generation and LF Effects

3.3.2.2.1. Interpretation of the Shifted XP

3.3.2.2.2. LF Effects

3.4. Situating the Prosodic Phase Hypothesis in Previous Analyses of XP-Shift

3.4.1. Movement Approaches

3.4.1.1. Phonological Movement

3.4.1.1.1. Stylistic Rule: Ross (1967/86) and Rochemont (1978/85)


3.4.1.2. Syntactic Movement

3.4.1.2.1. Rightward Movement: Guéron (1980) and Rochemont and Culicover (1990)

3.4.1.2.2. Leftward Movement: Larson (1989) and Kayne (1994)

3.4.1.2.3. Leftward Movement and Phonological Deletion: Takano (1998)

3.4.2. Base-Generation Approaches

3.4.2.1. Representation Theory: Williams (2003)

3.4.3. Summary

Appendix. Details of the Elicitation

CHAPTER IV  
Linearization of Verbal Dependents in Japanese

4.1. Prosodic Properties of Linearization in the Verbal Domain

4.1.1. Overview

4.1.2. Broad-Focus Sentences
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.2.1. Default Sentence Stress</td>
<td>107</td>
</tr>
<tr>
<td>4.1.2.2. Absence of the Distinction between Canonical and Marked Orders</td>
<td>110</td>
</tr>
<tr>
<td>4.1.2.3. Derivations</td>
<td>114</td>
</tr>
<tr>
<td>4.1.3. Narrow-Focus Sentences</td>
<td>116</td>
</tr>
<tr>
<td>4.1.3.1. Avoidance of Focal Stress and the Prosodic Economy Effect</td>
<td>116</td>
</tr>
<tr>
<td>4.1.3.2. Derivations</td>
<td>126</td>
</tr>
<tr>
<td>4.1.4. Residue: Unexpected Results of the Elicitation</td>
<td>129</td>
</tr>
<tr>
<td>4.2. More on the Absence of the Distinction between Canonical and Marked Orders</td>
<td>132</td>
</tr>
<tr>
<td>4.2.1. The Prosodic Economy Condition and Production Efficiency</td>
<td>132</td>
</tr>
<tr>
<td>4.2.2. Absence of the Prosodic Markedness Effect and Case Morphology</td>
<td>133</td>
</tr>
<tr>
<td>4.3. Deriving the Properties of Order-II</td>
<td>134</td>
</tr>
<tr>
<td>4.3.1. Arguments for the Prosodic Economy Condition in Japanese</td>
<td>135</td>
</tr>
<tr>
<td>4.3.1.1. More on the Absence of a Prosodic Weight Effect in XP-Scrambling</td>
<td>135</td>
</tr>
<tr>
<td>4.3.1.2. Locality of XP-Scrambling</td>
<td>136</td>
</tr>
<tr>
<td>4.3.2. Arguments for Incremental Structure-Building in the Syntax</td>
<td>138</td>
</tr>
<tr>
<td>4.3.2.1. Syntactic Constituency Tests</td>
<td>138</td>
</tr>
<tr>
<td>4.3.2.1.1. Coordination Test</td>
<td>139</td>
</tr>
<tr>
<td>4.3.2.1.2. Ellipsis Test</td>
<td>140</td>
</tr>
<tr>
<td>4.3.2.2. Base-Generation and LF Effects</td>
<td>141</td>
</tr>
<tr>
<td>4.3.2.2.1. Interpretation of the Scrambled XP</td>
<td>141</td>
</tr>
<tr>
<td>4.3.2.2.2. LF Effects</td>
<td>142</td>
</tr>
<tr>
<td>4.4. Situating the Prosodic Phase Hypothesis in Previous Analyses of XP-Scrambling</td>
<td>147</td>
</tr>
<tr>
<td>4.4.1. Movement Approaches</td>
<td>147</td>
</tr>
<tr>
<td>4.4.1.1. Phonological Movement: Ross (1967/86)</td>
<td>148</td>
</tr>
<tr>
<td>4.4.1.2. Syntactic Movement</td>
<td>148</td>
</tr>
<tr>
<td>4.4.1.2.1. Leftward Movement: Minimalist Approaches</td>
<td>148</td>
</tr>
<tr>
<td>4.4.1.2.2. LF-Lowering: Bošković and Takahashi (1998)</td>
<td>150</td>
</tr>
<tr>
<td>4.4.2. Base-Generation Approaches</td>
<td>152</td>
</tr>
<tr>
<td>4.4.3. Summary</td>
<td>153</td>
</tr>
<tr>
<td>Appendix</td>
<td>153</td>
</tr>
</tbody>
</table>

CHAPTER V Implications and Conclusions .................................................. 161

5.1. Summary of the Thesis ........................................................................ 161
| 5.1.1. Summary from a Typological Perspective                           | 161  |
| 5.1.1.1. Syntax-Prosody Mapping Algorithm                              | 161  |
| 5.1.1.2. Generalizations about Linearization of Verbal Dependents      | 166  |
| 5.1.2. Theoretical Significance of the Prosodic Phase Hypothesis      | 168  |
| 5.1.2.1. Determining Linearization                                     | 168  |
| 5.1.2.2. A Division of Labor between Syntax and Prosody                | 168  |
| 5.1.2.3. The Role of Intonational Phrases in Defining Phonological Processes | 169  |
| 5.2. Implications and Further Issues: The Lexical Pitch Parameter      | 169  |
| 5.3. Final Remarks                                                    | 171  |

References ........................................................................................................ 174
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>adjective</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative case</td>
</tr>
<tr>
<td>Adv</td>
<td>adverb</td>
</tr>
<tr>
<td>AdvP</td>
<td>adverbial phrase</td>
</tr>
<tr>
<td>AP</td>
<td>adjectival phrase</td>
</tr>
<tr>
<td>C</td>
<td>complementizer</td>
</tr>
<tr>
<td>CP</td>
<td>complementizer phrase</td>
</tr>
<tr>
<td>c-command</td>
<td>constituent command</td>
</tr>
<tr>
<td>C-order</td>
<td>canonical order</td>
</tr>
<tr>
<td>D</td>
<td>dative</td>
</tr>
<tr>
<td>Dat</td>
<td>dative case</td>
</tr>
<tr>
<td>DP</td>
<td>determiner phrase</td>
</tr>
<tr>
<td>EXNP</td>
<td>extraposition from noun phrase</td>
</tr>
<tr>
<td>Fnc</td>
<td>functional word</td>
</tr>
<tr>
<td>[F]</td>
<td>focus feature</td>
</tr>
<tr>
<td>[Focus]_{po}</td>
<td>focus feature at prosodic structure</td>
</tr>
<tr>
<td>[Focus]_{so}</td>
<td>focus feature at syntactic structure</td>
</tr>
<tr>
<td>HL</td>
<td>high-low (tone)</td>
</tr>
<tr>
<td>HNPS</td>
<td>heavy noun phrase shift</td>
</tr>
<tr>
<td>I</td>
<td>inflection</td>
</tr>
<tr>
<td>IntP</td>
<td>intonational phrase</td>
</tr>
<tr>
<td>IP</td>
<td>inflection phrase</td>
</tr>
<tr>
<td>L</td>
<td>left (edge)</td>
</tr>
<tr>
<td>LCA</td>
<td>linear correspondence axiom</td>
</tr>
<tr>
<td>Lex</td>
<td>lexical word</td>
</tr>
<tr>
<td>LF</td>
<td>logical form</td>
</tr>
<tr>
<td>M-order</td>
<td>marked order</td>
</tr>
<tr>
<td>N</td>
<td>noun</td>
</tr>
<tr>
<td>Nom</td>
<td>nominative case</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>P</td>
<td>pre/postposition</td>
</tr>
<tr>
<td>PF</td>
<td>phonetic form</td>
</tr>
<tr>
<td>PO</td>
<td>prosodic object</td>
</tr>
<tr>
<td>PO_{n}</td>
<td>prosodic object that is spelled out finally</td>
</tr>
<tr>
<td>PO_{n-1}</td>
<td>prosodic object that is spelled out before PO_{n}</td>
</tr>
<tr>
<td>PP</td>
<td>pre/postpositional phrase</td>
</tr>
<tr>
<td>Pwd</td>
<td>prosodic word</td>
</tr>
<tr>
<td>Q</td>
<td>question marker</td>
</tr>
<tr>
<td>R</td>
<td>right (edge)</td>
</tr>
<tr>
<td>SO</td>
<td>syntactic object</td>
</tr>
<tr>
<td>S-V-O</td>
<td>subject-verb-object (order)</td>
</tr>
<tr>
<td>Top</td>
<td>topic marker</td>
</tr>
<tr>
<td>v</td>
<td>small/light verb</td>
</tr>
<tr>
<td>vP</td>
<td>small verb phrase</td>
</tr>
<tr>
<td>V</td>
<td>verb</td>
</tr>
<tr>
<td>VP</td>
<td>verb phrase</td>
</tr>
<tr>
<td>X</td>
<td>syntactic head</td>
</tr>
<tr>
<td>XP</td>
<td>syntactic phrase</td>
</tr>
<tr>
<td>X'</td>
<td>intermediate syntactic category</td>
</tr>
</tbody>
</table>
The following notational conventions are used with numbered sentences and structures:

*  ill-formed in any context
%  variable acceptability depending on speakers
#  infelicitous in a given context or without an appropriate context
$\alpha > \beta$  $\alpha$ is more acceptable than $\beta$
default  A lexical item in bolding carries default sentence stress.
HEAVY  A lexical item in upper case letters carries extra prosodic prominence.
focused  A narrowly focused lexical item is underlined.
in-a-word  Lexical items that belong to the same prosodic word are connected by hyphens.
small letters  A deaccented lexical item is in small letters.

$\rightarrow$  mapping from one grammatical component to another
derivational steps

$\Rightarrow$  A syntactic object that is already spelled out is shaded.
[mapped]  An intonational phrase that is spelled out into the phonological component is contained in a square.

[syntax]  edge alignment
[prosody]
ACKNOWLEDGEMENTS

This thesis has greatly benefited from thought provoking discussion with my thesis advisor, Henry Davis. His words that came from his flexible and creative mind have always been inspiring and encouraging.

I am grateful to Rose-Marie Déchaine not only for insightful comments and discussion, but also for making me realize how important it is not to be satisfied with an easy solution, and to carefully focus on many small implications of a big idea before moving to the next. I am also grateful to Martina Wiltschko, who always got back to me with helpful comments in time. Also, the financial support of SSHRC 410-2002-1078 to Martina allowed me to present pieces of my work at conferences and concentrate on writing during the last phase of this thesis, which is gratefully acknowledged. I feel very lucky that I worked with this wonderful and well-balanced committee, who helped me survive, and even enjoy sometimes, this long and winding journey.

Beyond my committee, I would like to thank Guy Carden for greatly helping me design an elicitation, and friends and colleagues at UBC who spared time for me to act as language consultants. Thanks also go to Dirk Bury, Gunnar Hansson, Noriko Imanishi, Scott Shank and Noriko Yamane-Tanaka for substantial and helpful comments on earlier drafts, Sugunya Ruangjaroon for substantial and warm friendship, and Dianne Cook for stylistic suggestions.

Last but not least, my sincere gratitude goes to family and friends for being around.
CHAPTER 1
Introduction: Modeling Linearization

1.1 Major Goals of the Thesis

The major goal of this thesis is to account for a certain class of word order alternations in natural languages, in particular heavy NP shift in English and short-scrambling in Japanese. My central theoretical claim is that the properties of these alternations are best accounted for as PF interface phenomena constrained by correspondence conditions on the mapping from syntax to prosody. I develop a model of grammar in which linearization is distributed between core syntax and the prosody-syntax interface: using an incremental structure-building mechanism based on that of Phillips (1996, 2003), I provide a derivational model of the syntax-prosody mapping in which the unit of spell-out is defined by correspondence relations between syntactic objects and prosodic objects. This approach, which I refer to as the Prosodic Phase Hypothesis, provides a prosodically based account of the distinctive properties of heavy NP shift and short-scrambling, including not only clearly prosodic factors such as weight and sentence level stress, but also, indirectly, sensitivity to semantic/pragmatic factors such as focus.

English and Japanese VPs are ideal as a starting point for this research, because the two languages show a clear contrast in syntax and prosody. With respect to syntactic headedness, English is head-initial and alternations between dependents take place post-verbally, whereas Japanese is head-final and alternations between dependents take place pre-verbally (Fukui 1993). With respect to general prosodic properties, both languages can, in principle, use prosodic prominence to mark semantic/pragmatic focus, which enables us to look into the relation between semantic/pragmatic focus and its prosodic realization. However, English uses various degrees of prosodic prominence (e.g. accenting and deaccenting) in encoding semantic/pragmatic information in a sentence, whereas Japanese has lexically distinctive pitch accents and does not allow mobile prosodic prominence as freely as English (Abe 1955, Ladd 1996:196).

1.2 Empirical Domain of the Thesis: Heavy NP Shift in English and Short-Scrambling in Japanese

1.2.1. Heavy NP Shift in English

The process traditionally called “heavy NP shift” is schematized in (1b) and exemplified in (2) and (3).

(1) Word order alternations between dependents in English VPs
   b. Shifted order: V – PP – DP
(2) a. Meg donated [DP five hundred dollars] [PP to the city library].
   b. %Meg donated [PP to the city library] [DP five hundred dollars].
(3) a. Jack ate [DP more than ten California rolls] [PP at One More Sushi].
   b. %Jack ate [PP at One More Sushi] [DP more than ten California rolls].

The theme DP and the goal PP of the verb ‘donated’ appear in V – DP – PP order in (2a), and in V – PP – DP order in (2b). The theme DP and the place PP appear in V– DP – PP order in (3a), and in V – PP – DP order in (3b). The order in the (a) sentences is traditionally referred to as the “canonical” order and the order in the (b) sentences as the “shifted” order. Native speakers of English share the intuition that the (b) sentences in (2)-(3) are “less natural” than their counterparts in (2a) and (3a), though they are not ungrammatical. However, the exact level of acceptability varies between speakers (as is indicated by the “%” sign).

While the kind of judgments that speakers typically provide for heavy NP shift alternations are often gradient (that is, neither completely grammatical nor completely ungrammatical) and moreover are
subject to variation between speakers, they are by no means random. Rather, they are sensitive to two sets of conditioning factors: discourse context and prosodic prominence.

As far as discourse context is concerned, it turns out that we cannot evaluate judgments on shifted sentences accurately unless we situate the sentences within a larger discourse. For example, when a discourse context identifies the whole sentence as the focus of a question, speakers agree that a shifted sentence cannot serve as a felicitous answer (indicated by the "#" sign):

(4) A: What happened yesterday?
   B1: Meg donated five hundred dollars to the city library.
   B2: #Meg donated to the city library five hundred dollars.

(5) A: What happened yesterday?
   B1: Jack ate more than ten California rolls at One More Sushi.
   B2: #Jack ate at One More Sushi more than ten California rolls.

When the question asks specifically about the theme of the verb (as is indicated by underscore in the examples below), the answer can consist only of the theme DP as in (B1) in (6)-(7); the issue of word order obviously does not arise in this case since the answer is a sentence fragment rather than a full sentence. When the answer consists of a full sentence, however, the shifted order is available and can be more natural than the non-shifted order for some speakers (indicated by the "#" sign in (B2)), as is shown in the contrast between (B2) and (B3):

(6) A: How much did Meg donate to the city library?
   B1: Five hundred dollars.
   B2: #She donated five hundred dollars to them.
   B3: She donated to them five hundred dollars.

(7) A: What did Jack eat at One More Sushi?
   B1: More than ten California rolls.
   B2: #He ate more than ten California rolls there.
   B3: He ate there more than ten California rolls.

In (6)-(7), it appears that heavy NP shift is licensed, and hence the (B3) sentences can be preferred over the corresponding (B2) sentences.

However, even controlling for discourse contexts is not sufficient for evaluating the acceptability of sentences involving heavy NP shift. Another important factor is prosody. For example, the shifted sentences are felicitous answers only if the focused element is also prosodically more prominent (indicated by upper case letters in (6')-(7')) than the neighboring deaccented element (indicated by small letters):

(6') A: How much did Meg donate to the city library?
   B1: She donated FIVE HUNDRED DOLLARS to them.
   B2: She donated to them FIVE HUNDRED DOLLARS.

(7') A: What did Jack eat at One More Sushi?
   B1: He ate MORE than TEN CALIFORNIA rolls there.
   B2: He ate there MORE than TEN CALIFORNIA rolls.

Thus, heavy NP shift is licensed to the extent that the focused DP is prosodically more prominent than the neighboring element. In (6'B2) and (7'B2), this prosodic condition is satisfied because the neighboring element consists of function word(s), 'to them' in (6') and 'there' in (7'), which are inherently stressless (cf. Selkirk 1996). Descriptively, the degraded acceptability of (6'B1) and (7'B1) is known as a
"widespread conspiracy against VP-final weak pronoun" (Erteschik-Shir and Strahov 2004:315-316). If the neighboring element contains a lexical word, however, shifted sentences are degraded:

(8)  
A: How much did Meg donate to the city library?  
B1: She donated FIVE HUNDRED DOLLARS to the institution.  
B2: #She donated to the institution FIVE HUNDRED DOLLARS.

(9) A: What did Jack eat at One More Sushi?  
B1: He ate MORE than TEN CALIFORNIA rolls at the yummy place.  
B2: #He ate at the yummy place MORE than TEN CALIFORNIA rolls.

In the answer sentences (8B) and (9B), the neighboring element to the focused DP is anaphoric to a discourse antecedent, just like the deaccented pronouns in (6'B) and (7'B). (In (8B) ‘the institution’ refers to ‘the city library’ in the question sentence (8A), and in (9B) ‘the yummy place’ refers to ‘One More Sushi’ in the question sentence (9A).) However, the anaphoric elements in (8B) and (9B) contain full lexical words and hence carry lexical stress, whereas the stressless function words in (6'B) and (7'B) are deaccented. This shows that it is the prosodic properties of non-focused elements that are crucial to the acceptability of heavy NP shift, not merely their informational status as discourse anaphors. Therefore, the generalization is that the shifted order is available and can be motivated when the focused DP is prosodically heavy relative to the neighboring element. Focus in itself is not sufficient to license heavy NP shift.

It is worth emphasizing that the particular cluster of properties which characterize word order alternations such as heavy NP shift are not shared by all syntactic processes which affect linear order. For example, wh-movement and DP-movement are not subject to these effects, as shown in (10)-(11).

(10) Wh-movement
   a. *Add is cooking [what] in the kitchen?  
   b. [What] is Add cooking in the kitchen?

(11) DP-movement
   a. *(It/(j)) seems [Add] to be cooking Thai curry in the kitchen.  
   b. [Add] seems to be cooking Thai curry in the kitchen.

In (10)-(11), there is no variation in grammaticality judgments: speakers agree that the (a) sentences are ungrammatical and the (b) sentences are grammatical, irrespective of discourse context and prosodic prominence (except that (10a) is possible as an echo-question with special intonation).

The contrast between heavy NP shift and other syntactic processes with similar properties, notably extrapolation from NP on the one hand (Ross 1967, Rochemont and Culicover 1990), and “core” syntactic processes such as wh-movement and DP-movement on the other, has led many linguists to relegate heavy NP shift to a marked “periphery” (Chomsky 1981:7-9). The view that heavy NP shift is not a core syntactic phenomenon, but an optional stylistic rule in the phonological component dates at least back as far as Ross (1967), and has been explicitly argued for by Rochemont (1978) amongst others.

While it is clear that the distinction between the two types of syntactic process is a real one, I argue that the distinction should not be viewed as one of core versus periphery or obligatory grammatical process versus optional stylistic rule. I argue that heavy NP shift in English is not purely optional, but prosodically conditioned: a particular word order is determined by the prosodic pattern associated with that order, which, in turn, may be partially determined by the locus of a focus in the sentence (Shiobara 2001, 2002a). More broadly, I argue that a proper understanding of the syntax-prosody interface allows us an explanatory account of the properties of heavy NP shift and related processes without relegating them to the marked periphery. Instead, their distinctive properties fall out as interface phenomena, very much in

1 There seems to be a dialectal variation regarding to what extent VP-final pronouns are tolerated (Henry Davis, p.c. 2004). See also fn.11 in chapter 3.
the spirit of contemporary work in the Minimalist Program (cf. Chomsky 1995 and subsequent work).

1.2.2. Short-Scrambling in Japanese

If we look at a strictly head-final language such as Japanese, we find the mirror image of English heavy NP shift in a process known as “short-scrambling.” Short-scrambling is schematized in (12) and exemplified in (13B).

(12) Word order alternations between dependents in Japanese VPs
a. Non-scrambled order: PP – NP – V
b. Scrambled order: NP – PP – V

(13) A: Kinō nani-ga atta no? yesterday what-Nom happened Q
‘What happened yesterday?’

Meg-Nom city-library-to five-hundred-dollar-Acc donated
‘Meg donated five hundred dollars to the city library’


The answer in (13B1) is usually taken to reflect the canonical word order of verbal dependents in Japanese. The answer in (13B2), conversely, is assumed to be derived via a short-scrambling operation.

As with English heavy NP shift, Japanese short-scrambling appears to be optional, and the contrast between the non-scrambled and scrambled orders is judged to be much weaker in Japanese than in English. Authors since Ross (1967) have treated short-scrambling in the same way as heavy NP shift, as a stylistic rule operating in the periphery of the grammar (e.g. Williams 1984:649-650, Chomsky 1995:325, Chomsky 1998:21). I also treat heavy NP shift and short-scrambling as parallel operations, but in the sense that both of them are PF interface phenomena which, though conditioned by prosodic factors, nevertheless form part of core grammar.

It is worth mentioning that Japanese short-scrambling differs in significant ways from English heavy NP shift. This is partly due to independent differences between the two languages. First of all, the contrast in the availability of two orders is much weaker in Japanese short-scrambling than in English heavy NP shift, as was noted above regarding (13B1) versus (13B2). This difference is probably related to the lack of a Case adjacency effect between the verb and an NP in Japanese (to be elaborated in chapter 4).

Secondly, Japanese, unlike English, makes extensive use of null-NP anaphora in discourse. Therefore, any answer to a wh-question will normally involve deletion of all deaccented material; in this case, the issue of word order alternation does not even arise. This is illustrated in (14) below.

(14) A: Megu-wa ikura simin-tosyokan-ni kifusita no?
Meg-Top how.much city-library-to donated Q
‘How much did Meg donate to the city library?’

B1: Go-hyaku-doru (da).
five-hundred-dollars it.is
‘(It is) five hundred dollars’

she-Top there-to five-hundred-dollars-Acc donated

Following previous work (e.g. Saito 1985, Tada 1993, Miyagawa 1997, Takano 1998), I distinguish between long-distance, medial and short scrambling. Long-distance and medial scrambling differ from short-scrambling in moving a scrambled element outside the VP (or outside the clause, in the case of long-distance scrambling). I assume these processes are more closely related to “core” syntactic operations such as wh-movement and DP-movement than to prosodically conditioned linearization processes such as heavy NP shift, and will not discuss them further in this thesis.
‘She donated five hundred dollars to them’.

B3: #Kanojo-wa go-hyaku-doru-o soko-ni kifusita.

The natural answer to a question such as (14A) is (14B1) where all anaphoric elements are deleted. This means that full answers with both the non-scrambled order in (14B2) and the scrambled order in (14B3) are degraded relative to the sentence fragment answer in (14B1).

Once we control for null anaphora, for example when the informational content of the verbal dependents is not fully recoverable from previous discourse, there are still significant differences between short-scrambling and heavy NP shift. In particular, whereas the shifted order in the heavy NP shift alternation is associated with a particular prosodic pattern (i.e. the VP-final DP must be prosodically heavier than the preceding element), the scrambled order shows the same prosodic pattern as its non-scrambled counterpart. Furthermore, though short-scrambling resembles heavy NP shift in being sensitive to focus, it exhibits this sensitivity in a very different fashion: in Japanese, a focus appears immediately before the verb, whereas in English, a focus may appear immediately after the verb (e.g. (8B1) and (9B1)) or in the VP-final position (e.g. (6'B2) and (7'B2)). In other words, in Japanese, the focused element always occupies the position immediately left-adjacent to the verb. For example, when the wh-question specifically asks about the theme of the verb as in (15A), the non-scrambled order in (15B1), where the theme NP occupies the position immediately left-adjacent to the verb, is more natural than the scrambled order in (15B2). In both (15B1) and (15B2), neither increased prosodic prominence nor deaccenting is involved.

(15) A: Mamoru-to Mayumi-ga nanika-o motteiru keredo,
     Mamoru-and Mayumi-Nom something-Ace having and
     karera-wa nani-o siteiru no?
     they-Top what-Acc doing Q
     ‘Mamoru and Mayumi are holding something. What are they doing?’

B1: Mayumi-ga Mamoru-ni nuigurumi-o moratta mitai.
     Mayumi-Nom Mamoru-from doll-Ace received it.seems
     ‘It seems that Mayumi received a doll from Mamoru’

B2: #Mayumi-ga nuigurumi-o Mamoru-ni moratta mitai.

The contrast between (15B1) and (15B2) appears to suggest that Japanese short-scrambling is semantically/pragmatically conditioned by focus. However, I argue that Japanese short-scrambling is also prosodically conditioned in the following sense. In Japanese, the position immediately left-adjacent to the verb has been analyzed as the position for default sentence stress, which exhibits a higher pitch peak than the preceding phrasal stress (Selkirk and Tateishi 1991, Nagahara 1994, Ishihara 2000). Therefore, the generalization is that in Japanese, a focused element is placed into the “default sentence stress” position, which is associated with extra prosodic prominence independently of focus.

1.2.3. Summary of Major Empirical Findings

We have seen that heavy NP shift in English and short-scrambling in Japanese are both prosodically conditioned but in different ways. This is summarized in (16) and (17):

(16) Heavy NP shift in English
     When a DP is prosodically heavier than the neighboring element, it may appear in
     the VP-final position.

(17) Short-scrambling in Japanese
     When an XP is focused, it must appear in the default sentence stress position (i.e.
     the position immediately left-adjacent to the verb).
This embodies the claim that it is prosody that directly determines linearization of verbal dependents in English and Japanese, and focus in itself only indirectly influences the linearization processes (contra e.g. Rochemont 1986 and Rochemont and Culicover 1990 for heavy NP shift, and Miyagawa 1997 for scrambling), via increased prosodic prominence (e.g. focal stress in English) or extra prosodic prominence (e.g. default sentence stress in Japanese). The methodology that I employ in reaching this conclusion consist of systematic elicitation of acceptability judgments. In chapters 3 and 4, respectively, I examine English heavy NP shift and Japanese short-scrambling in more detail, and attempt to provide a unified account for their similarities and differences in terms of the Prosodic Phase Hypothesis (the definition of which will be given shortly in 1.3.1). I argue that independent prosodic properties of the two languages can be used to derive many of the syntactic and prosodic differences between heavy NP shift and short-scrambling, providing further support for the claim that both processes are prosodically driven.

I will show that once we control for the prosodic pattern of sentences involving heavy NP shift and short-scrambling, the gradient acceptability of these processes is much reduced. The remaining gradient effects can be accounted for as an inherent property of "prosodic prominence", which, like other forms of stress or pitch, is relational by nature, since it can only be evaluated relative to the prosody of neighboring elements.

1.3. Theoretical Background and Assumptions

This section previews the theoretical architecture that I adopt in this thesis. I begin by setting down the fundamental assumptions behind my theoretical claims.

1.3.1. The Model of the Grammar

The main proposal of this thesis is the Prosodic Phase Hypothesis in (18), which is represented in the model of the grammar illustrated in (19).

(18) The Prosodic Phase Hypothesis
A syntactic object is spelled out as a prosodic object.

(19) The model of the grammar in the Prosodic Phase Hypothesis

```
```
Lexicon
Numeration
(i) incremental incremental
SO-building & PO-building
(ii) SO-PO mapping
SO1 PO1
SO2 PO2
SO3 PO3
SO4 PO4
SOi POi
SOn POn

(iii) multiple spell-out

PF interface

LF interface
```

Many aspects of this model are familiar from standard models of grammatical organization within the tradition of Minimalist Program (cf. Chomsky 1995 and subsequent work). In particular, I assume syntactic objects (SOs) are built from a finite set of elements taken from the lexicon (a "Numeration"), and then assembled step by step in the computational component via "Merge", before being interpreted at
the two interfaces: on the semantic side, at the LF interface, and on the phonetic side, at the PF interface.

However, I adopt certain nonstandard assumptions that are critical in accounting for the phenomena which form the empirical domain of this thesis. First of all, following Phillips (1996, 2003), I adopt a left-to-right incremental structure-building algorithm that provides partial syntactic objects (SOs), which can then be mapped to prosodic objects (POs) during the course of the derivation. This means that I am adopting a strictly derivational approach to the computational component.

Second, I adopt the idea of multiple spell-out from work in the recent Minimalist framework (Chomsky 1998, 1999, 2001). However, rather than using designated syntactic categories (e.g. CP and vP) as units of spell-out (“Phases”), I will assume that spell-out is constrained by correspondence relations between syntactic and prosodic categories; any SO can be spelled out if a language provides a certain corresponding PO, in particular, an Intonational Phrase.

In the model of the grammar in (19), the syntax constructs a series of syntactic objects (SO_i, SO_2,...) each of which can potentially be realized by a corresponding prosodic object (PO_i, PO_2,...). However, not all POs can be directly realized by the phonology; only certain POs (in particular, Intonational Phrases), are large enough to correspond to units of spell-out. Thus, in the model above, PO_1, PO_2 and PO_3 cannot be realized independently because they are too small to count as Intonational Phrases, even though they are legitimate POs (for example, Prosodic Words).

1.3.2. Linearization in the Prosodic Phase Hypothesis

One of the main purposes of this thesis is to explain certain word order alternations in terms of the Prosodic Phase Hypothesis. In this section, in order to understand how the model works, I give a brief account of the heavy NP shift alternation.

Recall the basic word order alternation which characterizes heavy NP shift:

(20) A: What happened yesterday?
    B1: [\text{\textit{IntP} Meg donated five hundred dollars to the city library}].
    B2: #[\text{\textit{IntP Meg donated to the city library} \text{\textit{IntP five hundred dollars}}}].

Under the Prosodic Phase Hypothesis, both word orders are generated in the syntax, and both may correspond to legitimate prosodic objects. However, the two word orders correspond to two different sets of prosodic objects. In the non-shifted order in (20B1), the string 'Meg donated five hundred dollars to the city library' corresponds to a single Intonational Phrase (IntP); in the shifted order in (20B2), on the other hand, there are two IntPs: the string 'Meg donated to the city library' and the string 'five hundred dollars'. The reason for this (to be elaborated in chapter 3) is that a condition on the mapping from syntactic objects to prosodic objects forces the insertion of an IntP boundary at the right edge of a dependent PP in English. In this way, the particular prosodic properties associated with alternate word orders fall out from the syntax-prosody mapping constraints: while either word order is available in the syntax, the syntax-prosody mapping forces each order to be associated with a particular set of prosodic properties. The syntax-prosody mapping algorithm based on the Prosodic Phase Hypothesis is extended to cover focus effects via a focus feature which is realized prosodically in the form of increased prosodic prominence. Thus, both semantic/pragmatic and purely prosodic properties of alternate word orders are accounted for through the syntax-prosody mapping.

Notice that both non-shifted and shifted orders are directly base-generated in the syntax by the structure-building operation. In other words, there is no movement associated with heavy NP shift (or by extension with short-scrambling in Japanese, or with any other syntactic process which shows the cluster of properties associated with prosodically determined word order alternations). This is schematized in (21) below. Under the Prosodic Phase Hypothesis, the linear order between two dependents, DP and PP

---

7 As a consequence, I need to relax a theta-locality condition so that the semantic association between the verb and a displaced dependent (i.e. PP in the non-shifted order and DP in the shifted order) can be established. See 3.3.2.2 for further discussion.
for example, is fixed in the syntax, but not specified for a particular language. The information about linear ordering is transparently mapped from syntax to phonology, and it is conditions on the syntax-prosody mapping that determine possible and impossible linear orders.

(21) Linearization in the Prosodic Phase Hypothesis

a. Derivation of syntactic objects

```
  DP   PP   PP   DP
  \   \   \   \  \\
  A   A   A   A  
```

b. Linearization in the phonology

```
  Syntax
  \   \   \   \  \\
  DP   PP   PP   DP
  \   \   \   \  \\
  A   A   A   A  

  Phonology
  Language A:
  DP – PP, *PP – DP
  Language B:
  *DP – PP, PP – DP
  Language C:
  DP – PP, PP – DP
  Language D:
  *DP – PP, *PP – DP
```

The base-generation approach to heavy NP shift and short-scrambling adopted here is a welcome result for heavy NP shift and its close relatives, since it has long been acknowledged that as syntactic movement rules, these processes are extremely problematic. For example, in contrast to “core” syntactic movement rules such as wh-movement and DP-movement (see (10)-(11)), they involve a rightward dependency, are strictly bounded (Ross 1967, Akmajian 1975, Baltin 1981, 1983), and, as we have seen, are sensitive to prosodic and semantic/pragmatic factors.

It is important to note that the approach to linearization taken here is not meant to account for all linearization phenomena in natural languages. For example, “core” syntactic processes that affect linearization (e.g. wh-movement and DP-movement) are treated as syntactic movement rules (in the sense of Chomsky 1995 and other work within the Minimalist framework). The syntax-prosody mapping constraints do not determine the linear order between a head and a dependent either. Instead, I assume that the structure-building mechanism in the syntax directly incorporates a head-directionality parameter which specifies the linear order between a head and a dependent. In other words, the linearization processes the Prosodic Phase Hypothesis governs are limited to local alternations, such as permutation of dependents within the same domain (e.g. VP).

Moreover, linearization of verbal dependents involving heavy NP shift or short-scrambling is also subject to purely syntactic effects. For example, I assume that a Case adjacency requirement on DP/NP (Stowell 1981) specifies that V – DP – PP order is the canonical and “non-shifted” order in English, whereas a Case adjacency effect is much weaker in Japanese due to its overt Case morphology.

1.3.3. The Prosodic Phase Hypothesis as a Division of Labor between Grammatical Components
The major claim of the Prosodic Phase Hypothesis is that the different behavior of certain word order alternations in English and Japanese is due to difference in the general prosodic properties of the two languages. English makes use of mobile prosodic prominence (e.g. accenting and deaccenting) in order to encode particular focus interpretations whereas Japanese does not. Instead, Japanese deletes anaphoric elements, or moves a focused element into the default sentence stress position, i.e. the position immediately left-adjacent to the verb.

The Prosodic Phase Hypothesis is inspired by the guiding idea of the Minimalist Program (Chomsky 1995), which seeks to derive the conditions imposed on the language faculty from external constraints on interfaces. The Prosodic Phase Hypothesis proposes a particular division of labor between syntax and prosody, and provides a prosodically based account of interface properties associated with English heavy NP shift and Japanese short-scrambling. My examination will center on the prosodic properties of these processes, because the prosody is the direct determinant of them, and prosodic prominence is always related to semantic prominence (i.e. focus), but not vice versa.

1.4. Organization of the Thesis

The thesis is organized as follows. Chapter 2 presents the Prosodic Phase Hypothesis as a general syntax-prosody mapping algorithm, which constrains linearization of verbal dependents. I outline the theoretical claims the Prosodic Phase Hypothesis makes, and argue for incremental structure-building in the syntax and a derivational mapping from the syntax to the prosody. I subsequently show how the Prosodic Phase Hypothesis explains the general prosodic properties associated with English and Japanese sentences: in particular, the differences between the two languages regarding the size of an Intonational Phrase, and how increased prosodic prominence changes the intonation pattern of a sentence. More specifically, I show how syntax-prosody mapping constraints formalize the following differences: (i) the size of an Intonational Phrase is bigger in English than in Japanese; and (ii) increased prosodic prominence is aligned at the right edge in English and at the left edge in Japanese.

Chapter 3 and chapter 4, respectively, examine prosodic properties associated with word order alternations in English and Japanese, in particular heavy NP shift in English and short-scrambling in Japanese. The key data dealt with in these chapters are from the results of a systematic elicitation I designed and implemented in order to control for prosody and focus. In the elicitation of acceptability judgments, stimuli sentences are constructed in a question-answer format in order to identify the focus in answer sentences, and they are subsequently presented as sound files so that the subjects judge the acceptability of sentences with particular prosodic patterns. In these chapters, I discuss the major findings of the elicitation: (i) in English heavy NP shift, the shifted order is licensed to the extent that the rightmost DP is prosodically heavy enough to be an Intonational Phrase; and (ii) in Japanese short-scrambling, the order where the focused element sits in the default sentence stress position is chosen.

Based on these findings, I illustrate the derivations of non-shifted and shifted sentences in English and non-scrambled and scrambled sentences in Japanese, and formalize the prosodic conditions on heavy NP shift and short-scrambling as the “Prosodic Weight Condition” and the “Prosodic Economy Condition”, respectively. Then I show how the prosodic and syntactic properties associated with these alternations support the claims of the Prosodic Phase Hypothesis, including the incremental approach to structure-building in the syntax and the strictly derivational approach to the syntax-prosody mapping.

Chapter 5 recapitulates the empirical and conceptual significance of the Prosodic Phase Hypothesis and considers its implications from a broader typological perspective.

Appendix. Linearization in Other Contemporary Grammatical Models

A way to treat linearization is a necessary property of any theory of the language faculty that assumes hierarchical representations. If we assume that linearization, which “flattens” tree structure, is ultimately a PF requirement imposed by the sensori-motor system, there are two possibilities as to where in the grammar the linear precedence relation is determined: syntax or phonology. The Prosodic Phase
Hypothesis claims that certain instances of linearization, in particular local alternations between dependents, are determined by corresponding conditions on the syntax-prosody mapping, and provides a theory of the grammar that accommodates prosodic conditions on linearization. Under this view, linearization is distributed between the syntax proper and the syntax-prosody interface. In this appendix, I compare the Prosodic Phase Hypothesis to other contemporary models of linearization, with regards to general linearization processes.

A1. Syntax-Based Models

A1.1. Language-Particular Phrase Structure Rules

In traditional syntax-based models of linearization, the linear precedence relation is directly determined by phrase structure rules, and then transparently mapped to the phonology. In classical phrase structure theories (Chomsky 1965, 1975), a particular language has a set of phrase structure rules which specify immediate dominance relations and linear precedence relations in the form of rewrite rules. For example, the phrase structure rule of Language A in (1a) states that 'VP immediately dominates V and NP (and optionally PP), and V precedes NP'.

(1) Language-particular phrase structure rules and linearization

a. Derivation of syntactic objects

| Phrase structure rules of Language A: | VP → V - NP (- PP) |
| Phrase structure rules of Language B: | VP → (PP -) NP - V |

b. Linearization in the syntax

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Phonology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language A:</td>
<td>VP</td>
</tr>
<tr>
<td>V</td>
<td>NP (PP)</td>
</tr>
<tr>
<td>Language B:</td>
<td>VP</td>
</tr>
<tr>
<td>(PP)</td>
<td>NP</td>
</tr>
</tbody>
</table>

In this model, all instances of linearization, not only the order between a head and a dependent but also the order between dependents, are determined in the syntax.

A1.2. X-Bar Theory

The X-bar theory of phrase structure is another syntax-based model of linearization (Jackendoff 1977, Chomsky 1981, 1986b). However, the X-bar theory differs from classical phrase structure theories in that it separates hierarchical relations from linear order (cf. also Gazdar et al. 1985). The X-bar schema specifies that "XP immediately dominates X' and specifier" and "X' immediately dominates a head X, and complement", but the linear precedence relation between sisters is unspecified. This is shown in (2a) below.
A syntactic parameter (e.g. head-parameter, Case-directionality parameter) determines the linear order between the head X and the complement YP in the syntax: if Language A specifies the value of the head-parameter as initial, the syntactic object X' that consists of X and YP with X – YP order is derived. And if Language B specifies the value of the head-parameter as final, the syntactic object X' with YP – X order is derived. The linear precedence relation is transparently mapped to the phonology. This is shown in (2b).

(2) The X-bar theory and linearization

a. Derivation of syntactic objects

Phrase structure rules:

\[ XP \rightarrow X', ZP \text{ (specifier)} \]

\[ X' \rightarrow X, YP \text{ (complement)} \]

b. Linearization in the syntax

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Phonology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language A:</td>
<td></td>
</tr>
</tbody>
</table>
| \[ X' \]
| \[ X \]
| \[ YP \]
| \[ \triangle \]
| \[ \rightarrow \]
| \[ X - YP \] |

Language B:

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Phonology</th>
</tr>
</thead>
</table>
| \[ X' \]
| \[ YP \]
| \[ X \]
| \[ \triangle \]
| \[ \rightarrow \]
| \[ YP - X \] |

The X-bar theory also distinguishes between complements and adjuncts in terms of their hierarchical relations: a complement is a sister to a head X and immediately dominated by X', and an adjunct is a sister to X' and immediately dominated by X'. Note that the X-bar theory, as originally formulated, does not specify any hierarchical relation between complements of the same predicate. However, subsequent developments, notably the adoption of uniform binary branching (Kayne 1984, 1994) and the single complement hypothesis, have established an asymmetry between complements as well as adjuncts, as we will see in the next section.

A1.3. A Single Complement Hypothesis

The first step towards a theory of phrase structure where all complements and adjuncts are hierarchically distinguished came with the adoption of Baker’s (1988) Uniformity of Theta Assignment Hypothesis (UTAH). The UTAH is a grammatical principle that includes the idea that the hierarchical relation among theta-roles is rigidly mapped to a hierarchical relation in syntactic (deep) structures.

(3) The Uniformity of Theta Assignment Hypothesis (UTAH) (Baker 1988:46)

Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure.
Adopting the UTAH, Larson (1988) develops an account of multiple complements in English which involves a VP-shell structure for verb phrases as in (4):

(4) a. Thematic Hierarchy (Larson 1988:382, due to Carrier-Duncan 1985)
   Agent > Theme > Goal > Obliques (manner, location, time,...)

b. \[
\begin{array}{c}
\text{vP} \\
\text{ZP} \\
\triangle \text{v} \\
\text{YP}
\end{array}
\]
   \[
\begin{array}{c}
\text{VP} \\
\triangle \text{V} \\
\text{WP}
\end{array}
\]

Following Carrier-Duncan (1985), Larson assumes that Theme is higher than Goal in the Thematic Hierarchy (but see e.g. Jackendoff 1972, Bresnan and Kanerva 1989, and Grimshaw 1990 for different versions of the Thematic Hierarchy). Therefore, the Theme YP sits in a higher position than the Goal WP in D-structure in (4b).

What about the linear precedence relation between the two complements? Larson (1988) proposes a “single complement hypothesis” for X-bar structure (p.381), which entails the same number of verbal heads as complements to V: for example, in (4b), the head \(v\) takes the VP complement and the head \(V\) takes the WP complement. The Theme argument (YP) occupies the specifier position of the VP, and the Goal argument (WP) occupies the complement position of the V. Therefore, the linear precedence relation between YP and WP is determined by the linear order between the specifier (YP) and the head \(V\), and the order between the head \(V\) and the complement (WP). Note that unlike standard X-bar theory, Larson’s theory does not distinguish between complements and adjuncts: adjuncts are generated as complements in lower VP shells.

A1.4. Antisymmetry Theory

The most radical syntax-based model of linearization is Kayne’s (1994) Antisymmetry theory of phrase structure. It differs from X-bar theory in how linear order is determined. Kayne proposes the Linear Correspondence Axiom (LCA), which derives specifier-head-complement order (or complement-head-specifier) as the universal order which is base-generated in the syntax. This is shown in (5a) below.

Within the Antisymmetry theory, only \(X - YP\) order is base-generated, and any \(YP - X\) order must be derived via syntactic movement. The linear precedence relation is transparently mapped from the syntax to the phonology. This is shown in (5b).

(5) The Antisymmetry theory and linearization

a. Derivation of syntactic objects

\[
\begin{array}{c}
\text{(specifier)} \\
\text{ZP (head)} \\
\triangle X \\
\text{(complement)} \\
\text{YP}
\end{array}
\]
b. Linearization in the syntax

Syntax

Language A:

\[
\begin{array}{c}
X \\
\triangle
\end{array} \quad \begin{array}{c}
YP \\
\rightarrow
\end{array}
\]

Language B:

\[
\begin{array}{c}
YP \\
\triangle
\end{array} \quad \begin{array}{c}
X \\
\rightarrow
\end{array} \quad \begin{array}{c}
t_{YP} \\
\rightarrow
\end{array}
\]

Regarding the linear precedence relation between verbal dependents, the LCA forces Larson's single complement hypothesis, and hence Larsonian shell structures are derived (Kayne 1994:69).

A2. PF-Based Models

Let us now turn to PF-based models of linearization. In a PF-based model, the linear precedence relation is unspecified in the syntax and determined in the phonology. That is to say, the syntax is only in charge of the hierarchical relation. The Minimalist Program represents a PF-based model of linearization (Chomsky 1995).

In the Minimalist Program, syntactic objects are built from bottom up by "Merge" of two syntactic objects, e.g. \(\alpha\) and \(\beta\). The structures are "bare" and there is no node labeling (Chomsky 1994). The linear order between specifier, head and complement is determined in the phonological component. In particular, Chomsky (1995) adopts Kayne's (1994) syntactically motivated LCA and restates it in phonological terms (Chomsky 1995:334-340).

(6) The Minimalist Program and linearization

a. Derivation of syntactic objects

\[
\begin{array}{c}
\{\alpha, \beta\} \\
\alpha \\
\beta
\end{array}
\]

b. Linearization in the phonology

Syntax

\[
\begin{array}{c}
\{\alpha, \beta\} \\
\alpha \\
\beta
\end{array} \quad \rightarrow
\]

Phonology

Language A:

\(\alpha - \beta\)

Language B:

\(\beta - \alpha\)
Since the Minimalist Program adopts Larsonian shell structures (Chomsky 1995:315), presumably the linear order between complements is also determined by the LCA in the phonological component. In the Minimalist Program, all instances of linearization are determined in the phonology.

A3. A Combination of Syntax-Based and PF-Based Models

Phillips (1996, 2003) proposes an analysis that combines the syntax-based and PF-based models of linearization. His proposal is situated in a larger research framework which he calls the “Parser Is Grammar” view of the language faculty. The Parser Is Grammar view does not employ the traditional division of labor between the grammar and the parser, and there is no independent component comprised solely of the parser. In the Parser Is Grammar view, phrase structures are built from left to right in the order in which lexical items are pronounced. For example, if there is a syntactic object $\alpha$, and $\beta$ is merged into $\alpha$ and then $\gamma$ is merged into $\beta$, the order $\alpha - \beta - \gamma$ is derived. This is illustrated in the Language A pattern in (7). If $\gamma$ is merged into $\alpha$ first and then $\beta$ is merged into $\gamma$, the order $\alpha - \gamma - \beta$ is derived. This is illustrated in the Language B pattern in (7).

(7) The Parser Is Grammar view and linearization

\begin{center}
\textbf{Syntax/PF}
\end{center}

\begin{center}
\begin{align*}
\text{Language A:} & \quad \Rightarrow \\
& \quad \alpha \quad \beta \\
& \quad (\text{Linear order is fixed})
\end{align*}
\end{center}

\begin{center}
\begin{align*}
\text{Language B:} & \quad \Rightarrow \\
& \quad \alpha \quad \gamma \\
& \quad (\text{Linear order is fixed})
\end{align*}
\end{center}

The assumption behind the Parser Is Grammar view is that the linear precedence relation is part of grammatical knowledge that is used in structure-building. In many ways, this is a return to a more traditional view of structure-building where both immediate dominance and linear precedence relations are specified by the same rules.

A4. Summary: The Prosodic Phase Hypothesis in the Context of Theories of Linearization

The models of linearization reviewed in this appendix all attribute linearization to a single component of the grammar: either syntax, phonology, or a unified component of syntax/phonology. This contrasts with the Prosodic Phase Hypothesis in this thesis, which divides the labor of linearization between syntax and phonology. In particular, while the linear order between a head and a complement is determined in the syntax, the linear order between verbal dependents is determined by correspondence conditions on the syntax-prosodic mapping, as well as by purely syntactic mechanisms such as a Case adjacency condition.
CHAPTER II
The Prosodic Phase Hypothesis

2.1. The Architecture of the Language Faculty

The goals of this thesis are two-fold. Empirically, the major goal is to shed light on the interface (i.e. prosodic and semantic/pragmatic) properties associated with linearization processes. Theoretically, this thesis considers how linearization in general is influenced and shaped by the requirements of interface conditions. I propose that a certain type of linearization is constrained by the syntax-prosody mapping condition which requires a syntactic object to be mapped to a certain type of prosodic object and then spelled out into the phonological component. I call this the “Prosodic Phase Hypothesis”.

(1) The Prosodic Phase Hypothesis
A syntactic object is spelled out as a prosodic object.

The theoretical background assumed in this thesis is the theory of Generative Grammar, in particular, the recent versions of the Minimalist Program (Chomsky 1995, 1998, 1999, 2001). First, this thesis follows the traditional assumption that the language faculty, a domain in the human mind/brain dedicated to language, has at least two distinct systems: “competence” that stores the knowledge of language, and “performance” that accesses this knowledge and puts it to use (Chomsky 1965, 1986a). The MP shares with its predecessors the working assumption that competence, i.e. so called “grammar” or the “language cognitive system” in Minimalist terms (Chomsky 1995:2), and performance are two distinct systems. The guiding idea of the program is to seek to derive the conditions imposed on the language faculty from “external constraints” such as general conditions on cognitive systems and general considerations of conceptual naturalness (Chomsky 1995:1). The Minimalist Program assumes that the grammar interacts with two performance systems, the “sensori-motor system” and the “conceptual-intentional system”, at two interface levels, “Phonetic Form” (PF) and “Logical Form” (LF). PF and LF do not interact with each other directly, and external constraints are captured by “interface conditions” that apply at PF or LF.

Another feature of the Minimalist Program adopted here is a strictly derivational model of the computational component (Chomsky 1995:224). The grammar consists of the lexicon and the computational systems. The computation selects relevant lexical items from the lexicon and constructs a “Numeration”, from which the syntactic derivation proceeds. I argue that the computational component includes parallel building of syntactic objects (SOs) and prosodic objects (POs), which are related to each other by the syntax-prosody mapping and spelled out into the phonological component:

(2) The parallel building of syntactic structure and prosodic structure
POs show systematic relations to SOs, though they have been argued not to be isomorphic to SOs in prosodic structure theories (e.g. Selkirk 1984, 1986, Nespor and Vogel 1986, Inkelas 1989, Zec and Inkelas 1990). The set of rules or constraints relating SOs and POs are referred to collectively as the syntax-prosody mapping. In the model in (2), the syntax constructs a series of SOs (SŌ, SO₂, ...), each of which can potentially be realized by a corresponding PO (PŌ, PO₂, ...). When the computation exhausts the lexical items in the Numeration, the final SO (which is referred to as “SŌ” in (2)) is built, mapped to a PO (“PŌ” in (2)), and spelled out into the phonological component.

Unlike the single spell-out in the model in (2), recent work in the Minimalist framework proposes that the spell-out of an SO can apply more than once (Chomsky 1998, 1999, 2001). This is called “multiple spell-out”:

(3) Multiple spell-out

In the model in (3), any POs (PO₁, PO₂, ... POₙ) can be spelled out and directly realized by the phonology. More recently, Chomsky (1998, 1999, 2001) proposes that multiple spell-out occurs in the cycle called “Phase”. The computation forms SOs by the bottom-up structure-building operation called “Merge”, where SOs are cyclically spelled-out into the phonological component Phase-by-Phase. Chomsky takes vP and CP to be Phases. On the LF side, they are propositional: “verbal phrases with full argument structure and CP with force indicators” (Chomsky 1999:9); on the PF side, they have “a degree of phonetic independence” (ibid.).

The size of a Phase is not pre-theoretically determined, and should ultimately be an empirical issue. Epstein and Sheely (2002) point out that the syntactic definition of a Phase is problematic in that it requires a computational look-ahead: “how can we know that they are relatively independent at the interface if Spell-Out applies before the interface is reached, and without access to interface properties?” (p.78); “why should PF care about the propositional content of what is spelled out?” (ibid.) Given that a Phase is the unit of spell-out into the phonological component, looking into PF interface properties of Phases may be useful in illuminating the properties of the grammar. Departing from the syntactic definition of a Phase, I propose that a Phase is SO-PO pairing, which derivationally takes place in the computational component. The Prosodic Phase Hypothesis in (1), as a syntax-prosody mapping algorithm, formulates this, and the model of the language faculty illustrated in (4) represents the Prosodic Phase Hypothesis:
The model of the language faculty in the Prosodic Phase Hypothesis

In this model, not all POs can be directly realized by the phonology. I propose that only Intonational Phrases (IntPs) are large enough to correspond to units of spell-out: in (4), PO₁, PO₂ and PO₄ cannot be realized independently because they are too small to count as IntPs, even though they are legitimate POs (for example, prosodic words). The Prosodic Phase Hypothesis in (1), as a general syntax-prosody mapping algorithm, constrains certain linearization processes by imposing a condition on the size of the POs mapped from SOs. In particular, an SO-PO pair which is phonologically identified as the domain of an intonation contour (i.e. PO = IntP) acts as a unit of spell-out. This way, we solve the look-ahead problem of the syntactic definition of Phases, because the PF interface condition that POs must be IntPs directly determines the size of an SO-PO pair that is spelled out into the phonological component.

The model of the language faculty in (4) embodies the following claims of the Prosodic Phase Hypothesis: (i) in the computational component, SOs and POs are built in parallel from left to right in the order in which lexical items are pronounced, (ii) SOs are mapped to POs to establish SO-PO pairs, and (iii) SO-PO pairs are spelled out into the phonological component iteratively when the POs are as large as IntPs. In the rest of this chapter, I show how SOs are built (2.2), how POs are built (2.3), and how the mapping from the syntax to the prosody is constrained (2.4). Lastly, I discuss what the representations at interface levels of PF and LF look like under the Prosodic Phase Hypothesis (2.5).

2.2. Building Syntactic Objects

This section illustrates the mechanism of incremental structure-building in the syntax, and shows how the Prosodic Phase Hypothesis motivates incremental structure-building.

2.2.1. The Mechanism of Left-to-Right Structure-Building

I assume that syntactic structure-building is subject to the conditions in (5):

(5) Conditions on syntactic structure-building
i. Operation: Syntactic Merge
ii. Elementary unit of the syntactic object (i.e. terminal element): syntactic word
iii. Direction of Merge: left-to-right
iv. Branching: binary
v. Locus of Merge: any accessible target
vi. Other available information: e.g. head-directionality, selectional restrictions

The following sections discuss the details of these conditions.

2.2.1.1. The Structure-Building Operation

First, it is one of the minimum requirements that the computational component should be able to construct syntactic structures from lexical items. The Numeration consists of syntactic words, and the operation "Syntactic Merge", defined as in (6), builds syntactic structure.

(6) Syntactic Merge
Merge takes a pair of syntactic objects (SO$_i$, SO$_j$) and replaces them by a new combined syntactic object (SO$_k$). (Chomsky 1995:226)

I assume that Merge takes a syntactic word and merges it into another syntactic object (SO) (= conditions in (5i)-(5ii)). Simplifying somewhat, I assume that syntactic words are lexical and functional heads including inflectional morphemes but not derivational morphemes. (Some functional heads such as Infl are omitted in the following discussion when irrelevant.)

2.2.1.2. The Direction of Structure-Building

Spell-out of a pair of a syntactic object (SO) and a prosodic object (PO) established by the derivational SO-PO mapping, rather than spell-out of CPs or vPs, may be implemented when syntactic structure is built from left to right. First, let us compare bottom-up and left-to-right structure-building mechanisms, and then see why left-to-right structure-building allows an implementation of the Prosodic Phase Hypothesis.

An assumption of the Minimalist Program is that syntactic structure is created from the bottom up, with an element lower in the structure being introduced before an element higher in the structure. For example, the verb is merged with its internal argument first, and then the resulting constituent is merged with the external argument, creating a sentence:

(7) Bottom-up structure-building
Numeration: {Betty, saw, Jack}

a. \[
\begin{array}{cc}
\text{saw} & \text{Jack} \\
\end{array}
\]  \Rightarrow  \begin{array}{cc}
\text{Betty} & \text{\quad saw \quad Jack} \\
\end{array}
\]

(NB: Linear order is irrelevant.)

However, the way elements are merged is not pre-theoretically determined and therefore is subject to debate. Phillips (1996, 2003) has proposed that syntactic structure should be built from left to right. In his theory, the external argument is merged with the verb first yielding (8a), and then the internal argument is merged with the verb via "Merge-Right", destroying the constituent of S(ubject)-V in (8a) and creating the English S-V-O sentence as in (8b):

(8) Left-to-right structure-building
Merge-Right is a specific structure-building operation and based on the general hypothesis that structures are built incrementally:

(9) a. Incrementality Hypothesis (Phillips 2003:37)
Sentence structures are built incrementally from left to right, i.e. in the order in which terminal elements are pronounced.

b. Merge-Right (Phillips 1996:18)
New items must be introduced at the right edge of a structure.

Phillips' proposal is grounded in the Parser Is Grammar view of the architecture of the language faculty, which does not employ the traditional division of labor between the grammar and performance systems (or more specifically, the “Parser”). Under the Parser Is Grammar view, there is no independent component comprised solely of the parser, and the grammar and the parser are the same system. Therefore, the term “structure” in (9) refers to syntactic and performance structures at the same time.

Although this thesis adopts Phillips' incremental structure-building mechanism as one way of implementing the derivational SO-PO mapping (as we will see shortly), I assume a model of the language faculty with a conventional division between the grammar and performance systems (see (4)). Moreover, I distinguish between syntactic structure and prosodic structure in the grammar. This is because the Prosodic Phase Hypothesis in (1) subsumes the existence of SOs that are defined purely in syntactic terms and POs that are defined purely in prosodic terms. However, not to commit to the Parser Is Grammar view does not imply that left-to-right structure-building is an accidental property to be stipulated in the grammar. On the contrary, I assume that ultimately left-to-right structure-building reflects a general linear nature of temporality, or “the asymmetry of time” (Kayne 1994:36-38): time flows only from left to right (= condition in (5iii)).

2.2.1.3. The Course of Structure-Building

Let us examine the course of syntactic derivations. First, it is not obvious in Phillip's proposal in (9) how the starting point of the derivation is defined. The usual definition of phrase structures demands that “a single designated symbol be admissible as the start symbol in a derivation” (Komai and Pullum 1990:32). The label of the start symbol can be exocentric S (which stands for “sentence” or just “start”, e.g. Chomsky 1970), or endocentric $V^n$ (e.g. n = 3 in Jackendoff 1977), IP, or CP (Chomsky 1986b), depending on the theory of phrase structure. More recently, Chomsky (1994) proposes Bare Phrase Structure theory, where syntactic nodes are not annotated with labels. The operation Syntactic Merge in (6) defines the starting point of the derivation as when a pair of syntactic words are selected from the Numeration and merged. I annotate the topmost node as IP just for expository purposes, but node labels are not important and will be omitted in the structures below when irrelevant.

The endpoint of structure-building is when all the syntactic words are selected from the Numeration and merged into one structure dominated by a single node. This is generally assumed, and dictated as the Single Root Condition which states that “in every well-formed constituent structure tree, there is exactly one node that dominates every node” (Wall 1972:146). I assume this condition.

In the definition of Merge-Right in (9b), the terms “introduced” and “the right edge” are
undefined, and it is the economy condition termed "Branch-Right" that determines where a syntactic word is merged.

(10) Branch-Right (Phillips 1996:19)
Metric: select the attachment that uses the shortest path(s) from the last item in the input to the current input item.
Reference set: all attachments of a new item that are compatible with a given interpretation.

The principle Branch-Right favors the building of maximally right-branching structures, all other things being equal. This idea of local attachment preference is independently argued for and closely related with the same preference in parsing (e.g. "Right Association" in Kimball 1973, "Late Closure" in Frazier 1978, and "Recency" in Gibson 1991).

For example, when a simple sentence such as 'Betty saw Jack' is built, "the right edge" should be the rightmost terminal element and "introduce" should be Chomsky-adjoin, because a syntactic word adjoins to the rightmost terminal element and becomes a sister of the element, as we saw in (8). Therefore, if we were to continue the derivation of the sentence and merge the adverb 'yesterday', the adverb targets the rightmost terminal noun 'Jack', and becomes the sister of the noun as in (11c).

(11) a. Betty saw ⇒ Betty saw Jack ⇒ Betty saw Jack yesterday
(NB: Linear order is relevant.)

What seems to be implicitly excluded in this case is ternary branching or flat structure as in (12a), and targeting the node that immediately dominates the rightmost terminal element (which I call the "rightmost node") and adjoining the element above the rightmost terminal element as in (12b):

(12) a. * Betty saw Jack * yesterday
(b. Betty saw Jack yesterday)

I assume binary branching in syntactic structures (= condition in (5iv)) and consequently exclude flat structures like (12a) (cf. Kayne 1984, 1994).

Concerning the locus of Merge-Right, Merge-Right that targets only the rightmost terminal element yields maximally right-branching structures and excludes structures like (12b). However, Merge-Right in (9b) may target the rightmost node as in (12b) for certain cases involving the internal structure of phrases. Consider the following examples:

(13) Internal structure of phrases

a. Merge-Right to the rightmost node.
In (13a), the verb ‘saw’ adjoins to the node that immediately dominates the rightmost terminal noun ‘cat’, not to the noun itself, and becomes the sister of the node. Then, the determiner ‘the’ adjoins to the rightmost verb ‘saw’, and then the noun ‘rat’ adjoins to the rightmost determiner ‘the’. In the resulting structure, the subject ‘the cat’ is a syntactic phrase with its own internal structure. On the other hand, if the verb were to adjoin to the terminal noun ‘cat’, the subject ‘the cat’ does not form a syntactic constituent in the resulting structure as is seen in (13b). At the step 2, for example, the principle Branch-Right in (10) chooses the structure in (13b) over (13a), because the path from the verb ‘saw’ to the preceding lexical item ‘cat’ is shorter in (13b) (two paths, as indicated by dotted arrows) than in (13a) (three paths).

In order to distinguish between heads and phrases in syntactic structure as in (13a), I restrict the application of Merge in terms of the lexical property of the rightmost terminal element: when the terminal is a head that lexically selects a dependent, Merge applies to the rightmost terminal element; else, Merge applies to the rightmost node.

(14) Conditions on Syntactic Merge

Merge
a. to a rightmost terminal element if the terminal is the head that lexically selects a dependent, else
b. to the rightmost node.
(Rightmost node: the node that immediately dominates the rightmost terminal element)

Having restricted the application of Merge-right as in (14), I abandon Branch-Right.

Going back to the sentence ‘the cat saw the rat’, the conditions in (14) force the syntactic derivation of the sentence to proceed in four steps as in (13a), not (13b). First, a pair of syntactic objects, ‘the’ and ‘cat’ are selected from the Numeration and merged. Since the noun ‘cat’ does not select a following dependent in this case, the verb ‘saw’ targets the rightmost node and merges to the node in step

---

1 Later to my question of how the internal structure of NPs is created, Phillips says that “the answer to your question *must* be that NP-internal structure is built up incrementally, in the same way as NP-external structure. This is because NPs can be arbitrarily complex, and can contain all of the sentential syntax that can appear outside of an NP. However, I have not yet shown exactly how this can work” (p.c. April 2003).

2 The condition in (14a) could be too restrictive for prenominal modifiers such as possessor phrases (PossP), which require a following head noun, yet may have its own internal structure under some analyses (e.g. [[PossP [John] ['s] cat]]).
2. Next, since the verb ‘saw’ lexically selects an internal argument, the determiner ‘the’ is merged to the rightmost terminal ‘saw’ in step 3. Finally the noun ‘rat’ is merged to the rightmost terminal ‘the’, because the determiner requires a following noun. In the final structure in (13a), the rightmost terminal element is the noun ‘rat’, which does not select a following dependent. Therefore, if we were to merge an adverb ‘briefly’ after the noun, the adverb adjoins to the node that immediately dominates the rightmost terminal noun ‘rat’, not to the noun itself:

(13a) step 5.

\[
\text{the} \quad \text{cat} \quad \text{saw} \quad \text{briefly} \\
\text{the} \quad \text{rat}
\]

The conditions on Merge in (14) impose a specific way of implementing left-to-right structure-building in an incremental fashion, in accordance with the general temporality condition. Because the direction of Merge is from left to right, Merge always targets the right edge of the structure. Therefore, the locus of Merge can be specified as “any accessible target” (= condition in (5v)).

As we have seen, syntactic structure-building is guided by the lexical selectional information of heads (= condition in (5vi)). This is independently argued by certain analyses of left-to-right sentence processing mechanisms, where it is considered lexical information guides ongoing sentence processing (Berwick and Weinberg 1984:152). Ungrammatical strings such as ‘*cat the the rat saw’ are not generated due to the direction of lexical selection by a head. For example, the determiner ‘the’ in English requires a following, not preceding, head noun, whereby the string ‘*cat the’ is impossible. Likewise, the head verb ‘saw’ in English requires a following, not preceding, theme argument, and the string ‘the rat saw’ (in the sense that ‘the rat was seen’) is not possible. Notice here that the head-parameter is incorporated as syntactic information regarding the direction of lexical selection by a head, e.g. rightward in English, and this kind of information is accessible to syntactic structure-building.3

I propose that the conditions on Merge in (14) are universal and hold in head-final languages such as Japanese too.

(15) Neko ga nezumi o mita.

\[
\begin{array}{llll}
\text{cat} & \text{Nom} & \text{rat} & \text{Acc} \\
\text{saw}
\end{array}
\]

‘The cat saw the rat’

Numeration: \{neko, ga, nezumi, o, mita\}

\[
\begin{align*}
\text{step 1. } & \Rightarrow \\
\text{step 2. } & \Rightarrow \\
\text{step 3. } & \Rightarrow \\
\text{step 4. } & \Rightarrow \\
\text{nezumi} & \text{neko} & \text{ga} & \text{nezumi} & \text{o} & \text{neko} & \text{ga} & \text{mita} \\
\text{nezumi} & \text{o}
\end{align*}
\]

Note that the case-markers in Japanese are taken as minimum SOs in the sense that they have syntactic functions (i.e. case-marking). Selecting a pair of syntactic words, the noun neko (‘cat’) and the nominative marker -ga, starts the derivation. Since the case-marker does not select a following dependent, the next

---

3 It would be conceptually simpler if every instance of linearization could be reduced to interface requirements. This thesis limits the scope of investigation to local alternations of dependents, and leaves investigation of the prosodic properties of head-dependent orderings for future work. (For PF-based approaches to head-dependent orderings, see, for example, Neeleman and Weerman 1999 for an analysis of Case adjacency as a PF requirement, and Donati and Nespor 2003 for a phonologically based analysis of head-complement orderings.)
noun *nezumi* ('rat') targets the rightmost node and merges to the node in step 2. Then, the accusative case-marker -o merges to the rightmost terminal noun *nezumi* in step 3, because the noun requires a following case-marker. Finally, the verb *mita* ('saw') merges to the rightmost node in step 4.

Having seen how left-to-right structure-building differs from bottom-up structure-building, let us go back to the original question of how left-to-right structure-building is motivated by the Prosodic Phase Hypothesis.

### 2.2.2. Motivating Left-to-Right Structure-Building

First of all, the left-to-right structure-building mechanism provides an explanation for the alleged non-isomorphism between syntactic and prosodic constituencies. Consider the often-cited example from Chomsky and Halle (1968:372):

\[(16)\]

\[a.\quad \text{This is } [\text{DP}, \text{the cat } [\text{CP}, \text{that caught } [\text{DP}, \text{the rat } [\text{CP } \text{that stole } [\text{DP}, \text{the cheese } \ldots \]

\[b.\quad \text{[IntP1: This is the cat]} \quad [\text{IntP2: that caught the rat}] \quad [\text{IntP3: that stole the cheese}] \ldots \]

The bracketing in (16a) represents the syntactic phrasing by DPs and CPs, and the bracketing in (16b) represents the prosodic phrasing by Intonational Phrases (IntPs). In left-to-right structure-building, DP, CP and IntP are constituents at a certain point of the derivation. For example, at the point when the noun ‘cat’ is merged as in (17a) below, the constituent indicated by α corresponds to the first IntP1. Then, at the point when the noun ‘rat’ is merged, the constituent indicated by β in (17b) corresponds to the first DP1 ‘the cat that caught the rat’, and the constituent indicated by γ corresponds to the relative clause CP1 and the second IntP2 ‘that caught the rat’. At the point in (17b), the first IntP1 (= α in (17a)) does not form a constituent any more, and the topmost node only represents the entire sentence (= IP).

\[(17)\]

\[a.\quad \alpha \quad (= \text{IP, IntP1}) \quad \to \quad \beta \quad (= \text{DP1})\]

\[\text{this} \quad \text{is} \quad \text{the cat} \quad \Rightarrow \quad \text{this} \quad \text{is} \quad \text{the cat that caught the rat} \quad \beta \quad (= \text{DP1}, \text{IntP2})\]

\[\gamma \quad (= \text{CP1, IntP2})\]

In other words, not only syntactic constituency but also prosodic constituency are read off from the structure, but at different points of the derivation (cf. Phillips 1996:ch.4). In contrast, in bottom-up structure-building, prosodic constituents such as IntPs are not read off from the structure because there is no such point where the string of words ‘this is the cat’ or ‘that caught the rat’ exists as an independent constituent. For example, the syntactic node that dominates the string of words ‘this is the cat’ is the topmost IP, and the IP also dominates the following elements ‘that caught...’ which are in the lower position of the structure and hence should already be merged. The Prosodic Phase Hypothesis argues in favor of left-to-right structure-building, because it is a pair of syntactic objects (SOs) and prosodic objects (POs) (e.g. α or γ in (17)), not just an SO, that qualifies as the unit of spell-out into the phonological component.

The role played by POs is independently recognized in the performance systems. For example, Nespor and Vogel (1986) point out that they are “the first level of processing in the perception of speech” (p.249). This line of reasoning has been motivated in the psycholinguistic literature as well. For example, Gee and Grosjean (1983) argue for the psychological reality of prosodic structures as a mediator between syntactic and performance structures. Gee and Grosjean define “ϕ-phrase” as a unit of performance
structure-building. The $\phi$-phrase is created by breaking the input string after each syntactic head that is a lexical word. Therefore, a $\phi$-phrase is a PO, which is formed using syntactic information such as head. Based on psycholinguistic experiments, they give independent motivation for prosodic structure which is proposed elsewhere by a linguistic (competence) theory as the domain of phonological rules (e.g. Nespor and Vogel 1986).\footnote{Refining the definition of $\phi$-phrase proposed by Gee and Grosjean, Abney (1991) pursues the idea that performance structure-building (i.e. parsing) proceeds by "chunks". Abney discusses not only how input strings are converted into chunks, but also how chunks are converted into a sentence. It is noted that "the order in which chunks occur [in a sentence] is much more flexible than the order of words within chunks" (Abney 1991:257). The flexibility in the ordering of chunks is likely to be related with the flexibility of intonational phrasing in English (as we will see in 2.4.2.2).}

Note that the syntactic definition of a Phase (e.g. CP and vP) proposed in Chomsky (1998, 1999, 2001) cannot be accommodated in left-to-right structure-building.\footnote{This does not mean that the Prosodic Phase Hypothesis would completely dispense with the syntactic notion of Phases (i.e. CP and vP). On the contrary, syntactic phenomena such as Strict Cyclicity still need to be explained with reference to syntactic cycles. See 2.5.2 for related discussion.} For example in (17b), if we were to spell out the first CP (= $\gamma$) encountered during incremental structure-building as a Phase, the structure dominated by the node $\gamma$ would not exist and we would not be able to continue to build the structures which were dominated by this CP. In order to maintain purely syntactic structure while defining a Phase as SO-PO pairing, I distinguish between syntactic structure and prosodic structure, and argue that POs are incrementally built in parallel with SOs, and derivationally paired with SOs by the syntax-prosody mapping. Let us now look at the properties of POs, and then how the mapping from syntax to prosody takes place.

2.3. Building Prosodic Objects

I assume that prosodic structure-building is subject to the conditions in (18):

(18) Conditions on prosodic structure-building
i. Operation: Prosodic Merge
ii. Elementary unit of the prosodic object: prosodic word
iii. Direction of Merge: left-to-right
iv. Branching: multiple
v. Locus of Merge: the root node
vi. Other available information: e.g. presence or absence of lexical stress

First, the operation "Prosodic Merge" defined as in (19) merges a prosodic word into another prosodic object (PO) (= conditions in (18i)-(18ii)).

(19) Prosodic Merge

Merge takes a pair of prosodic objects (PO$_n$, PO$_o$) and replaces them by a new combined prosodic object (PO$_{n,o}$).

Following the standard prosodic structure theory (Selkirk 1984, 1986, Nespor and Vogel 1986), I assume that the elementary unit of the PO is a prosodic word. A prosodic word is phonologically identified as the domain of lexical stress, and usually corresponds to a lexical word and its satellite function word(s) (Selkirk 1996). (Further discussion on prosodic words will be given in 2.4.2.1.)

Since the mapping from syntax to prosody is uni-directional, the way prosodic structure is built follows the way syntactic structure is built. First, the starting point of prosodic structure-building is when two prosodic words are merged. Then, another prosodic word is merged to the existing PO incrementally in the order in which prosodic words are pronounced (= condition in (18iii)). However, the resulting prosodic structure looks different from syntactic structure because of the properties specific to prosodic
structure such as the non-recursivity condition (Selkirk 1984, Nespor and Vogel 1986, Selkirk 1996):

(20) Non-recursivity: No PO' dominates PO, j = i. (Selkirk 1996:190)

For example, as we saw in the example (16b) (repeated below), no Intonational Phrase (IntP) dominates an IntP.

(16) a. This is [DPI the cat [CPi that caught [CP3 the rat [CP3 that stole [CP3 the cheese ... 
   b. [IntP This is the cat] [IntP that caught the rat] [IntP that stole the cheese]...

The non-recursivity condition does not allow embedded IntPs, and hence requires prosodic structure to be flat with possibly multiple branching (= condition in (18iv)). Consequently, a newly selected prosodic word must be merged at the root node (= condition in (18v)). This contrasts with syntactic structure that may contain recursive DPs or CPs as in (16a). Ultimately, the non-recursivity condition on prosodic structure may follow from exhaustiveness of the prosodic parse, in the sense that the sensori-motor system can produce or process one and only one PO at one time. I assume, however, that the non-recursivity condition in (20) applies only to IntPs, but not necessarily to prosodic words (for the reason to be discussed in 2.4.3.2).

The derivations of the prosodic structures for the English sentence 'the cat saw the rat' in (13) and the Japanese sentence neko-ga nezumi-o mita 'the cat saw the rat' in (15) are illustrated in (21). (Syntactic words within a prosodic word are connected by hyphens in prosodic structure.)

(21) a. Derivation of the prosodic structure for the English sentence in (13): 'the cat saw the rat'

step 1.⇒

the-cat saw

step 2.⇒

the-cat saw the-rat

b. Derivation of the prosodic structure for the Japanese sentence in (15):
neko ga nezumi o mita 'the cat saw the rat'

step 1.⇒

neko-ga nezumi-o

cat-Nom rat-Acc'

step 2.⇒

neko-ga nezumi-o mita
'saw'

In (21a), prosodic structure-building starts when a pair of prosodic words, 'the-cat' and 'saw', are merged. Then, the prosodic word 'the-rat' is merged at the root node of the PO 'the-cat saw', resulting in a ternary branching structure. Note that the string of two syntactic words 'the-cat' is identified as one prosodic word, because the prosodic structure-building mechanism can access the information about lexical stress (= (18vi)), and hence knows that the stressless determiner 'the' does not form a prosodic word on its own.

In the Japanese example in (21b), prosodic structure-building starts when a pair of prosodic words, neko-ga ('cat-Nom') and nezumi-o ('rat-Acc'), are merged. Then, the prosodic word mita ('saw') is merged at the root node of the PO neko-ga nezumi-o. Again, the resulting structure has a ternary branching.

We have seen how SOs and POs are built: both of them are incrementally built from left-to-right, but syntactic structure and prosodic structure exhibit different configurations due to recursivity of syntactic constituency and non-recursivity of IntPs. Next, let us look at how the syntax-prosody mapping takes place.
2.4. Pairing Syntactic Objects and Prosodic Objects

The central claim of this thesis is the Prosodic Phase Hypothesis in (1), repeated below:

(1) The Prosodic Phase Hypothesis
A syntactic object is spelled out as a prosodic object.

Prosodic objects (POs) are built in parallel with syntactic objects (SOs), and they are related to each other by uni-directional mapping from syntax to prosody. The conditions on syntactic and prosodic structure-building are summarized in (22):

(22) Conditions on syntactic and prosodic structure-building

<table>
<thead>
<tr>
<th>Operation</th>
<th>Syntactic structure</th>
<th>Prosodic structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary unit</td>
<td>syntactic word</td>
<td>prosodic word</td>
</tr>
<tr>
<td>Branching</td>
<td>binary</td>
<td>multiple</td>
</tr>
<tr>
<td>Merge</td>
<td>Direction</td>
<td>left-to-right</td>
</tr>
<tr>
<td>Locus</td>
<td>any accessible target</td>
<td>root node</td>
</tr>
</tbody>
</table>

The SO-PO pairs that are identified as the domain of an intonation contour at the PF interface (i.e. PO = Intonational Phrase) play an important role as units of spell-out. Since the intonation pattern of a sentence changes depending on semantic/pragmatic factors such as focus, let me present operational definitions of certain semantic/pragmatic notions I will use in the rest of the thesis.

2.4.1. Operational Definitions of Semantic/Pragmatic Notions

This section begins by introducing the syntactic feature \([\text{Focus}]_{SO}\), which is realized at both PF and LF interfaces. Two types of focus, “broad-focus” and “narrow-focus” are subsequently distinguished in terms of whether or not the feature \([\text{Focus}]_{SO}\) can project onto higher syntactic nodes at LF.

2.4.1.1. The Feature \([\text{Focus}]\)

First of all, I distinguish focus from presupposition. If, for example, we follow Jackendoff (1972), the focus-presupposition structure of the sentence in (23a) is (23b). (The element with extra prosodic prominence is indicated by upper case letters.)

(23) a. Kay donated FIVE novels to the library.
    b. Focus = five

   Presupposition: \(\lambda x (\text{Kay donated } x \text{ novels to the library})\)

In (23a), the numeral ‘five’ carries extra prosodic prominence (i.e. increased prominence relative to normal lexical stress) when the sentence encodes the focus-presupposition structure in (23b). In other words, the semantic/pragmatic notion of focus has a prosodic correlate in this case. Since I assume the architecture of the language faculty where PF and LF do not interact with each other directly (see the model of the language faculty in (4)), I postulate a syntactic feature \([\text{Focus}]_{SO}\) (following e.g. Jackendoff 1972, Rochemont 1986), which is mapped to prosodic structure as \([\text{Focus}]_{PO}\) and realized as prosodic prominence at PF, and realized as a focus interpretation at LF.\(^6\)

\(^6\) The definition of the feature \([\text{Focus}]_{SO}\) in (24) does not cover structural focus such as clefts.
(24) Syntactic feature [Focus]$_{so}$
   a. can be freely introduced in the Numeration,
   b. is mapped to prosodic structure as [Focus]$_{po}$ and realized as prosodic
      prominence (e.g. focal stress) at the PF interface, and
   c. is realized as a focus interpretation at the LF interface.

Although semantic focus and prosodic prominence are correlated, their relation is not isomorphic. In the following sections, I will introduce two types of focus, “broad-focus” and “narrow-focus”, which are associated with different prosodic properties.

2.4.1.2. Broad-Focus

A semantically focused element does not necessarily carry extra prosodic prominence. This is exemplified in “broad-focus” sentences (Ladd 1980, 1996:162-163,198-199). Broad-focus sentences are elicited by setting up a question that requires an answer with the whole sentence (IP) as an informational focus, as is illustrated in (25):

(25) A: What happened yesterday?
    B: Kay donated five novels to the library.

In (25B), the whole sentence corresponds to the focus domain, and the sentence is pronounced in one intonation contour without any extra prosodic prominence (Downing 1970). I call the intonation pattern associated with a broad-focus sentence the “default intonation pattern”. In a sentence with the default intonation pattern, the rightmost prosodic prominence (e.g. lexical stress) in the sentence is defined as the “default sentence stress” (or “nuclear stress” in e.g. Chomsky and Halle 1968). In (25B), the rightmost prosodic word ‘library’ carries lexical stress, which is defined as default sentence stress (indicated by bolding).

The default sentence stress is not always the most prominent prosodic peak in a sentence (Bolinger 1989, Ladd 1996:203, Hirst and Di Cnsto 1998:19). The “rightmost-ness” of default sentence stress, however, is significant in explaining possible focus interpretations associated with a sentence with the default intonation pattern. For example, a sentence with the default intonation pattern can serve as an answer to questions not only with IP-focus as in (26B1), but also with VP-focus as in (27B1), and with Adv-focus as in (28B1) (focus domains are underlined and bracketed in the examples below):

(26) [IP]-focus
    A: That class room is very quiet. What is going on there?
    B1: [IP The students are following the lecture attentively].

(27) [VP]-focus
    A: What are the students doing?
    B1: They are [VP following the lecture attentively].
        >
    B2: The students are [VP following the lecture attentively].
    B3: [VP Following the lecture attentively].

(28) [Adv]-focus
    A: How are the students following the lecture?
    B1: They are following it [Adv attentively].
        >
    B2: The students are following the lecture [Adv attentively].
    B3: [Adv Attentively].

The possible focus interpretations related to default sentence stress are an instance of “focus projection”
In principle, any phrase that dominates the prosodic word with default sentence stress can be interpreted as a focus. The focus projection formulates this by a percolation of a syntactic feature \([\text{Focus}]_{SO}\) ("\([F]\) so" for short) onto higher syntactic nodes (indicated by dotted arrows):

\[
\begin{align*}
\text{(29) a. Focus projection} \\
[\text{Focus}]_{SO} & \text{can project to any node that dominates it, when } [\text{Focus}]_{SO} \text{ dominates the default sentence stress position.} \\
& \text{(NB: In English, the rightmost prosodic word in a sentence is the default sentence stress position, as will be defined in (41a).)} \\
\end{align*}
\]

\[
\begin{align*}
\text{b.} \\
\text{IP } [F]_{SO} & \Rightarrow \text{focus in (26)} \\
\text{VP } [F]_{SO} & \Rightarrow \text{focus in (27)} \\
\text{V} & \Rightarrow \text{focus in (28)} \\
\text{Adv}[F]_{SO} & \Rightarrow \text{focus in (28)} \\
\text{the lecture} & \Rightarrow \text{focus in (28)} \\
\text{(default sentence stress)} & \\
\end{align*}
\]

Since focus projection is a bottom-up operation, it is considered to apply to an LF representation after the whole syntactic structure is built from top to down.

Comparing the (B1) sentences in (26)-(28), the default sentence stress on the Adv in (26B1) is not prosodically distinct, whereas the default sentence stress in (27B1) or (28B1) exhibits the most prominent peak in the sentence (cf. Cinque 1993:263). This follows from the fact that (26B1) is a broad-focus sentence whereas (27B1) and (28B1) are "partial" broad-focus sentences which contain a presuppositional part: in (27B1), ‘they are’ is presuppositional, and in (28B1), ‘they are following it’ is presuppositional. In English, a DP argument is normally a deaccented pronoun when presupposed. It is neither expressed by a lexical DP as in (27B2) or (28B2), nor deleted as in (27B3) or (28B3). (The greater than “\(\alpha > \beta\)” sign in the examples indicates that \(\alpha\) is more acceptable than \(\beta\).) Since there is a prosodic contrast between a focused element and a presuppositional element in the partial broad-focus sentences (27B1) and (28B1), default sentence stress is perceived as relatively more prominent than any other stress.

Thus, the prosodic way of encoding the focus-presupposition structure of a sentence is holistic rather than bivalent (Williams 1997:577), and semantic focus and extra prosodic prominence are not isomorphic to each other. This is because prominence may be realized by deaccenting as well as accenting. Following Williams (1997), I refer to presupposition as "anaphora", contrasting it with "focus" (or "disanaphora" in Williams 1997:594), and call the focus-presupposition information of a sentence "disanaphoricity". In (27B1), for example, the string ‘they are’ is anaphoric whereas the VP ‘following the lecture attentively’ is focused (or ‘disanaphoric’). English deaccents anaphoric elements in encoding disanaphoricity in partial broad-focus sentences.

2.4.1.3. Narrow-Focus

In contrast with a broad-focus that is obtained by focus projection, I call the focus whose \([\text{Focus}]_{SO}\) does not project "narrow-focus". This is exemplified in the sentence ‘Kay donated FIVE novels to the library’ we saw above in (23a). If we associate the sentence in (23a) with the default intonation pattern as in (30), the numeral ‘five’ in the theme DP is not included in the possible focus interpretations (which is called the "focus set"), because the numeral does not dominate the PP ‘to the library’ which is marked with \([\text{Focus}]_{SO}\) and dominates the default sentence stress position.
(30)  \[\text{[IP Kay \[VP \text{donated} \[DP \text{five novels}] \[PP \text{to the library}]\]}\]. \[= (25B)\]

a. Focus set: \{IP, VP, PP\}

b. Focus projection

\[
\begin{array}{c}
\text{IP} [F]_{SO} \\
\text{Kay} \\
\text{VP} [F]_{SO} \\
\text{donated} \\
\text{five novels} \\
\text{to the library}
\end{array}
\]

As the example in (23a) illustrates, a narrow-focus sentence does not exhibit the default intonation pattern, and extra prosodic prominence (e.g. "focal stress") is employed in order to mark a narrow-focus. The question-answer pair in (31) clarifies this point:

(31)  \[[\text{Numeral}]-\text{focus}\]

A:  How many books did Kay donate to the university library?
B:  She donated \([\text{Numeral}_FIVE]\) novels to the institution.

Likewise, if the question sentence identifies two non-adjacent elements as focused, the V ‘donated’ and the PP ‘to the library’ as in (32) for example, the intended focus interpretation cannot be obtained by focus projection, because there is no node that exhaustively dominates the V and the PP. In this case, one can assign focal stress to the V and the PP in order to encode disanaphoricity.\footnote{However, focal stress on the focused V is not necessary. See 3.1.4 for further discussion on non-adjacent foci examples.}

(32)  \[[\text{V}]+[\text{PP}]-\text{focus}\]

A:  What did Kay do with her collection of novels by Mishima?
B:  She \([\text{V}_F\text{DONATED}]\) them \([\text{PP}_F\text{to the LIBRARY}]\).

In (31)-(32), the element carrying focal stress is a narrow-focus (indicated by underscore), in the sense that the feature [Focus]_{SO} on it does not project. For expository purposes, I refer to sentences where disanaphoric elements are distributed on more than one syntactic object (e.g. V and PP in (32B)) as "double narrow-focus" sentences. However, this does not mean any commitment to a particular analysis of focus that allows a sentence to contain more than one semantic/pragmatic focus. On the contrary, I regard focus as disanaphora, i.e. a complement set of anaphora (Williams 1997:594), which implies that a sentence contains only one focused element semantically/pragmatically.

2.4.1.4. Summary

In sum, broad-focus and narrow-focus sentences are distinguished with respect to their semantic and prosodic properties in the following way:

(33)  a. A broad-focus sentence

i. contains a focus achieved by focus projection, and

ii. exhibits the default intonation pattern.
In order to encode disanaphoricity, English uses different degrees of prosodic prominence. I formulate the prosodic encoding of disanaphoricity using a descriptive scale of prosodic prominence:

(34) Disanaphoricity Scale

<table>
<thead>
<tr>
<th></th>
<th>anaphoric</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>deletion</td>
<td>(i)</td>
<td>deaccenting</td>
<td>(ii)</td>
<td>normal prominence</td>
</tr>
<tr>
<td></td>
<td>focused/disanaphoric</td>
<td></td>
<td>(iii)</td>
<td>extra prominence</td>
</tr>
<tr>
<td></td>
<td>(i.e. lexical stress)</td>
<td></td>
<td></td>
<td>focused/disanaphoric</td>
</tr>
</tbody>
</table>

In the Disanaphoricity Scale in (34), deletion (= (34i)) is regarded as the weakest prosodic prominence (Tancredi 1992, Williams 1997). As we saw in (27)-(28), English does not delete anaphoric elements, but retains them as pronominals. English pronominals are necessarily deaccented (= (34ii)) when used anaphorically.

Default sentence stress in English is not necessarily the most prominent peak in a sentence: it is only the rightmost lexical stress (= (34iii)) in broad-focus sentences, but can be perceived as relatively prominent in partial broad-focus sentences. Lexical items with normal lexical stress can be either anaphoric (as in narrow-focus sentences) or focused/disanaphoric (as in broad-focus sentences). A narrow-focus in English must carry extra prosodic prominence called focal stress (= (34iv)).

In the following sections, I will show how the syntax-prosody mapping algorithm establishes pairs of syntactic and prosodic objects in broad-focus sentences (2.4.2) and in narrow-focus sentences (2.4.3) in English and Japanese.

2.4.2. Broad-Focus Sentences with Default Sentence Stress

2.4.2.1. General Mapping Algorithm

A class of constraints on the syntax-prosody mapping is constituted by constraints on alignment of edges of syntactic objects (SOs) and prosodic objects (POs). I adopt the syntax-prosody mapping mechanism based on the edge alignment (Selkirk 1986), and use the generalized alignment constraint proposed by McCarthy and Prince (1993):

    Align (αCat, E; βCat, E): align the Edge of α with the Edge of β.
    (αCat/βCat ranges over syntactic and prosodic categories; E = Right or Left)

The Prosodic Phase Hypothesis in (1) (repeated below) is a general SO-PO mapping algorithm. When the relevant PO is an Intonational Phrase, which is phonologically identified as the domain of an intonation contour, the Prosodic Phase Hypothesis is implemented by the alignment constraints in (36).
(1) **The Prosodic Phase Hypothesis**
A syntactic object is spelled out as a prosodic object.

(36) **SO-PO mapping for maximal prosodic objects (= Intonational Phrases)**

a. **English**: Align (SO, R; PO, R) and Align (SO, L; PO, L), where SO = clause (CP/IP).

\[
\text{SO-PO mapping}
\]

\[
\text{SO: \{\text{the cat saw the rat}\}}
\]

\[
\text{PO: \{\text{the-cat saw the-rat}\}}
\]

b. **Japanese**: Align (SO, L; PO, L), where SO = phrase (XP).

\[
\text{SO-PO mapping}
\]

\[
\text{SO: \{\text{\[NP neko ga\] [VP\[NP nezumi o] mita]\}}
\]

\[
\text{PO: \{\text{\[np neko-ga\] [np nezumi-o] mita]\}}
\]

\[
\text{\text{cat-Nom rat-Acc saw}}
\]

\[
\text{\text{The cat saw the rat}}
\]

The alignment constraints for English in (36a) are read as ‘align the right edge of an SO to the right edge of a PO, and align the left edge of an SO to the left edge of a PO, where the SO is a clause (CP/IP)’. In English, a clause corresponds to the domain of an intonation contour in the default case, i.e. in broad-focus sentences. For example, the root clause ‘the cat saw the rat’ is pronounced in one intonation contour. I use a boundary tone (either Low or High) at the right edge as a phonological criterion for the prosodic level of Intonational Phrase (IntP) in English (Pierrehumbert and Beckman 1988:234).

(37) **Intonational Phrase in English**

is marked by a boundary tone (either L% or H%) at the right edge.

The IntP in English is identified as the domain of an intonation contour, and can be preceded or followed by an intonational break (Nespor and Vogel 1986:188, Levelt 1989:308).

The alignment constraint for Japanese in (36b) is read as ‘align the left edge of an SO to the left edge of a PO, where the SO is a syntactic phrase (XP)’. In Japanese, the maximal string that can be assigned an intonation contour is smaller than in English (Beckman and Pierrehumbert 1986:287), and the left edge of a syntactic phrase starts the domain of an intonation contour in the default case, i.e. in broad-focus sentences. For example, in the example in (36b), each part of the subject NP neko-ga (‘cat-Nom’) and the VP nezumi-o mita (‘rat-Acc saw’) forms the domain of an intonation contour. Notice that the verb does not form an intonational domain on its own, which suggests that it is the left edge, not the right edge, that corresponds to an intonational break in Japanese (Nagahara 1994). I use a boundary Low tone, which is traditionally referred to as ‘Initial Lowering’ at the left edge (McCawley 1968), as a phonological criterion for the IntP in Japanese.

(38) **Intonational Phrase in Japanese**

is marked by a boundary tone (L%) at the left edge.

---

8 The left edge and right edge alignment constraints can be unified in terms of “Wrap-XP” (XP = CP/IP in this case) proposed by Truckenbrodt (1995, 1999). I use two alignment constraints in order to clarify the contrast with Japanese.

(39) Downstep is a change in pitch register, which is manifested as a marked lowering in the stretch of an utterance following an accented syllable.

Japanese is a pitch accent language, where pitch features are specified in the lexicon and hence lexically distinctive. For example, the noun hasi [hashi] means two different things depending on the pitch feature: when the pitch is High-Low (accented), it means ‘chopsticks’, and when the pitch is High-High (unaccented), it means ‘end’. An accented word has a sharp pitch fall from High to Low while an unaccented word involves no pitch fall. This differs from English whose pitch features are “only postlexical or intonational” (Ladd 1996:156).

Let us look at more examples. The sentence given in (40a) below consists of accented words, and a pitch fall is found after each accented syllable. (Word-pitch is shown in parentheses under each word.) In contrast, the sentence in (40b) consists of only unaccented words, i.e. no words involve a falling pitch. In (40b), the first syllable of the prosodic word at the left edge of each syntactic phrase exhibits a Low tone, and downstep is triggered after these initial Low tones (Ladd 1996:148). Therefore, the left edge of an IntP corresponds to the left edge of a syntactic phrase, assuming the right-branching syntactic structure built incrementally via Merge. The domain of downstep does not start at the left edge of the head verb.

(40) a. Yuji ga biru ni wain o mazeta yo.9
   Yuji Nom beer to wine Acc mixed yo
   ‘Yuuji mixed wine to beer’
   F₀ contour:
   \[ \text{IntP} \text{ Yuji ga } (\text{HLL}) \text{ biru ni } (\text{HLL}) \text{ wain o } (\text{HLL}) \text{ mazeta yo } (\text{HLL}) \].

b. Mayumi ga Mamoru ni nuigurumi o moratta yo.
   Mayumi Nom Mamoru from doll Acc received yo
   ‘Mayumi received a doll from Mamoru’
   F₀ contour:
   \[ \text{IntP} \text{ Mayumi ga } (\text{LHH}) \text{ Mamoru ni } (\text{LHH}) \text{ nuigurumi o } (\text{LHHHH}) \text{ moratta yo } (\text{LHH}) \].

Thus, the SO-PO mapping constraints for maximal prosodic objects in (36) determine the maximal possible SO that can be mapped to an IntP. The English sentence, ‘the cat saw the rat’,

---

9 Two possible ways of ending a declarative sentence in Japanese are provided. One is the particle -yo, which is used to make the conversation sound colloquial and hence natural. The other is -mitai/-rasii yo ‘it seems that’/‘they say that’, which subordinates the sentence and also makes the conversation sound natural (Miyagawa 2001:329, fn.9). These particles may induce a final rise in F0 contour independently of the intonation contour of the rest of the sentence, which is not relevant to the present discussion and hence ignored.
corresponds to at least one IntP, whereas the Japanese sentence, *neko ga nezumi o mita* 'the cat saw the rat', corresponds to at least two IntPs. Given the difference in the size of the intonational domains in English and Japanese, the intonational domain in Japanese is often given a different name than "Intonational Phrase" in prosodic structure theories. For example, the intonational domain in Japanese is called "Major Phrase" in McCawley (1968) and Selkirk and Tateishi (1991), and "Intermediate Phrase" in Beckman and Pierrehumbert (1986). Hayes and Lahiri (1991) suggest that the intonational domain in Japanese is equivalent to Phonological Phrase, which is proposed as a level below the Intonational Phrase in the prosodic hierarchy (Nespor and Vogel 1986). The SO-PO mapping constraints in (36) derive the difference in the size of an intonational domain between English and Japanese from the difference in the size of a syntactic category (i.e. CP/IP versus XP), without assigning different prosodic categories to them.

The level of IntP plays a significant role under the Prosodic Phase Hypothesis in the following respects. First, when the relevant PO is an IntP, the SO-PO pair acts as a unit of spell-out. Secondly, IntPs are used to determine the domain of certain phonological processes. For example, an IntP is useful in defining the locus of default sentence stress. Default sentence stress is usually found in the rightmost syntactic phrase in English (cf. Chomsky and Halle 1968, Cinque 1993), and in the phrase immediately preceding the V in Japanese (cf. Selkirk and Tateishi 1991, Nagahara 1994). I formulate the assignment of default sentence stress as a PF rule which makes reference to the edge of the rightmost IntP, which I call "POn," in the sense that the rightmost IntP is the PO that is finally spelled out into the phonological component.11

(41) Default Sentence Stress Assignment
Default sentence stress is assigned to a prosodic word in the rightmost Intonational Phrase (= POn).
   a. In English, it is assigned to the rightmost prosodic word in the POn.
   b. In Japanese, it is assigned to the leftmost prosodic word in the POn.

In the English sentence, ['IntP the-cat saw the-rat'], default sentence stress (indicated by bolding) falls on the rightmost prosodic word 'the-rat' in the rightmost IntP, which corresponds to the whole sentence. In the Japanese sentence, ['Sup neko-ga][InP nezumi-o mita] ('cat-Nom, rat-Acc, saw'), default sentence stress falls on the leftmost prosodic word, *nezumi-o*, in the rightmost IntP, which corresponds to the VP, *nezumi-o mita*. The rightmost IntP (= POn) can be recognized as the PO which exhausts the prosodic words in the Numeration.

Let us now turn to the minimum prosodic object. The elementary unit of prosodic structure, i.e. a prosodic word, is empirically motivated as the domain of lexical stress. There is a systematic prosodic difference between lexical and functional category words (Gee and Grosjean 1983, Nespor and Vogel 1986, Selkirk 1996, Truckenbrodt 1999). In particular, a function word may be prosodically weak (i.e. stressless) and forms a prosodic word only with an adjacent lexical word. For example, in English a function word normally cliticizes onto the following lexical word (Selkirk 1999), whereas in Japanese a function word cliticizes onto the preceding lexical word (Uechi 1998). Therefore, when the relevant PO is a prosodic word, the Prosodic Phase Hypothesis is implemented by the alignment constraints in (42):

10 This view of IntPs is compatible with the results of corpus examination reported in Croft (1995). Croft shows that intonation units are almost always grammatical constituents, which he argues is an instance of grammaticalization. The Prosodic Phase Hypothesis is more restrictive than that, and argues that all IntPs have syntactic correlates.
11 My argument that it is an IntP that determines the unit of spell-out is empirically motivated by prosodic differences between English heavy NP shift and Japanese short-scrambling, as we will see in detail in chapter 3 and 4. Wagner (2003) also argues for prosodically based spell-out, but proposes a cyclic algorithm where syntactic objects are spelled out into the phonological component "as Soon as Possible" (p.7). The empirical domain Wagner concerns is stress pattern of complex predicate in West Germanic.
SO-PO mapping for minimal prosodic objects (= prosodic words)

a. English: Align (SO, R; PO, R), where SO = Lex.

![SO-PO mapping for English example]


![SO-PO mapping for Japanese example]

The alignment constraint in (42a) requires the right edge of a lexical word to be aligned with the right edge of a prosodic word, and the alignment constraint in (42b) requires the left edge of a lexical word to be aligned with the left edge of a prosodic word. Determiners in English and case-markers in Japanese are instances of function words. Therefore, the right edge of the determiner 'the' does not have to correspond to the right edge of a prosodic word in English, and the left edge of a case-marker does not have to correspond to the left edge of a prosodic word in Japanese. The first Prosodic Merge operation applies to the first two prosodic words 'the cat' and 'saw' in English, and neko-ga ('cat-Nom') and nezumi-o ('rat-Acc') in Japanese.

2.4.2.2. English

The SO-PO mapping constraint for maximal prosodic objects in English in (36a) specifies that an Intonational Phrase (IntP) corresponds to a clause in the default case in English. This is exemplified in (43):

(43) a. [\textit{IntP} The cat saw the rat].
    b. [\textit{IntP} My friend's baby hamster always looks for food in the corners of its cage].
    c. [\textit{IntP} This is the cat] [\textit{IntP} that caught the rat] [\textit{IntP} that stole the cheese] ...

In (43a), the whole sentence is the domain of an intonation contour, as well as in (43b), even though (43b) contains more lexical items. In (43c), each embedded relative clause (CP) forms the domain of an intonation contour.

Although a clause corresponds to an IntP as the default case, the intonational pattern in English is characterized by relatively large degree of variability (Nespor and Vogel 1986, Hirst 1998 citing Couper-Kuhlen 1986:153). The number of intonation contours in an utterance may be affected by prosodic factors such as weight (Nespor and Vogel 1986:194), semantic/pragmatic factors related to focus (Selkirk 1986:385, Levelt 1989:385), and performance factors such as rate of speech and style (Nespor and Vogel 1986:187,194, Levelt 1989:307). For example, it is possible that a sentence with a large number of lexical items as (44a) is pronounced with more than one intonation contour as in (44b) and (44c).
(44) a. \[\text{IntP My friend's baby hamster always looks for food in the corners of its cage}. \] \[(=43b)\]

b. \[\text{IntP My friend's baby hamster} \] always looks for food in the corners of its cage

c. \[\text{IntP My friend's baby hamster} \] always looks for food \[\text{IntP in the corners of its cage}. \)

(cf. Nespor and Vogel 1986:194)

However, variability in intonational phrasing does not mean that the phrasing is random. For example, the sentence in (45a) may be divided into smaller IntPs in a particular way as in (45b), but not arbitrarily as is shown in (45c).

(45) a. \[\text{IntP The cat saw the rat}. \] \[(=43a)\]

b. \[\text{IntP The cat} \] \[\text{intP saw the rat}. \]

c. \[\text{IntP cat saw the} \] \[\text{IntP rat}. \]

Let us look at how the examples in (45) are derived or excluded. (An SO that is already spelled out is shaded in syntactic structure.)

(46) Derivations in English

a. Example (45a): \[\text{IntP the cat saw the rat} \]

b. Example (45b): \[\text{IntP the cat} \text{IntP saw the rat} \]

c. Example (45c): \[\text{IntP cat saw the} \text{IntP rat}. \]

---

12 Nespor and Vogel (1986) do not use the greater-than “>” sign, but the sign in (44) reflects what they report in the text. The acceptability judgments presented in this thesis are based on judgments I collected from native speakers, unless the reference is given or otherwise noted.

13 Nespor and Vogel (1986) have pointed out that “there is a tendency to establish intonation contours of a more or less uniform, ‘average’ [weight]” (p194). (See Gee and Grosjean 1983 and Levelt 1989 for the same observation.) Levelt (1989) notes that the reason for this tendency is unclear, and speculates that “such a tendency to equal durations (if it can be substantiated for spontaneous speech at all) may have its raison d'être in articulatory, motor programming, either in size of the Articulatory Buffer, or in the convenient amount of air to be inhaled, or in both” (p386). I also take the view that the variable intonational phrasing patterns shown in (44) are performance variants. See 2.5.1 for further discussion.
In (46a), the whole sentence (IP) corresponds to an IntP (as indicated by dotted alignment lines), and gets spelled out at once. This is the default derivation of the sentence ‘the cat saw the rat’ in English. In (46b), the subject DP ‘the cat’ is paired with an PO and spelled out as an IntP, and then the VP ‘saw the rat’ is paired with another PO (= PO₂) and spelled out as an IntP. Since the SO-PO mapping constraint for IntPs in English in (36a) dictates that a clause corresponds to an IntP as the default case, this is not the default derivation of the sentence although the derivation does not violate any mapping constraint. I take the intonation pattern in (45b) as a performance variant that is allowed but not motivated in the grammar (see fn. 13).

In contrast, the derivation crashes in (46c). First, the function word ‘the’ is paired with a PO and spelled out as an IntP. This does not violate the SO-PO mapping constraint for prosodic words in English in (42a), because the constraint is only one-way from syntax to prosody, and does not require the right edge of a prosodic word to correspond to the right edge of a lexical word. In the next step, however, the noun ‘cat’ and the following syntactic words do not form a syntactic constituent and hence cannot be paired with a PO.

The alleged variability of intonational phrasing in English sentences aside, there are certain types of syntactic constructions that form intonation domains on their own. According to Nespor and Vogel (1986), these constructions include relative clauses, tag questions, vocatives, exclamative expletives, certain dislocated elements shown in (47a)-(47e) respectively, and parenthetical expressions such as ‘as you know’ in (48):

(47) a. My brother [IntP who absolutely loves animals] just bought himself an exotic tropical bird.
   b. That’s Theodore’s cat [IntP isn’t it]?
   c. [IntP Clarence] I’d like you to meet Mr. Smith.
d. [IntP Good heavens] there’s a bear in the back yard.
e. They are so cute [IntP those Australian koalas].

(48) a. [IntP As you know] Isabelle is an artist.
b. Isabelle [IntP as you know] is an artist.
c. Isabelle is [IntP as you know] an artist.
d. Isabelle is an artist [IntP as you know].

What is common among these constructions is that their edges correspond to the edge of clauses (CP/IP). Relative clauses and parentheticals are independent clauses on their own, and the other constructions always appear at the edge of the root clause. For example, a tag question in (47b) and a dislocated element in (47c) appear at the right edge of the root clause, and a vocative in (47e) and an exclamative expletive in (47d) appear at the left edge of the root clause. This conforms to the tendency in English for a syntactic clause to correspond to an intonational domain. Notice that when a clause is embedded in another clause as in (47a) and (48b)-(48c), the root sentence is divided into three IntPs as in (49b) (Nespor and Vogel 1986:189):

(49) a. * IntP
    ... IntP ... e.g. [IntP Isabelle [IntP as you know] is an artist]
b. IntP1 IntP2 IntP3
    ... ... ... e.g. [IntP1 Isabelle] [IntP2 as you know] [IntP3 is an artist]

The recursive intonational structure in (49a) is excluded by the non-recursivity condition for IntPs, which prohibits an IntP from containing another IntP (see (20)).

2.4.2.3. Japanese

The domain of an intonation contour in Japanese is smaller than that in English (Beckman and Pierrehumbert 1986:287), and the SO-PO mapping constraint for maximal prosodic objects in Japanese in (36b) specifies that the left edge of a syntactic phrase starts an Intonational Phrase (IntP) corresponds to a syntactic phrase in the default case in Japanese.

Apart from the difference in size, intonational phrasing in Japanese differs from that in English in that it is not subject to the same degree of variability. For example, the intonational phrasing pattern in (50a) is the only possible pattern for a broad-focus sentence, and the intonational patterns in (50b)-(50c) are unacceptable.

(50) a. [IntP Neko ga][IntP nezumi o mita].
    cat Nom rat Acc saw ‘The cat saw the rat’
b. #[IntP Neko ga nezumi o mita].
c. #[IntP Neko][IntP ga nezumi o][IntP mita].

Let us look at how the Japanese examples in (50) are derived or excluded.

(51) Derivations in Japanese
a. Example (50a): [IntP neko ga][IntP nezumi o mita]

syntactic structure
b. Example #(50b): \[\text{IntP} \text{neko ga nezumi o mita}\]

syntactic structure


\[\text{[neko ga]} \quad \text{[[neko ga] nezumi]} \quad \text{[[neko ga][nezumi o]]} \quad \text{[[neko ga][nezumi o] mita]}\]

(* The left edge of VP is not aligned with the left edge of an IntP.)

prosodic structure

\[\text{PO}_n \quad \text{[[neko ga] nezumi-o mita]}\]

c. Example #(50c): \[\text{IntP neko ga nezumi o} \text{IntP mita}\]

syntactic structure


\[\text{[[neko ga]} \quad \text{[[neko ga] nezumi]} \quad \text{[[neko ga][nezumi o]]} \quad \text{[[neko ga][nezumi o] mita]}\]

(* There is no SO such as 'ga nezumi o'.)

prosodic structure

\[\text{IntP neko-o} \quad \text{[neko ga] nezumi-o mita}\]

In (51a), the subject NP neko ga ('cat-Nom') is paired with an PO and spelled out as an IntP, and then the VP nezumi o mita ('rat-Acc, saw') is paired with another PO (= PO\(_n\)) and spelled out as an IntP. This is
the default derivation of the sentence, *neko ga nezumi o mita* ('cat-Nom, rat-Acc, saw'), in Japanese.

In (51b), the whole sentence (IP) corresponds to an IntP and gets spelled out at once. This derivation violates the SO-PO mapping constraint for IntPs in Japanese in (36b), because the left edge of the VP is not aligned with the left edge of an IntP.

In (51c), the derivation crashes. First, the lexical word *neko* ('cat') is mapped to a PO, satisfying the SO-PO mapping constraint for prosodic words in Japanese in (42b). At the next step, however, the nominative marker *-ga* and the following syntactic words do not form a syntactic constituent and hence cannot be paired with a PO.

2.4.3. Narrow-Focus Sentences with Focal Stress

This section shows how the syntax-prosody mapping algorithm establishes a pair of syntactic objects (SOs) and prosodic objects (POs) in narrow-focus sentences, where focal stress comes into play.

2.4.3.1. Mapping Constraints for [Focus]

The SO-PO mapping constraints for maximal objects in (36) (repeated below) apply to broad-focus sentences.

(36) **SO-PO mapping for maximal prosodic objects (= Intonational Phrases)**

a. English: Align (SO, R; PO, R) and Align (SO, L; PO, L), where SO = clause (CP/IP).

b. Japanese: Align (SO, L; PO, L), where SO = phrase (XP).

In narrow-focus sentences where a particular element is focused and carries extra prosodic prominence (e.g. focal stress), the intonation pattern of a sentence changes.

First of all, focal stress is prosodically distinguished from normal lexical stress of neighboring lexical items by a higher pitch peak. In English, an intonational break is often inserted after a lexical item with focal stress (Nespor and Vogel 1986:196). Consider the following examples. (The narrowly focused element is underlined and the element with focal stress is in upper case letters.)

(52) A: What ate the rat?
   B1: [\textit{\textup{IntP a CAT}} \textit{\textup{IntP ate it}}].
   >
   B2: [\textit{\textup{IntP a CAT ate it}}].

(53) A: What happened to the cat and the rat?
   B1: [\textit{\textup{IntP the cat ATE}} \textit{\textup{IntP the rat}}].
   >
   B2: [\textit{\textup{IntP the cat ATE the rat}}].

If the subject DP is a narrow-focus as in (52), the prosodic word 'a-cat' within the DP receives focal stress and an intonational break is inserted after the DP, as is illustrated in (52B1). Likewise, if the V is a narrow-focus as in (53), the V receives focal stress and an intonational break is inserted after the V, as is illustrated in (53B1). Thus, when focal stress is involved, an Intonational Phrase (IntP) gets divided into smaller IntPs, whereby the intonational pattern differs from the default intonation pattern in broad-focus sentences. The (B1) pattern is normally observed when the element following the prosodic word with focal stress contains a lexical word (e.g. 'ate' in (52B) and 'rat' in (53B)), and is not deaccented.

In Japanese, a prosodic word with focal stress starts a new domain of downstep, and the elements after the word get deaccented and incorporated into the same domain of downstep (Pierrehumbert and
Beckman 1988, Ishihara 2000).\(^{14}\) (A boosted pitch peak is indicated by an upward arrow and deaccented elements are in small letters.)

(54) A: Nani ga nezumi o tabeta no?
   what Nom rat Acc ate Q
   ‘What ate the rat?’

F0 contour

B1: \([\text{IntP NEKO ga}}\]
   \([\text{nezumi o tabeta rasii yo}].\]
   cat Nom rat Acc ate they.say yo
   ‘They say that a cat ate the rat’

B2: \(#_{\text{IntP NEKO}}\) ga
   [nezumi o tabeta rasii yo].

(55) A: Neko ga nezumi o dousita no?
   cat Nom rat Acc what.did Q
   ‘What did the cat do to the rat?’

B1: \([\text{IntP neko wa}][\text{IntF}}\]
   \([\text{nezumi o TABETA rasii yo}].\]
   cat Top rat Acc ate they-say yo

B2: \(#_{\text{IntP neko wa}}\)
   [nezumi o TABETA rasii yo].

When the subject NP is a narrow-focus, the prosodic word neko (‘cat-Nom’) receives focal stress and all following elements are deaccented. As a result, the whole sentence forms one IntP, as in (54B1). When the V is a narrow-focus, the V with focal stress starts a new domain of downstep and constructs an IntP on its own, as in (55B1). Thus, when focal stress is involved, the intonation pattern of a sentence changes from the default intonation pattern.

Based on these observations, I incorporate the prosodic effect of focal stress into the SO-PO mapping constraints in (56), and formulate the extension of an IntP in Japanese as a specific prosodic condition in (57):

(56) SO-PO mapping for maximal prosodic objects (= Intonational Phrases) with [Focus]

a. English: Align (SO, R; PO, R), where SO is with [Focus]SO.

\[
\begin{align*}
\text{SO:} & \quad [[\text{the cat}]_{\text{DP[PO]}}][\text{ate the rat}]_{\text{IP}} \\
\text{SO-PO mapping} & \quad \downarrow \\
\text{PO:} & \quad [[\text{the-CAT}[F]_{\text{PO IP}}][\text{ate the-rat}_{\text{IntP}}]
\end{align*}
\]

\[^{14}\text{Japanese normally deletes anaphoric elements when their informational content is recoverable from discourse context. For example, the most natural answers to the questions in (54A) and (55A) would be neko rasii yo ‘they say (that) a cat (ate the rat)’, and tabeta rasii yo ‘they say that (the cat) ate (the rat)’, respectively. (See chapter 4 for further discussion.) In this chapter, I discuss Japanese examples where anaphoric elements are retained, for the purpose of comparison with English. The point I would like to emphasize here is that the (B2) pattern is clearly worse than the (B1) pattern in (54)-(55).}\]
b. Japanese: Align (SO, L; PO, L), where SO is with [Focus]_SO.

\[\text{SO: } [\text{ipbpp}[\text{F}]_\text{so} \text{neko ga} [\text{nezumi o tabeta}] \]
\[\text{PO: } [\text{ipbpp}[\text{F}]_\text{po} \text{ga nezumi-o tabeta}]\]

\[\text{cat-Nom} \quad \text{rat-Acc} \quad \text{ate} \]

\[\text{The cat ate the rat'}\]

(57) **The Prosodic Condition on Japanese [Focus]_PO**

In Japanese, [Focus]_PO must be in the rightmost Intonational Phrase (= PO_\text{v}).

The SO-PO mapping constraint for English in (56a) requires the right edge of an SO with [Focus]_SO to be aligned with the right edge of an IntP. The alignment constraint for Japanese in (56b) requires the left edge of an SO with [Focus]_SO to be aligned with the left edge of an IntP. As I noted in 2.4.2.2-2.4.2.3, intonational phrasing in English is more variable than in Japanese. The effect of focal stress on intonational phrasing is also more variable in English than in Japanese. For example, not every native speaker of English shows preference for (B1) over (B2) in (52)-(53), whereas every native speaker of Japanese prefers (B1) over (B2) in (54)-(55). The SO-PO mapping constraint for IntPs with [focus] in English in (56a), however, does not accommodate the variability of intonational phrasing in English, and hence needs to be relaxed in order to rule in the (B2) pattern.

**The Prosodic Condition on Japanese [Focus]_PO** in (57) specifies that the prosodic feature [Focus]_PO must be contained in the rightmost IntP in Japanese, and derives the deaccenting phenomenon. Concerning the deaccenting phenomenon in Japanese, Ishihara (2000), following Nagahara (1994), proposes that any element following focal stress (or “additional-stress” in Ishihara 2000:155) is deaccented in head-final languages such as Japanese. I restate Ishihara’s deaccenting rule as a cross-linguistic generalization which makes specific reference to the prosodic level of IntP:

(58) **Deaccenting**

Any elements that follow a prosodic word with extra prosodic prominence are deaccented within an Intonational Phrase.

In English, the deaccenting phenomenon is normally invisible because a prosodic word with focal stress is always at the right edge of an IntP, as in (52B1) and (53B1). On the other hand, in the (B2) pattern in (52)-(53) where an IntP does not end in a prosodic peak, elements after the prosodic word with focal stress are deaccented. This suggests that the deaccenting phenomenon defined as in (58) is not specific to Japanese but holds for English too. In Japanese, deaccenting is realized as the extension of the size of an IntP, as in (54B1), and formulated by the Prosodic Condition in (57).\(^\text{15}\)

Furthermore, note that the SO-PO mapping constraints with specific reference to [Focus] in (56) are vacuously satisfied in broad-focus sentences where no focal stress is involved. In English, the prosodic word with [Focus]_PO that carries default sentence stress is at the right edge of the whole sentence, and hence at the right edge of an IntP. Also in Japanese, the prosodic word with [Focus]_PO that carries default sentence stress is at the left edge of an IntP. The Prosodic Condition in (57) is also satisfied in Japanese broad-focus sentences, because a prosodic word with [Focus]_PO (i.e. the default sentence stress position) is

---

\(^{15}\) The domain condition on deaccenting yields the same effect as the argument put forth by Donati and Nespor (2003) that one element bears main prominence within all prosodic units. For the level of IntP, they propose that “main prominence gives a cue for the informational structure of a sentence by signalling the constituent(s) that are to be interpreted as focused” (p.1126-1127). This statement implies that there is only one prosodic cue for focus per IntP, even though the number of focused constituent(s) in the sentence is more than one. Furthermore, Donati and Nespor explicitly argue that when there are more than one word with focal stress, the sentence is acceptable “only if they belong to two different intonational phrases” (p.1132, fn.14).
at the left edge of the rightmost IntP. Therefore, the SO-PO mapping constraints for maximal prosodic objects with [Focus] in (56) apply as a specific case, and the SO-PO mapping constraints for maximal prosodic objects in general in (36) apply elsewhere as the default case. In the following sections, we will look at more English and Japanese examples and see how they are derived or excluded.

2.4.3.2. English

In 2.4.2.2, we saw that intonational phrasing in English is characterized by a relatively large degree of variability. For example, a sentence with a large number of lexical items may be divided into smaller IntPs (see (44)). Containing a large number of lexical items is not the only way for a sentence to be divided into small IntPs; another way is to contain a lexical item with extra prosodic prominence (e.g. focal stress), as we saw in the previous section.

For example, while pronouns are inherently stressless, under certain discourse contexts “they become semantically prominent and must receive stress” (Nespor and Vogel 1986:196):

(59) a. [\text{ImP} Paul called Paula] [\text{ImP} before Carla called Carl],
b. #[\text{ImP} Paul called Paula] [\text{ImP} before SHE called HIM],
c. [\text{IntP} Paul called Paula] [\text{IntP} before SHE] [\text{IntP} called HIM].

(Nespor and Vogel 1986:196)

The sentences in (59) all have the same syntactic structure. In (59b)-(59c), however, the presence of the pronoun requires the listener to interpret this sentence in a specific way, namely, the pronoun ‘she’ referring to Paula and the pronoun ‘he’ referring to Paul, and the intended interpretation must be expressed prosodically by placing extra prominence on the pronouns. According to Nespor and Vogel (1986), this, in turn, causes the single IntP to be broken down into two smaller IntPs as in (59c). In other words, carrying extra prosodic prominence is another way to increase the prosodic weight of the PO, and the PO comes to qualify as an independent IntP (see Zubizarreta 1998:149 for the same point).

Thus, in English, certain types of focus must be realized as extra prosodic prominence with a higher pitch peak than a normal stress, and affect the intonational pattern of a sentence. Notice here that focus only indirectly affects intonational phrasing, via its prosodic realization as extra prosodic prominence (Shiobara 2002a).

Let us look at how sentences with focal stress are derived. First, consider again the example in (52) we saw above, where the subject DP is a narrow-focus.

(60) Derivations in English

a. Example (52B1): [\text{ImP [a CAT][ImP ate it]}

<table>
<thead>
<tr>
<th>syntactic structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>step 1.⇒</td>
</tr>
<tr>
<td>[DP[F]a cat]</td>
</tr>
<tr>
<td>prosodic structure</td>
</tr>
<tr>
<td>[\text{ImP a-CAT[F]c3}]</td>
</tr>
<tr>
<td>spell-out</td>
</tr>
</tbody>
</table>
b. Example (52B2): [\text{IntP} a \text{CAT ate it}]

syntactic structure

\begin{align*}
\text{step 1.} & \Rightarrow \\
[\text{DP}\{\text{a cat}\}] & \Rightarrow \\
[[\text{F}]_{\text{SO}} \text{a cat}] & \Rightarrow \\
[[\text{F}]_{\text{SO}} \text{a cat}] \text{ ate} & \\
\end{align*}

\text{prosodic structure}

\begin{align*}
\text{step 2.} & \Rightarrow \\
[\text{PO}] & \Rightarrow \\
[[\text{F}]_{\text{PO}} \text{a-CAT ate-it}] & \\
\end{align*}

(* The right edge of [\text{F}]_{\text{SO}} is not aligned with the right edge of an IntP.)

The sentence \[\text{'[IntP a CAT] IntP saw it'}\] is spelled out in two steps as is illustrated in (60a). When the subject DP \('a cat'\) is a narrow-focus and marked with [Focus]_{\text{SO}}, the PO \('a-cat'\) with [Focus]_{\text{PO}} is spelled out into the phonological component as an independent IntP, because the SO-PO mapping constraint for maximal objects with [Focus] in English in (56a) requires the right edge of the DP with [Focus]_{\text{SO}} to be aligned with the right edge of an IntP. If the whole sentence is paired with a PO and gets spelled out as one IntP as in (60b), it violates the SO-PO mapping constraint in (56a), because the DP with [Focus]_{\text{SO}} is not at the right edge of an IntP. This means that the constraint in (56a) is too restrictive to allow the intonation pattern in (52B2) and hence needs to be relaxed to the extent that the pattern is acceptable.

Note in passing that when a function word is in the sentence-final position as in \('saw it'\) in the examples in (52B1)-(52B2), it must cliticize onto the preceding lexical word to form a prosodic word (Pwd), e.g. \[\text{Pwd saw-it}\]. In this case, the right edge of the lexical word \('saw'\) is not aligned with the right edge of a prosodic word, violating the SO-PO mapping constraint for prosodic words in (42a). I stipulate here that the non-recursivity condition on prosodic structure (in (20)) applies only to Intonational Phrases (IntP), and not to prosodic words. Then we can assume a recursive structure for prosodic words, e.g. \[\text{Pwd [Pwd saw]-it]\], where the SO-PO mapping constraint for prosodic words is satisfied.\footnote{See Selkirk (1996) for phonological evidence for the recursivity of prosodic words.}

Next, consider the example in (53), where the V is a narrow-focus.

(61) Derivations in English

a. Example (53B1): [\text{IntP the cat ATE [IntP the rat]}

syntactic structure

\begin{align*}
\text{step 1.} & \Rightarrow \\
[\text{the cat}] & \Rightarrow \\
[[\text{V}]_{\text{SO}} \text{ate}] & \Rightarrow \\
[[\text{V}]_{\text{SO}} \text{ate}] & \Rightarrow \\
[[\text{V}]_{\text{SO}} \text{ate [the rat]}] & \\
\end{align*}

prosodic structure

\begin{align*}
\text{SO-PO} & \Rightarrow \\
\text{spell-out} & \\
\end{align*}

\text{PO}_n

\text{[IntP the-rat]}
b. Example (53B2): $\text{[IntP the cat ATE the rat]}$

syntactic structure

step 1. $\Rightarrow$  step 2. $\Rightarrow$  step 3. $\Rightarrow$  step 4.

[the cat]  \[[(\text{the cat})_{F} \text{ate}]\]  \[[(\text{the cat})_{F} \text{ate} \text{the}]\]  \[[(\text{the cat})_{F} \text{ate \text{the rat}}]\]

(* $[F]_{SO}$ is not at the right edge of an IntP.)

prosodic structure

The sentence "[IntP the cat ATE IntP the rat]" is spelled out in two steps, as in (61a). If the whole sentence is spelled out in one step as in (61b), it violates the SO-PO mapping constraint for IntPs with [Focus] in (56a), because the right edge of the V 'ate' with [Focus]$_{SO}$ is not aligned with the right edge of an IntP. Again, the constraint in (56a) needs to be relaxed if the intonation pattern in (53B2) is possible.

2.4.3.3. Japanese

As we saw in 2.4.3.1, the intonational pattern in Japanese is also systematically affected by the presence of focal stress. In particular, the left edge of a lexical item with focal stress starts a new domain of downstep and the domain of downstep spreads to the right edge of the sentence, with following elements deaccented. First, consider the derivation of the example we saw above in (54), where the subject NP is a narrow-focus:

(62) Derivations in Japanese

a. Example (54B1): $\text{[IntP NEKO ga nezumi o tabeta]}$ ('cat-Nom rat-Acc ate')

syntactic structure

step 1. $\Rightarrow$  step 2. $\Rightarrow$  step 3. $\Rightarrow$  step 4.

[neko ga]  \[[(neko ga)_{F} \text{nezumi}]\]  \[[(neko ga)_{F} \text{nezumi}]\]  \[[(neko ga)_{F} \text{nezumi o tabeta}]\]

SO-PO mapping

prosodic structure

b. Example #(54B2): $\text{[IntP NEKO ga]IntP nezumi o tabeta}$

syntactic structure
In (62a), the whole sentence corresponds to an IntP and gets spelled out at once. When the subject NP *neko ga* ('cat Nom') is a narrow-focus and marked with [Focus]SO, the feature [Focus]SO is mapped to prosodic structure as [Focus]PO and the derivation satisfies the SO-PO mapping constraint for maximal objects with [Focus] in Japanese in (56b), which requires that the left edge of the NP with [Focus]SO to be aligned with the left edge of an IntP. The derivation also satisfies the Prosodic Condition on Japanese [Focus]PO in (57), because the IntP is the rightmost IntP (= POn).

In contrast, when the sentence is spelled out in two steps as in (62b), the derivation violates the Prosodic Condition, because the PO containing the prosodic feature [Focus]PO is not in the rightmost IntP.

Next, consider the example in (55), where the V is a narrow-focus.

(63) Derivations in Japanese

a. Example (55B1): [{IntP neko wa}][IntP nezumi o]TABETA]
syntactic structure


[neko wa] [[neko wa] nezumi] [[n wa] [nezumi o]] [[n wa] [nezumi o] [vF3o tabeta]]

spell-out

b. Example #(55B2): [{IntP neko wa}][IntP nezumi o]TABETA]
syntactic structure


[neko wa] [[neko wa] nezumi] [[n wa] [nezumi o]] [[n wa] [nezumi o] [vF3o tabeta]]

(* [F]SO is not at the left edge of an IntP.)
The sentence, ‘[ *[neko wa]*[nezumi o]*[TABETA]’’, is spelled out in three steps as in (63a). When the V tabeta (‘ate’) is a narrow-focus and marked with [Focus]SO, the left edge of the V is aligned with the left edge of the rightmost IntP (= POa). If the sentence is spelled out in two steps as in (63b), it violates the SO-PO mapping constraint in (56b), because the V tabeta (‘ate’) with [Focus]SO is at the right edge, and not at the left edge of the rightmost IntP (= POa).

2.4.4. Summary: A Parallel Derivational Approach to Prosodic Structure

We have seen that English and Japanese intonation patterns differ in that the size of an Intonational Phrase (IntP) in broad-focus sentences is bigger in English than in Japanese (2.4.2). More specifically, an IntP corresponds to a clause (CP/IP) in English, and to a syntactic phrase (XP) in Japanese.

In addition, English and Japanese differ in how extra prosodic prominence (e.g. focal stress) changes the intonation pattern of a sentence (2.4.3). In narrow-focus sentences where focal stress is involved, an intonational break is inserted after a prosodic word with focal stress in English, whereas all elements after a prosodic word with focal stress are deaccented and incorporated into one intonational domain in Japanese.

(64) Intonation patterns in English and Japanese

<table>
<thead>
<tr>
<th>English</th>
<th>Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Broad-focus sentences</td>
<td>[ ... stress [hetp [=IP]] ] [hetp [=XP][ ] ][stress ... ]</td>
</tr>
<tr>
<td>ii. Narrow-focus sentences</td>
<td>[ ... FOCUS[ ] ] [ ... ][FOCUS ... ]</td>
</tr>
</tbody>
</table>

Thus, the two languages encode disanaphoricity (i.e. focus-presupposition structure) differently. This difference is attributable to the difference in the degree of mobility of prosodic prominence in English and Japanese: English can use mobile prosodic prominence overlaid in a sentence in order to encode a particular disanaphoricity pattern in a sentence, whereas prosodic prominence in Japanese is less mobile, and the elements after a prosodic word with extra prosodic prominence are all deaccented.

Ultimately, the unavailability of mobile prosodic prominence in Japanese may be reduced to the distinctiveness of lexically specified pitch features in the language. A traditional way of thinking about the relation between intonation and lexical features such as pitch accent or tone is “that all languages have ‘intonation’, and that in addition some languages have local pitch perturbations for ‘word accent’ or ‘tone’ overlaid on the global intonation” (Ladd 1996:30). In this sense, the correspondence between a clause (CP/IP) and an IntP in English is regarded as the default SO-PO mapping pattern cross-linguistically, and the size of the IntP is smaller than a clause in languages with lexically distinctive pitch accent such as Japanese. I will consider this issue from a broader cross-linguistic perspective in chapter 5.

The SO-PO mapping constraints for IntPs in (36) specify the general SO-PO correspondence relation by associating an IntP with a maximal possible SO, i.e. a clause in English and a phrase in Japanese. Let me compare the parallel derivational approach to the syntax-prosody mapping pursued in this thesis with a representational approach, where syntactic and prosodic representations are bidirectionally related to each other by mapping rules (as indicated by the bi-directional, dotted arrow in (65b)): 

(65) Derivational versus representational approaches to prosodic structure
A parallel derivational approach to the syntax-prosody mapping differs from a representational approach in that it may refer to the derivational steps of syntactic or prosodic structure-building. One of the theoretical consequences of a derivational approach is that syntactic structure must be built from left to right in the order in which syntactic words are pronounced. Incremental structure-building provides a natural way of implementing the Prosodic Phase Hypothesis which claims that it is certain SO-PO pairs that act as units of spell-out. Incremental structure-building mimics the way incoming sounds are parsed in ongoing sentence processing, and therefore has an independent motivation from the PF interface. Moreover, incremental structure-building can account for certain syntactic phenomena that are problematic for standard bottom-up structure-building (Phillips 1996, 2003). We will see some of them in the following chapters (in particular, 3.3.2 and 4.3.2).

A parallel derivational approach to prosodic structure is more restricted than a representational approach, in that the alignment constraints that determine legitimate and illegitimate SO-PO pairs are stated only uni-directionally, i.e. from the edge of an SO to the edge of a PO, and not vice-versa. For example, in English, the right edge of a clause must be aligned with the right edge of an IntP, but the right edge of an IntP does not have to be aligned with the right edge of a clause. In other words, alignment constraints do not prevent a clause in English from being divided into smaller IntPs. I will consider such examples in the next section from a performance perspective.

2.5. Interface Levels

2.5.1. PF Interface

The spell-out into the phonological component is a general linearization process. The Prosodic Phase Hypothesis argues that spell-out may apply more than once, and proposes that an Intonational Phrases (IntP) defines the unit of spell-out.

The POs spelled out into the phonological component are constrained by PF interface conditions imposed by the sensori-motor system. For example, an IntP is the domain of an intonation contour, and a prosodic word is the domain of lexical stress. Look at the model of the language faculty that represents the Prosodic Phase Hypothesis again:

(66) The model of the language faculty in the Prosodic Phase Hypothesis [= (4)]
We noted in 2.4.2.2 that the intonation pattern in English is affected by performance factors such as rate of speech and style. For example, it is possible that a sentence with a large number of lexical items is pronounced with more than one intonation contour.

(67) a. \[\text{My friend's baby hamster always looks for food in the corners of its cage}.\]
   >
   b. \[\text{My friend's baby hamster} \text{ always looks for food in the corners of its cage}].\]
   c. \[\text{My friend's baby hamster} \text{ always looks for food} \text{ in the corners of its cage}].\]

In particular, division of an IntP into smaller IntPs happens when the rate of speech is relatively slow (Nespor and Vogel 1986:187, 194, Levelt 1989:307). In contrast, when the sentence is pronounced relatively fast, it is likely that the whole sentence is pronounced in one intonation contour. I argue that the division of an IntP as in (67b) and (67c) only happens as the sentences are spelled out into the phonological component. That is to say, the intonational pattern illustrated in (67a) is derived from the default SO-PO mapping algorithm, and the ones in (67b) and (67c) are performance variants.

In addition to output conditions on POs, there are phonological rules and processes that apply in the phonological component. The Default Sentence Stress Assignment (in (41)) and the deaccenting phenomenon (in (58)) are such examples.

2.5.2. LF Interface

I adopt a conventional view of the LF interface in which LF is mapped from surface syntactic structure (Chomsky 1981). That is to say, an LF representation is a structure after all syntactic words are merged into one syntactic structure and spelled out into the phonological component (see the model of the language faculty in (66) above). There may or may not be covert movement operations (i.e. so called “LF movement”) applying after the last spell-out takes place. The conceptual-intentional system calculates the truth-conditional meaning of the root sentence from its LF representation.

Under the Prosodic Phase Hypothesis, the unit of spell-out is defined by correspondence relation between SOs and POs. While the Prosodic Phase Hypothesis emphasizes the significance of the PF properties of the unit of spell-out as an Intonational Phrase, it does not deny the significance of the LF properties associated with certain SOs. For example, propositional units such as CPs and vPs, which were
originally proposed as Phases (Chomsky 1999:9), are still considered to define the cycle of Strict Cyclicity in the syntax (see fn. 5). Furthermore, multiple mapping from SOs to LF representations is also compatible with a strictly derivational model of the grammar (Epstein et al. 1998). The mapping on the LF side of the grammar, however, is beyond the scope of this thesis.
CHAPTER III
Linearization of Verbal Dependents in English

This chapter shows how the Prosodic Phase Hypothesis explains a cluster of properties associated with linearization in the verbal domain in English. First, I show that the heavy NP shift alternation changes the intonation pattern of a sentence: the whole sentence consists of one intonation contour in the non-shifted order, whereas the rightmost DP forms the domain of an intonation contour on its own in the shifted order. Section 3.1 shows that constraints on the syntax-prosody mapping derive the generalization that the heavy NP shift alternation in English is prosodically conditioned, and the shifted order is more marked than the non-shifted order with respect to prosody. In 3.2, I speculate on the possibility that the prosodic condition on the marked, shifted order reflects efficiency of on-line sentence processing. Section 3.3 shows that the prosodic and syntactic properties associated with the marked order provide supporting evidence for the Prosodic Phase Hypothesis and for the claim that structure is built incrementally in the syntax. Lastly in 3.4, I give a brief review of previous analyses of shifted sentences, and compare them with the Prosodic Phase Hypothesis.

3.1. Prosodic Properties of Linearization in the Verbal Domain

3.1.1. Overview

This section shows that the Prosodic Phase Hypothesis in (1), as a general syntax-prosody mapping algorithm, accounts for the prosodic properties of the heavy NP shift alternation in English.

(1) The Prosodic Phase Hypothesis
A syntactic object is spelled out as a prosodic object.

In the heavy NP shift alternation, sentences with the non-shifted V – DP – PP order consist of one intonation contour, whereas in the shifted V – PP – DP order, the rightmost DP forms its own intonation contour. The generalization is that the shifted-order is available when the rightmost DP is “heavy” enough to form a prosodically legitimate Intonational Phrase (IntP), which will be formulated as a “prosodic weight condition”. The non-shifted order is always available and must be chosen when the DP is not prosodically heavy enough to form an IntP. I call the non-shifted order “canonical order” (C-order), and the shifted order “marked order” (M-order), in order to highlight the prosodic markedness of the shifted order.

(2) Linearization in the verbal domain in English
Let the C-order = V – DP – PP with the prosodic pattern \[ \text{IntP} \ldots \text{} \text{V} – \text{DP} – \text{PP} \], and
the M-order = V – PP – DP with the prosodic pattern \[ \text{IntP} \ldots \text{V} – \text{PP} \] \[ \text{IntP} \text{DP} \],
a. select M-order or C-order when M-order satisfies the Prosodic Weight Condition, and
b. select C-order elsewhere.

The methodology that I employed in reaching this generalization consists of acceptability judgments of pronounced stimuli sentences. In order to examine interface properties of the heavy NP shift alternation, I designed and conducted a systematic elicitation that includes two tasks, a pronunciation task and a judgment task. The purpose of the pronunciation task was to examine how a native speaker of English pronounces non-shifted and shifted sentences with particular focus interpretations. These sentences were recorded and subsequently used as stimuli for the judgment task, which was designed as a test of how natural or unnatural the stimuli sentences sound to other native speakers of English with particular focus interpretations, and with particular prosodic patterns. (See appendix to this chapter for the details of the pilot study.)
3.1.2. Broad-Focus Sentences

This section shows that broad-focus sentences with two different word orders, V – DP – PP and V – PP – DP, exhibit different intonation patterns in English, and the V – PP – DP order is prosodically more marked than the V – DP – PP order. I argue that the linear order between DP and PP dependents is determined by the relative prosodic weight of the DP dependent, and formulate this effect as a “prosodic weight condition”.

3.1.2.1. Default Sentence Stress

By setting up a question that requires an answer with the whole sentence (IP) as a focus, we can elicit a broad-focus sentence with the default word order and with the default intonation pattern. In the broad-focus sentence in (3B), the two internal arguments of the V ‘donated’, the theme DP ‘her collection of novels by Mishima’, and the goal PP ‘to the library’, appear in DP – PP order. In the broad-focus sentence in (4B), the theme argument DP, ‘more than ten California rolls’, of the V ‘ate’ precedes the adverbial adjunct PP ‘in five minutes’.

(3) A: What happened yesterday?
    B: Kay donated [DP her collection of novels by Mishima] [PP to the library].

(4) A: What happened yesterday?
    B: Ken ate [DP more than ten California rolls] [PP in five minutes].

Both in (3B) and (4B), the whole sentence is pronounced in one intonation contour. The constraint on the mapping of maximal prosodic objects in English in (5a) formulates the default intonation pattern in English:

(5) SO-PO mapping for maximal prosodic objects (= Intonational Phrases) [= (36) in 2.4.2.1]

a. English: Align (SO, R; PO, R) and Align (SO, L; PO, L), where SO = clause (CP/IP).
b. Japanese: Align (SO, L; PO, L), where SO = phrase (XP).

When the V is ‘donated’ or ‘ate’, the canonical order (C-order) is V – DP – PP, and it is associated with the default intonation pattern.

The prosodic level of Intonational Phrase (IntP) is significant in that it defines the unit of spell-out as well as the domain of certain phonological processes. For example, the Default Sentence Stress Assignment in (6) is formulated with reference to the rightmost IntP:

(6) Default Sentence Stress Assignment [= (41) in 2.4.2.1]

Default sentence stress is assigned to a prosodic word within the rightmost Intonational Phrase (= PO_r).

a. In English, it is assigned to the rightmost prosodic word in the PO_r.
b. In Japanese, it is assigned to the leftmost prosodic word in the PO_r.

In C-order, default sentence stress is found in the rightmost PP, namely on the rightmost prosodic word ‘library’ in (3B), and ‘minutes’ in (4B) (as indicated by bolding).

Recall that the “rightmost-ness” of default sentence stress is theoretically significant in explaining possible focus interpretations of a sentence with the default intonation pattern. For example, a sentence with the default intonation pattern is felicitous as an answer to the question not only with IP-focus as in (3B)-(4B), but also with VP-focus as in (7B1) and with PP-focus as in (8B1) (the intended focus is underlined in the examples below):
The possible focus interpretations related to default sentence stress (which is called the “focus set”) are an instance of “focus projection” of the syntactic feature [Focus]_S0 onto higher syntactic node(s).

In partial broad-focus sentences where [Focus]_S0 does not project up to the entire IP, English pronominalizes and hence deaccents anaphoric arguments as is seen in (7B1) and (8B1). English does not generally retain anaphoric arguments as lexical words ((7B2) and (8B2)), nor does it delete them ((7B3) and (8B3)) except in very informal conversations.

In broad-focus sentences, default sentence stress in English is not always the most prominent prosodic peak and the whole sentence may consist of prosodic words with normal lexical stress. In partial broad-focus sentences, default sentence stress may be perceived as relatively more prominent than other lexical stress, due to its prosodic contrast with deaccented anaphoric elements. The prosodic way English encodes (partial) broad-focus is summarized in (10).

<table>
<thead>
<tr>
<th>Disanaphoricity</th>
<th>Prosodic process</th>
<th>Deaccenting</th>
<th>Normal prominence</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Broad-focus (e.g. (3)-(4)) [ip [dp _] ]</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>b. Partial broad-focus (e.g. (7)-(8)) [ip [xp [dp _] ] ]</td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disanaphoricity Scale</th>
<th>(i) deletion (ii) deaccenting (iii) normal prominence (i.e. lexical stress) (iv) extra prominence</th>
</tr>
</thead>
<tbody>
<tr>
<td>focused/disanaphoric</td>
<td></td>
</tr>
</tbody>
</table>

e.g. In English,

(ii) pronominal
(iii) default sentence stress
(iv) focal stress
Therefore, English uses only normal prosodic prominence (= (11iii)) in broad-focus sentences, and the contrast between deaccenting (= (11ii)) and normal prominence (= (11iii)) in partial broad-focus sentences, in order to encode disanaphoricity (i.e. focus-presupposition structure).

3.1.2.2. The Distinction between Canonical and Marked Orders

Contrasting with C-order, I call the order where the two post-verbal dependents are flipped marked order (M-order), which corresponds to what is traditionally analyzed as the shifted order. Usually, M-order is unacceptable for broad-focus sentences.

(12) A: What happened yesterday?
   B: #Kay donated [pp to the library] [dp her collection of novels by Mishima].

(13) A: What happened yesterday?
   B: #Ken ate [pp in five minutes] [dp more than ten California rolls].

In (12B), the goal PP precedes the theme DP, and in (13B), the adverbial PP precedes the theme DP. When the V is followed by the two dependents, DP and PP, it is normally the case that a DP dependent appears immediately after the V. A DP may appear after the other dependent only as a marked case, as we will see shortly.

The adjacency effect between V and DP has been explained in terms of an adjacency requirement on structural Case assignment from V to DP (Stowell 1981). I assume that the distinction between C-order and M-order is originally due to this requirement.

(14) Case Adjacency
DPs must be adjacent to the Case-assigning verb.

I will show, however, that the adjacency requirement is violable and in fact overridden in sentences with M-order by the prosodic condition that motivates M-order.\footnote{This suggests that Case Adjacency should not be a hard principle in the grammar, but, for example, a prosodically based requirement, as is argued by Neeleman and Weerman (1999:40, fn.17).}

The existence of acceptable shifted sentences shows that M-order is not entirely prohibited for broad-focus sentences. For example, M-order may be used when the sentence-final DP is "heavy", e.g. contains more than a certain number of prosodic words and can be preceded by an intonational break (Rochemont and Culicover 1990, Zubizarreta 1998:149). Consider the following shifted sentences (syntactic words within a prosodic word are connected by hyphens):

(15) A: What happened yesterday?
C-order
B1: [int Kay donated five hundred Canadian dollars and her-collection of novels by Mishima to the library].
M-order
B2: [int Kay donated to the library] fiv five hundred Canadian dollars and her-collection of novels by Mishima.

(16) A: What happened yesterday?
C-order
B1: [int Ken ate six BC-rolls] [int four Alaska-rolls] [int two or three dynamite-rolls] [int and more than ten California-rolls] [int in five minutes].
M-order
B2: [int Ken ate in five minutes] [int six BC-rolls] [int four Alaska-rolls] [int two or-
Compared with the infelicitous shifted sentences in (12B) and (13B), the DPs in (15B2) and (16B2) contain more prosodic words and form Intonational Phrases (IntPs) on their own. In (15B2), the DP contains seven prosodic words (‘five, hundred,… by-Mishima’) and in (16B2) the DP contains nine prosodic words (‘six, BC-rolls,… California-rolls’). In these cases, M-order is available and optionality between C-order and M-order arises. Notice that in (16B2), each listed DP is explicitly surrounded by intonational breaks (see Nespor and Vogel 1986:200 and Levelt 1989:386 for the intonation pattern of lists). Regarding the locus of default sentence stress (i.e. the rightmost lexical stress) in M-order, it is found in the rightmost DP. This is predicted by the Default Sentence Stress Assignment in (6), because the rightmost DP is indeed the rightmost IntP. More specifically, the rightmost prosodic word ‘Mishima’ in (15B2) and the rightmost prosodic word ‘California-roll’ in (16B2) carry default sentence stress.

In sum, there is a distinction between canonical and marked orders in English, and the marked order is associated with a marked intonation pattern. In particular, the rightmost DP must form an IntP on its own, and achieve “right-edge heaviness”. This predicts that a parenthetical expression such as ‘I think’, which has its own intonation contour and forms an independent IntP (cf. 2.4.2.2), can occur before the rightmost DP. This prediction is borne out by the acceptable shifted sentence in (17B1):

(17) A: What happened yesterday?
B1: Kay donated to the library, I think, her collection of novels by Mishima.
B2: #Kay donated, I think, to the library her collection of novels by Mishima.

The degraded acceptability of the shifted sentence in (17B2), where the parenthetical expression occurs between the V and the PP, suggests that there can be no intonational break at this place (see Rochemont and Culicover 1990:119, Wasow 2002:ch.2 for the same effect). Given this observation, I propose the English-specific condition on the SO-PO mapping in (18), which derives prosodic markedness of M-order:

(18) The SO-PO Mapping Condition on English PPs
Align (SO, R; PO, R), where SO = a PP dependent, and PO = Intonational Phrase.

The alignment constraint in (18) is read as ‘align the right edge of a PP dependent with the right edge of an IntP’. Therefore, if a PP dependent is followed by a DP dependent, an intonational break is inserted between the PP and the DP. The canonical V – DP – PP order vacuously satisfies this condition, because a PP dependent is in sentence-final position in C-order. The idea behind the condition in (18) is an Optimality Theoretic competition (Prince and Smolensky 1993) with the Case Adjacency requirement in (14): M-order, which violates Case Adjacency, must be independently motivated by another requirement. In optimality theoretic terms, the SO-PO Mapping Condition on English PPs in (18) is given priority and wins over Case Adjacency in M-order (see fn.1). At the moment, I limit the category of non-DP dependents to PPs, but this can be extended to AdvPs that do not necessarily contain an overt P, as we will see in 3.3.1.1.

In light of (18), the degraded acceptability of the shifted sentences in (12B) and (13B), compared with the full acceptability of the shifted sentences in (15B2) and (16B2), is attributable to the prosodic property of the sentence-final DP: the DPs in (12B) and (13B) are not prosodically heavy enough to form IntPs on their own. Examine the following spontaneously produced shifted sentences, which support this conclusion.

(19) We-exclude from-consideration [DP candidates with-epenthetic syllables and-those with-multiple prosodic words dividing-up a-single morphological word].
(20) The Vatican said “no-ideology can-erase from-the-human spirit [DP the-certainty that-marriage exists solely between a-man and-a-woman]”…

(Vancouver Sun, 1 August, 2003)

(21) Iran has-refused to-send to-Canada [DP the-body of-a-man who-died after being flogged in-prison],…

(Vancouver Sun, 27 February, 2004)

The example in (19) is from a linguistics article, and (20)-(21) are from newspaper articles. Notice that the sentence-final DPs contain a fairly large number of prosodic words: eleven in (19) (‘candidates, with-epenthetic,… morphological, word’), seven in (20) (‘the-certainty, that-marriage,… a-man, and-a-woman’), and seven in (21) (‘the-body, of-a-man,… flogged, in-prison’).

As we have seen, the occurrence of the marked V – PP – DP order is more restricted than the occurrence of the canonical V – DP – PP order, in that the DP in M-order must be of some minimal weight so it can form an IntP. In Shiobara (2000, 2001), I argued that this right-edge heaviness effect is prosodic in nature, and the weight of the DP should be measured by the number of prosodic words, rather than syntactic words or phrases as is argued by Hawkins (1994) and Wasow (1997, 2002). This is based on the observation that the acceptability of shifted sentences varies depending on whether the DP contains a proper name or a pronoun:

(22) A: What happened yesterday?
   B1: Jack received from Betty [DP the-report on-Jim].
   B2: Jack received from Betty [DP the-report on-HIM]. (pointing at Jim in front of the speaker)
   B3: Jack received from Betty [DP the-report on-ME].
   >
   B4: #Jack received from Betty [DP the-report on-him/me].

The shifted sentences in (22B1)-(22B3) are not perfect but still acceptable compared with the shifted sentence in (22B4). In (B1)-(B3), the DP contains the proper name ‘Jim’ with default sentence stress (i.e. normal lexical stress) in (B1), or a stressed pronoun (indicated by upper case letters) in (B2)-(B3). In (B4), the rightmost word is a stressless pronoun, and the default sentence stress is shifted onto the lexical word ‘report’. The crucial difference between (B1)-(B3) and (B4) is prosodic in nature. In particular, a stressless pronoun does not form a prosodic word on its own. Therefore, DPs in (B1)-(B3) contain two prosodic words and are prosodically heavier than the DP in (B4) which contains only one prosodic word ‘report’.

The weight effect observed in (22) tells us that a large number of prosodic words is not the only way to achieve the status of a prosodically legitimate IntP. Another way is to assign extra prosodic prominence to a prosodic word, which can also make a shifted sentence acceptable:

(23) A: What happened yesterday?
   B1: Kay donated to the library [the novels by MISHIMA]. (cf. #(12B))
   B2: Ken ate in five minutes [TEN CALIFORNIA rolls]. (cf. #(13B))

3 One native speaker of English noted that the shifted sentence in (i) below sounds less acceptable than (22B1)-(22B3), but more acceptable than (22B4):

(i) Jack received from Betty [the REPORT on me] (and not a letter or anything else).

In (i), the head N ‘REPORT’ is contrastively focused and carries focal stress. If the gradient effect in acceptability is solely due to the prosodic weight of DPs, the measure of prosodic weight has to specify that a lexical word with extra prosodic prominence is heavier than a lexical word with default sentence stress (hence (i) > (22B4)), but not as heavy as two prosodic words (hence (22B1)-(22B3) > (i)).

(ii) The prosodic weight of DPs

[the report on Jim/HIM/ME] > [the REPORT on me] > [the report on him/me]

2 Pwds (underlined) > [n Pwds (where 2>n>1)] 1 Pwd

See related discussion on the examples in (23)-(26) below and also Shiobara (2000, 2001) for relevant examples.
When the rightmost DP contains a prosodic word with extra prosodic prominence as in (23B), it is either the case that the sentence encodes a broad-focus interpretation with the default sentence stress being realized as extra prosodic prominence, or that the speaker narrows down the focus onto the DP for some reason.

More instances of shifted sentences with DPs containing a small number of prosodic words are discussed in Wasow (1997, 2002):

(24) You'll find in your seat pocket [DP a copy of *Hemispheres*].
    (United Airlines flight attendant, as cited in Wasow 1997:83)
(25) ... take into account [DP inflation]
    (Aligned-Hansard corpus, as cited in Wasow 2002:22)
(26) ... assigns to the tokens [DP blocks]
    (talk by Jon Etchemendy, as cited in Wasow 2002:22)

Although Wasow does not indicate the discourse context these examples are situated in or the prosodic pattern of these examples, my analysis of the right-edge heaviness effect in M-order predicts that the rightmost DP carries extra prosodic prominence and forms an IntP.

Thus, M-order is prosodically more marked than C-order. For broad-focus sentences with M-order to be felicitous, the rightmost DP must form its own IntP, which is made possible by containing a large number of prosodic words or carrying extra prosodic prominence. In other words, the alternation between C-order and M-order in English is not entirely free but prosodically conditioned. The question remains as to exactly how heavy the rightmost DP must be in order to form a prosodically legitimate IntP. In the next section, I will illustrate how broad-focus sentences with C-order or M-order are derived, and formulate the prosodic property of the alternation in terms of a “prosodic weight condition” on DP dependents.

3.1.2.3. Derivations

Let us consider how broad-focus sentences are derived. First, let us look at the case where C-order is preferred over M-order.

(27) A: What happened yesterday?
    C-order
    B1: [\[s\] \(\text{Kay donated her collection of novels by Mishima to the library}\). \(= (3B)\)]
    M-order
    B2: #[\[s\] \(\text{Kay donated to the library}\) \[\text{her collection of novels by Mishima}\). \(= (12B)\)]

The derivation of the non-shifted sentence in (27B1) is shown in (28) below. Syntactic structure is incrementally built from left to right via Merge and spelled out into the phonological component. In the non-shifted sentence with C-order, the whole sentence IP corresponds to an IntP and gets spelled out at once. The default sentence stress is assigned to the rightmost prosodic word ‘library’ by the Default Sentence Stress Assignment in (6a).

(28) Derivation of the example (27B1): [\(s\) \(\text{Kay donated her collection of novels by Mishima to the library}\)]

Numeration {Kay, donated, her, collection, of, novels, by, Mishima, to, the, library}

\[
\downarrow
\]

syntactic structure
The derivation of the shifted sentence in (27B2) is shown in (29) below. Syntactic structure is built incrementally from left to right via Merge in the same way as the corresponding non-shifted sentence in (27B1). The difference is in when spell-out takes place. Because of the English-specific SO-PO Mapping Condition on PPs in (18), the right edge of the PP must be aligned with the right edge of an IntP. Therefore, when the PP is merged, the SO is paired with a PO and spelled out. Then, the rightmost DP is paired with another PO and spelled out as an independent IntP (= POₙ). When a native speaker of English is asked to pronounce this sentence, the rightmost lexical stress is found on the rightmost prosodic word ‘Mishima’ in the POₙ. (See appendix to this chapter for the details on how relevant sentences are pronounced.)

(29) Derivation of the example #(27B2): \[\text{Kay donated to the library} \rightarrow \text{her collection of novels by Mishima}\]

Numeration \{ Kay, donated, her, collection, of, novels, by, Mishima, to, the, library\}

syntactic structure

prosodic structure

The degraded acceptability of the M-order is attributable to the prosodic status of the sentence-final DP: although the SO-PO Mapping Condition on English PPs in (18) forces the DP to form an IntP on its own, the DP is not prosodically heavy enough to form the domain of an intonation contour. But exactly how heavy should the sentence-final DP be in order to qualify as a prosodically legitimate IntP? In order to answer this question, let us look at the case where M-order is available.

(30) A: What happened yesterday?
    C-order
    B1: \[\text{Kay donated five hundred Canadian dollars and her collection of novels by Mishima to the library}. \quad [= (16B1)]\]
M-order

B2: [$_{in_p}$ Kay donated to the library][$_{intP}$ five hundred Canadian dollars and her collection of novels by Mishima]. [= (16B2)]

The derivation of the shifted sentence in (30B2) is shown in (31). The derivation goes through the same steps as the derivation of the degraded shifted sentence we saw in (29). The difference is in how many prosodic words the PO$_n$ contains: in (30B2) the PO$_n$ contains seven prosodic words (‘five, hundred, Canadian, dollars, and-her-collection, of-novels, by-Mishima’) whereas in (29), the PO$_n$ contains only three prosodic words (‘her-collection, of-novels, by-Mishima’). In both cases, the PO that is spelled out before the PO$_n$ (which I call “PO$_{n-1}$”) is the same and consists of three prosodic words (‘Kay donated to-the-library’).

(31) Derivation of the example (30B2): [$_{in_p}$ Kay donated to the library][$_{intP}$ five hundred Canadian dollars and her collection of novels by Mishima]

<table>
<thead>
<tr>
<th>Numeration</th>
<th>{Kay, donated, five, hundred, Canadian, dollars, and, her, collection, of, novels, by, Mishima, to, the, library}</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>syntactic structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Kay donated] [=] [Kay [donated [to the library]]] == [Kay [donated [to the library]]]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>prosodic structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>[$<em>{intP}$ Kay donated to-the-library] == [$</em>{intP}$ five hundred... by-Mishima]</td>
</tr>
</tbody>
</table>

Remember that prosodic heaviness is achieved not only by a large number of prosodic words, but also by extra prosodic prominence:

(32) A: What happened yesterday?  
B: [$_{in_p}$ Kay donated to the library][$_{intP}$ the novels by MISHIMA]. [= (23B1)]

As a first step toward a formulation of the right-edge heaviness effect in M-order, I make the following generalization regarding the intonational status of the PO$_n$: in M-order, the rightmost DP must form a prosodically legitimate IntP, which is possible in one of the three ways in (33).

(33) PO$_n$ (= the rightmost Intonational Phrase) is a prosodically legitimate Intonational Phrase if  
   a. it is preceded by a parenthetical phrase, or  
   b. it is a CP/IP, or  
   c. it is prosodically heavy, satisfying the Prosodic Weight Condition (in (34)).

First, the condition in (33a) concerns the environment of the PO$_n$. As we saw in 3.1.2.2, a parenthetical phrase such as ‘I think’ in English is inherently an IntP, and surrounded by intonational breaks (see the example in (17)). Therefore, it serves the same function as an intonational break and can make the following phrase an independent IntP.
The condition in (33b) concerns the syntactic status of the PO. The SO-PO mapping constraint for maximal prosodic objects in (5a) specifies that a clause corresponds to an IntP in English. Therefore, when the rightmost phrase in a shifted sentence is a CP/IP, it forms an IntP and hence the shifted sentence should be acceptable. This is irrelevant to the current case where the rightmost phrase is a DP, and I will defer the discussion of this case until 3.3.1.1.

The condition in (33c) is a prosodic way of achieving the status of a legitimate IntP. Remember that there are two ways of achieving prosodic heaeviness, which is now formulated as the Prosodic Weight Condition in (34):

(34) The Prosodic Weight Condition
POₙ (= the rightmost Intonational Phrase) must be prosodically heavy by containing
a. a larger number of prosodic words than POₙ₋₁ (= the second rightmost
Intonational Phrase), or
b. a prosodic word with extra prosodic prominence.

The first option in (34a) concerns the prosodic weight of the POₙ relative to the weight of the POₙ₋₁. The second option in (34b) concerns the absolute prosodic weight of the POₙ.⁴ The Prosodic Weight Condition in (34) correctly rules out the shifted sentence in (27B2), where the POₙ contains only the same number of prosodic words as the POₙ₋₁. On the other hand, it correctly permits the shifted sentence in (30B2), where the POₙ contains seven prosodic words and the POₙ₋₁ contains only three prosodic words. The Prosodic Weight Condition also correctly rules in the shifted sentence in (32B), where the POₙ contains a prosodic word with extra prosodic prominence. At this point, I have no explanation for why the Prosodic Weight Condition is formalized the way it is in (34), and only speculate that the right-edge heaeviness effect may contribute to ease of sentence processing (cf. Hawkins 1990, 1994). I will come back to this issue in 3.2.

The evaluation of relative prosodic weight involves a pairwise comparison between the two IntPs at the right edge, POₙ and POₙ₋₁. In the derivational model of the computational component pursued in this thesis, the rightmost IntPs is identified as a PO that exhausts the Numeration, and the application of the first option of the Prosodic Weight Condition in (34a) is localized to a pair of IntPs at the right edge. The Prosodic Weight Condition is non-applicable to non-shifted sentences, because the whole sentence corresponds to an IntP and hence there is only one IntP in the non-shifted order.

Thus, we have reached the generalization we saw in (2) (repeated below), which states that M-order is available when it satisfies the Prosodic Weight Condition in (34), whereas C-order is always available.

(2) Linearization in the verbal domain in English
Let the C-order = V – DP – PP with the prosodic pattern [IntP ... V – DP – PP], and the M-order = V – PP – DP with the prosodic pattern [IntP ... V – PP] [IntP DP],

a. select M-order or C-order when M-order satisfies the Prosodic Weight Condition, and
b. select C-order elsewhere.

In other words, the Prosodic Weight Condition in (34) licenses the marked V – PP – DP order in English when the DP is prosodically heavier than the preceding part of the sentence. Word order optionality arises to the extent that the DP can be a prosodically legitimate IntP. Furthermore, the alleged gradient acceptability associated with English heavy NP shift is accounted for as an inherent property of prosodic weight, which is relative by nature, since it can only be evaluated relative to the number of prosodic words (as in (34a)) or degree of prosodic prominence (as in (34b)) of the neighboring element.

⁴ It would be desirable if we could unify the two disjunctive options in (34). One possible way would be to count a prosodic word with extra prominence as one and a half (or 'more than one and less than two') prosodic words as I suggested in fn.3. However, this does not work because it rules out the acceptable shifted sentence in (32B): if a prosodic word with extra prosodic prominence is worth one and a half prosodic words, the rightmost DP contains two and a half prosodic words ('the-novels, by-MISHIMA') and hence is lighter than the preceding IntP that contains three prosodic words ('Kay, donated, to-the-library').
So far, our observation has been limited to broad-focus sentences. The next step is to look at narrow-focus sentences with C-order or M-order, where focal stress comes into play.

3.1.3. Narrow-Focus Sentences

Narrow-focus sentences exhibit various prosodic patterns depending on the locus of focal stress. In this section, I will show that the Prosodic Weight Condition in (34) explains the acceptability of C-order as well as M-order by referring to the prosodic weight of the rightmost Intonational Phrase (= PO) in a sentence.

3.1.3.1. Focal Stress and the Prosodic Weight Effect

English uses extra prosodic prominence, called focal stress, to prosodically mark a focus that is not included in possible focus interpretations achieved by focus projection. For example, a V-focus interpretation is possible when the V carries focal stress (indicated by upper case letters) as in (35). Likewise, a V+PP-focus interpretation should be possible when both the V and the PP carry focal stress as in (36), although focal stress on the V is not necessary and degrades the sentence as is shown in the contrast between (36B1) and (36B2). I will discuss this in 3.1.4.2.

(35) V-focus
A: Did Kay sell her collection of novels by Mishima to the Asian library?
B: No. [\text{\textsc{imp}} She DONATED]\text{\textsc{her books to the institution}}.

(36) V+PP-focus
A: What did Kay do with her collection of novels by Mishima?
B1: #L, P\text{\textsc{She DONATED}}\text{\textsc{them to the Asian LIBRARY}}.
B2: \text{\textsc{IntP She donated them to the Asian LIBRARY}}.

In (35)-(36), the element carrying focal stress is a narrow-focus (indicated by underscore) in that the feature [Focus] so on it does not project onto higher syntactic nodes. For expository purposes, I refer to sentences where the focused element is distributed on one syntactic object (e.g. V in (35B)) as “single narrow-focus” sentences, and sentences where the focused elements are distributed on more than one syntactic object (e.g. V and PP in (36B)) as “double narrow-focus” sentences. This is not meant to imply any commitment to any particular theory of focus.

The narrow-focus sentences in (35B) and (36B) show that the locus of the main sentence stress may be shifted onto a narrow-focus, because of the discourse-based rule that determines that “the main stress of the phrase in 'focus' will win out, in relative terms, over that of the presupposed phrase” (Cinque 1993:258). Focal stress, whose locus varies depending on the disanaphoricity pattern of a sentence, differs from the default sentence stress which is assigned to the rightmost prosodic word by the Default Sentence Stress Assignment in (6a). Notice that the prosodic word ‘LIBRARY’ in (36B) is in the default sentence stress position but still carries extra prosodic prominence, not just normal lexical stress. This is because the sentence would encode a single V-focus interpretation if the PP did not carry focal stress. On the other hand, anaphoric arguments are pronominalized when possible, and prosodically deaccented.

Focal stress changes the intonation pattern of sentences too. In English, an intonational break is often inserted after a prosodic word with focal stress, and this effect is formulated by the SO-PO mapping.

---

3 According to Bolinger (1989), however, there is no absolute prosodic difference between default and focal stresses in English; rather, we hear focal stress more prominent than default stress because of "the degree to which it [= focal stress] overshadows all others" (p.354). The phonetic realization of different types of stress is a delicate issue (Ladd 1996:203), and Bolinger's statement contradicts the results of the pronunciation task of my elicitation, where a narrowly focused element often exhibits a higher pitch than the element in the default sentence stress position. (See appendix to this chapter for the details of elicitation.) In the examples reported in this chapter, I distinguish a prosodic word with extra prosodic prominence (indicated by upper case letters) from a prosodic word with normal lexical stress (indicated by holding if it is in the default sentence position) in relative terms.
constraint in (37):

(37) SO-PO mapping for maximal prosodic objects (= Intonational Phrases) with [Focus]
     [= (56a) in 2.4.3.1]
     in English, Align (SO, R; PO, R), where SO is with [Focus] so.

The way English encodes (double) narrow-focus prosodically is summarized in (38)

(38) Prosodic encoding of (double) narrow-focus in English

<table>
<thead>
<tr>
<th>Disanaphoricity</th>
<th>Prosodic process</th>
<th>(ii) Deaccenting</th>
<th>(iii) Normal prominence</th>
<th>(iv) Extra prominence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrow-focus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Single narrow-focus</td>
<td></td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>b. Double narrow-focus</td>
<td></td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

(39) Disanaphoricity Scale [= (11)]

<table>
<thead>
<tr>
<th>Anaphoric</th>
<th></th>
<th>(iii) normal prominence (i.e. lexical stress)</th>
<th>(iv) extra prominence</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) deletion</td>
<td></td>
<td>focused/disanaphoric</td>
<td></td>
</tr>
<tr>
<td>(ii) deaccenting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iii) default sentence stress</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iv) focal stress</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

e.g. In English,
(ii) pronominal
(iii) default sentence stress
(iv) focal stress

Thus, in narrow-focus sentences, English deaccents anaphoric elements (= (39ii)) and assigns focal stress to focused/disanaphoric elements (= (39iv)), regardless of where a narrow-focus appears in a sentence and how many.

The generalization in (38) makes the following prediction regarding the linearization of verbal dependents in narrow-focus sentences: when the DP is (a part of) a narrow-focus, it carries extra prosodic prominence and hence satisfies the Prosodic Weight Condition in (34), whereby the marked V – PP – DP order is available. Note that it is prosodic heaviness (e.g. focal stress), not a semantic/pragmatic focus, that directly determines the word order in English: a narrow-focus on the rightmost DP licenses M-order because a narrow-focus must be realized as prosodic prominence in English (Shiobara 2001, 2002a). On the other hand, the canonical V – DP – PP order is always available as the default option.

This prediction is borne out by the results of the elicitation I conducted in order to examine the prosodic properties associated with alternations between C-order and M-order. (The details of the elicitation are presented in appendix to this chapter.) The table in (40) summarizes the acceptability of C-order and M-order with five different focus interpretations. The sample sentences are presented in (41)-(45).

(40) Summary of acceptability
The examples in (41)-(45) contain the V ‘donated’ which takes two internal arguments, a theme DP and a goal PP. The results were: in broad-focus sentences, C-order is preferred over M-order, as is seen in (41); in narrow-focus sentences where the DP is (a part of) a narrow-focus, the word order preference varies depending on speakers, as is seen in (42) and (44); in narrow-focus sentences where the PP is (a part of) a narrow-focus, C-order must be chosen, as is seen in (43) and (45).

In the next section, I will illustrate how narrow-focus sentences with C-order or M-order are derived, and how the Prosodic Weight Condition in (34) explains the acceptability of narrow-focus sentences. After that, I will discuss the unexpected findings of the pronunciation and judgment tasks of the elicitation.

### 3.1.3.2. Derivations

We already saw the derivation of broad-focus sentences in 3.1.2.3. In broad-focus sentences, M-order is available when the rightmost Intonation Phrase (= PO) is prosodically heavier than the preceding one (= PO). For example, the shifted sentence in (41b) is degraded relative to the non-shifted sentence in (41a), because both PO (‘her-collection, of-novels, by-Mishima’) and PO (‘Kay, donated, to-the-library’) contain three prosodic words, and hence the PO is not heavier than the PO.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad</td>
<td>(41)IP</td>
<td>√ [Inp...V-DP-PP[F]]</td>
<td>√ [Inp...V-PP [DP,F]]</td>
</tr>
<tr>
<td>Single narrow</td>
<td>(42)DP</td>
<td>√ [Inp...V-DP-PP[F]] [PP]</td>
<td>√ [Inp...V-PP [DP,F]]</td>
</tr>
<tr>
<td>Double narrow</td>
<td>(43)PP</td>
<td>√ [Inp...V-DP-PP[F]]</td>
<td>√ [Inp...V-PP[F]] [DP,F]</td>
</tr>
<tr>
<td>Double narrow</td>
<td>(44)VP+DP</td>
<td>√ [Inp...V[F] [DP,F]] [PP]</td>
<td>√ [Inp...V[F] [PP,F] [DP,F]]</td>
</tr>
<tr>
<td>Double narrow</td>
<td>(45)VP+PP</td>
<td>√ [Inp...V[F] [DP-PP,F]]</td>
<td>√ [Inp...V[F] [PP,F] [DP,F]]</td>
</tr>
</tbody>
</table>

(41) Broad-focus (IP-focus)
[What happened yesterday?]
   a. [Inp Kay donated her collection of novels by Mishima to the library].
   b. #[Inp Kay donated to the library][Inp her collection of novels by Mishima].

(42) Single narrow-focus (DP[theme]-focus)
[What did Kay donate to the Asian library?]
   a. √#[Inp She donated her collection of novels by MISHIMA][Inp to the library].
   b. #[Inp She donated the library][Inp her collection of novels by MISHIMA].

(43) Single narrow-focus (PP[goal]-focus)
[Where did Kay donate her collection of novels by Mishima?]
   a. [Inp She donated them to the ASIAN LIBRARY].
   b. *[Inp She donated to the ASIAN LIBRARY][Inp them].

(44) Double narrow-focus (V+DP-focus)
[UBC Asian library is facing a serious financial problem. What did Kay do to save the library?]
   a. √#[Inp She DONATED][Inp her collection of novels by MISHIMA][Inp to the library].
   b. #[Inp She DONATED][Inp to the library][Inp her collection of novels by MISHIMA].

(45) Double narrow-focus (V+PP-focus)
[Kay will go back to Japan soon, but she had too many books to ship to Japan. Do you know what she did with her collection of novels by Mishima?]
   a. [Inp She donated/#DONATED][Inp them to the ASIAN LIBRARY].
   b. *[Inp She DONATED][Inp to the ASIAN LIBRARY][Inp them].
Notice here that the M-order is only degraded relative to the corresponding C-order (as is indicated by the "#" sign in (41b)). This contrasts with the severe ill-formedness of the shifted, PP-focus and V+PP-focus sentences (as indicated by the "*" sign in (43b) and (45b)). I hypothesize that the degree of acceptability of a sentence is determined by a general cumulativity measure for the violation of relevant constraints. For example, the more alignment constraints are involved in inducing an intonational break that separates PO_n from PO_{n+1}, the more severe the absence of right-edge heaviness becomes. The relevant alignment constraints are repeated in (46), and the Cumulativity Measure is formulated as in (47):

(46) a. The SO-PO Mapping Condition on English PPs [= (18)]
   Align (SO, R; PO, R), where SO = a PP dependent, and PO = Intonational Phrase.
   b. SO-PO mapping for maximal prosodic objects (= Intonational Phrases) with [Focus]
      [= (37)]
      in English, Align (SO, R; PO, R), where SO is with [Focus]_SO.

(47) The Cumulativity Measure
   The more constraints are involved, the more they should be respected.

The Cumulativity Measure in (47) concerns the application of constraints, and predicts that a violation of the Prosodic Weight Condition in (34) is more severe when the intonational break before the PO_n is due to both (46a) and (46b), than when the intonational break before the PO_n is due to only (46a) or (46b). Then, a violation of the Prosodic Weight Condition in C-order is never more severe than that in M-order, because the SO-PO Mapping Condition on English PPs in (46a) is specific to the right edge of PPs, and therefore only relevant to M-order.

In the shifted sentence in (41b), the alignment constraint in (46b) is irrelevant because the constraint is vacuously satisfied in broad-focus sentences. The rightmost DP is separated from the preceding PP due to only one alignment constraint, the Mapping Condition on English PPs in (46a), which applies to any shifted sentences. Therefore, the Cumulativity Measure in (47) correctly formalizes the observation that a violation of the Prosodic Weight Condition in the shifted sentence (41b) is not the most severe, but only degrades the sentence.

Let us move onto the derivation of narrow-focus sentences. In narrow-focus sentences, M-order is ill-formed when the DP is not (a part of) a narrow-focus, as in the PP-focus sentences in (43) and in the V+PP-focus sentences in (45). On the other hand, M-order is available for some speakers when the DP is (a part of) a narrow-focus. This is exemplified by the DP-focus sentences in (42) and the V+DP-focus sentences in (44).

Look at the derivations of DP-focus sentences given in (48) below. Both the non-shifted sentence in (42a) and the shifted sentence in (42b) consist of two IntPs, and hence spell-out applies twice in each case. However, an intonational break is inserted for different reasons: in the non-shifted sentence in (42a), the right edge of the DP with [Focus]_SO is aligned with the right edge of an IntP, as is required by the alignment constraint for [Focus] in (46b), whereas in the shifted sentence in (42b), the right edge of the PP is aligned with the right edge of an IntP due to the SO-PO Mapping Condition on English PPs in (46a).

In (42a), the Prosodic Weight Condition is violated because the PO_n ("to-the-library") contains fewer prosodic words than the PO_{n+1} ("she-donated, her-collection, of-novels, by-MISHIMA"), resulting in the variable acceptability of the sentence. In contrast, in (42b), the Prosodic Weight Condition is satisfied because the PO_n ("her-collection, of-novels, by-MISHIMA") contains more prosodic words than the PO_{n+1} ("she-donated to-the-library") and also the prosodic word with extra prosodic prominence 'MISHIMA'. This is the case where word order optionality arises.

(48) Derivation of the examples in (42): single DP[theme]-focus

a. C-order \(\sqrt[\#](42a)\): [Imp she donated her_{collection of novels by MISHIMA}]{imp to the library}
Note in passing that in the shifted sentence in (42b), focal stress on the N ‘MISHIMA’ should not be necessary because the DP-focus interpretation can be achieved by focus projection from the syntactic feature [Focus] on the N in the default sentence stress position. However, the result of the pronunciation task shows that the N carries focal stress (as we will see in 3.1.4.1).

Next, consider the PP-focus sentences in (43). Their derivations are shown in (49) below. The non-shifted sentence in (43a) consists of one IntP, because the PP with [Focus] is in sentence-final position and the right edge of the SO corresponds to the right edge of the whole sentence. On the other hand, the shifted sentence in (43b) consists of two IntPs, because the right edge of the PP is aligned with the right edge of an IntP. An intonational break between the PP and the DP is due to the two alignment constraints in (46): by the PP-specific condition in (46a), and by the SO-PO mapping constraint for [Focus] in (46b), which requires the right edge of the PP with [Focus] to be aligned with the right edge of an IntP. In this case, the shifted sentence violates the Prosodic Weight Condition, because the PO (`them') contains less prosodic words than the PO (`she-donated to-the-Asian, LIBRARY'). The

---

6 The syntactic feature [Focus] may be freely introduced in the Numeration and attached to any syntactic object. (See the definition of [Focus] in 2.4.1.1.) In the derivations presented hereafter, the feature is shown after a focused element in the Numeration, but this is only for the purpose of illustrating the derivation of a sentence with a particular focus interpretation.
violation results in the severe ill-formedness, as is predicted by the Cumulativity Measure in (47), because the intonational break between $PO_n$ and $PO_{n+1}$ is due to two alignment constraints.

(49) Derivation of the examples in (43): single PP[goal]-focus

a. C-order (43a): [\text{she donated them to the Asian LIBRARY}]

Numeration \{she, donated, them, to, the, Asian, library $[F]_30$\}

syntactic structure

\[
\begin{array}{c}
\text{[she donated]} \\
\Rightarrow \\
\text{[she [donated [them] [PP[goal] to the library]]]}
\end{array}
\]

prosodic structure

\[
\begin{array}{c}
\text{PO}_n \\
[\text{she-donated-them-to-the-Asian-LIBRARY}$[F]_30$] \\
\text{SO-PO mapping} \\
\downarrow \\
\text{spell-out}
\end{array}
\]

b. M-order *(43b): [\text{she donated to the Asian LIBRARY}]_30 [\text{them}]

Numeration \{she, donated, them, to, the, Asian, library $[F]_30$\}

syntactic structure

\[
\begin{array}{c}
\text{[she [donated [PP[goal] to the library]]]} \\
\Rightarrow \\
\text{[she-donated-to-the-Asian-LIBRARY}$[F]_30$]
\end{array}
\]

prosodic structure

\[
\begin{array}{c}
\text{PO}_{n+1} \\
[\text{she-donated-to-the-Asian-LIBRARY}$[F]_30$] \\
\downarrow \\
\text{(PO}_n \text{ is lighter than PO}_{n+1})
\end{array}
\]

Let us move on to double narrow-focus cases. First, consider the V+DP-focus sentences in (44). Their derivations are shown in (50) below. In this case, both the non-shifted sentence in (44a) and the shifted sentence in (44b) consist of three IntPs. In the non-shifted sentence in (44a), an intonational break is inserted at two places after each SO with [Focus]$_{32}$, after the V 'DONATED' and the N 'MISHIMA', due to the alignment constraint for [Focus] in (46b). The Prosodic Weight Condition is violated because the rightmost IntP (= PO$_n$) 'to-the-library' dominates only one prosodic word and is lighter than the PO$_{n+1}$ 'her-collection, of-novels, by-MISHIMA'. As a result, there arise non-adjacent occurrences of prosodic prominence on the V and the N 'MISHIMA', and the non-shifted sentence is subject to the variable acceptability.

In the shifted sentence in (44b), an intonational break is inserted after the V with [Focus]$_{32}$, due to the alignment constraint for [Focus] in (46b), and another intonational break is inserted between the PP and the DP due to the SO-PO Mapping Condition on English PPs in (46a). The Prosodic Weight Condition is satisfied because the PO$_n$ ('her-collection, of-novels, by-MISHIMA') contains more prosodic
words than the $PO_n$ (‘to-the-library’) and also the prosodic word with extra prosodic prominence ‘MISHIMA’. This is another case where word order optionality arises.

(50) Derivation of the examples in (44): double $V+DP$-focus

a. C-order $\sqrt[1]{#(44a)}$: $[\text{[she DONATED]}[\text{her collection of novels by MISHIMA}][\text{to the library}]]$

b. M-order $\sqrt[2]{#(44b)}$: $[\text{[she DONATED]}[\text{her collection of novels by MISHIMA}][\text{to the library}]]$

Finally, consider the $V+PP$-focus sentences in (45). The derivation of the non-shifted sentence in (45a) is shown in (51a) below. An intonational break is inserted at two places after each SO with [Focus]$_{SO}$, the V ‘DONATED’ and the N ‘LIBRARY’, due to the alignment constraint for [Focus] in (46b). Since the second break is at the sentence-final position, the whole sentence is divided into two IntPs, not three. This sentence satisfies the Prosodic Weight Condition because the rightmost IntP (= $PO_n$), ‘them-to-the-Asian, LIBRARY’, contains more prosodic words than the $PO_{n-1}$, ‘she-DONATED’.

The derivation of the shifted sentence in (45b) is shown in (51b) below. An intonational break is inserted at two places after each SO with [Focus]$_{SO}$, the V ‘DONATED’ and the N ‘LIBRARY’, due to the alignment constraint for [Focus] in (46b). The second intonational break is also due to the SO-PO
Mapping Condition on English PPs in (46a), which requires the right edge of the PP to be aligned with the right edge of an IntP. The shifted sentence violates the Prosodic Weight Condition because the PO_n ('them') contains less prosodic words than the PO_n+1 ('to-the-Asian, LIBRARY'). The violation results in the severe ill-formedness of the sentence, as is predicted by the Cumulativity Measure.

(51) Derivation of the examples in (45): double V+PP-focus
a. C-order (45a): IntP she donated [§DONATED] IntP them to the Asian LIBRARY

Numeration {she, donated [F]SO, them, to, the, Asian, library [F]SO}

\[\text{syntactic structure} \]

\[\text{prosodic structure} \]

\[\text{PO}_{n-1} \]

\[\text{spell-out} \]

b. M-order *(45b): IntP she DONATED IntP to the Asian LIBRARY IntP them.

Numeration {she, donated [F]SO, them, to, the, Asian, library [F]SO}

\[\text{syntactic structure} \]

\[\text{prosodic structure} \]

\[\text{PO}_{n-1} \]

\[\text{PO}_n \]

(PO_n is lighter than PO_{n-1}.)

Note here that the non-shifted sentence in (45a) is degraded when the V carries focal stress compared to the case when it does not. I will discuss this issue as an unexpected result of the judgment task in the next section.

Thus, the Prosodic Weight Condition in (34), together with the Cumulativity Measure in (47), account for the distribution of C-order and M-order in English heavy NP shift.

3.1.4. Residue: Unexpected Results of the Elicitation

The key data dealt with in this chapter are results of a systematic elicitation I conducted in order to examine the interface properties of the heavy NP shift alternation in English. I have shown representative
examples in (41)-(45), and illustrated in (48)-(51) how the Prosodic Phase Hypothesis accounts for their acceptability in terms of the Prosodic Weight Condition in (34) and the SO-PO mapping constraints in (46). In this section, I discuss unexpected results of the pronunciation task and the judgment task of the elicitation. (More details of the elicitation are found in appendix to this chapter.)

3.1.4.1. Pronunciation Task

The major findings of the pronunciation task are summarized in (52):

(52) Major findings of the pronunciation task
i. Finding predicted by the Prosodic Phase Hypothesis
   Default sentence stress falls on the rightmost prosodic word in the rightmost Intonational Phrase (= PO").
   ⇒ Default Sentence Stress Assignment in (6a)
ii. Findings unexpected by the Prosodic Phase Hypothesis
   a. Extra prosodic prominence (e.g. focal stress), not normal lexical stress, is employed for a narrow-focus in the default sentence stress position.
   b. Single focal stress is used for double narrow-focus.

First, in broad-focus sentences, default sentence stress (i.e. the rightmost lexical stress in a sentence) is always found in the rightmost Intonational Phrase (IntP), specifically, on the rightmost prosodic word (= finding in (52i)). This supports the prosodically based formulation of the Default Sentence Stress Assignment in English in (6a).

The unexpected findings in (52ii) have to do with narrow-focus sentences. First, let us consider the finding in (52ii-a), which is related to single narrow-focus cases. If a narrow-focus is on the rightmost syntactic phrase, e.g. PP in the canonical V - DP - PP order, and DP in the marked V - PP - DP order, the focused phrase is in the default sentence stress position. In such cases, two ways of encoding disanaphoricity should be available in principle: one is to assign default sentence stress to the rightmost prosodic word and achieve a narrow-focus interpretation by focus projection (as defined in (9)); the other is to assign focal stress to the narrowly focused phrase. The former is regarded as an instance of partial broad-focus, and the latter, as single narrow-focus. They encode the same disanaphoricity pattern, but differ in how they do so prosodically. The result of the pronunciation task suggests that the narrowly focused phrase should carry focal stress in such cases.

(53) Single narrow-focus: PP-focus in C-order and DP-focus in M-order
a. C-order
   \[ [\text{IntP V - DP - PP}] \]
   extra prominence
   e.g. [\text{donated them to the Asian LIBRARY}]

b. M-order
   \[ [\text{IntP V - PP - DP}] \]
   extra prominence
   e.g. [\text{donated to the library}[\text{her... MISHIMA}]]

Remember that default sentence stress in English is not necessarily the most prominent prosodic peak in a sentence (cf. 3.1.2.1). Therefore, I argue that English resorts to extra prosodic prominence (e.g. focal stress) in order to prosodically distinguish narrow-focus from (partial) broad-focus. I formalize this as a descriptive generalization in (54):

(54) Prosodic encoding of disanaphoricity in English when narrow-focus = partial broad-focus
When a narrow-focus is in the default sentence stress position, English employs focal stress rather than default sentence stress, in order to prosodically encode disanaphoricity.
Therefore, English always uses extra prosodic prominence (e.g. focal stress) in narrow-focus sentences. Note that the generalization in (54) is inconsistent with Neeleman and Reinhart’s (1998) economy condition:

(55) Economy entails that stress strengthening applies only to derive foci not already in the focus set. (Neeleman and Reinhart 1998:340)

My speculation regarding (54) versus (55) is that the economy condition in (55) concerns production-based economy whereas the generalization in (54) suggests perception-based economy: extra prosodic prominence (or “stress strengthening” in (55)) requires extra effort in articulation whereas it gives perceptual cue for a narrow-focus. Neeleman and Reinhart’s economy condition is based on the examination of Dutch scrambling. In 4.1.3.1, I will show that the prosodic property associated with Japanese short-scrambling is consistent with production-based economy in (55) rather than perception-based economy in (54).

Let us move on to another unexpected finding in (52ii-b). When two foci are non-adjacent as with the canonical V - DP - PP order in V+PP-focus sentences and the marked V - PP - DP order in V+DP-focus sentences, the prosodic pattern is similar to single PP-focus and single DP-focus sentences, respectively. That is, single focal stress on the focused phrase is sufficient and the V does not receive focal stress. However, having non-adjacent prosodic prominence on the V and the DP is judged to be also possible for some speakers in M-order (but not in C-order). I will come back to this shortly in the next section.

(56) Non-adjacent double narrow-focus: V + PP-focus in C-order and V + DP-focus in M-order
   a. C-order
      \[\text{[\text{[\text{DP} - \text{PP}] \uparrow}] \text{V}}\]
      extra prominence
      e.g. \[[\text{donated}][\text{them}][\text{to... LIBRARY}]\]
   b. M-order
      \[\text{[\text{[\text{PP} \uparrow}] - \text{[\text{DP} \uparrow}] \text{V}}\]
      (extra prominence) extra prominence
      e.g. \[[\text{donated}][\text{to the library}][\text{...MISHIMA}]\]

Note that in (56), focus projection from the default sentence stress position fails to achieve the intended focus interpretation, because there is no syntactic node that exhaustively dominates the non-adjacent focused elements, V and PP in (56a) and V and DP in (56b).

Therefore, it seems that English can encode different disanaphoricity patterns by using only one occurrence of prosodic prominence. I argue that this is possible because English can make use of mobile prosodic prominence, e.g. accenting and deaccenting, in encoding disanaphoricity. The question is how. I propose a sort of LF scrambling, which removes an anaphoric argument from the focus domain:

(57) LF scrambling of anaphoric argument

\[\text{DP} \quad \text{VP \{F\}_\text{so} \Rightarrow focus domain} \]
\[\text{V} \quad \text{focus projection} \]
\[t_{\text{DP}} \quad \text{PP\{F\}_\text{so}} \]
\[\text{N} \quad \text{focal stress} \]
\[\text{e.g. She \{[\text{donated}][\text{them}][\text{to the Asian LIBRARY}]. [=\{45a\}]\}}\]
In this case, the V can be interpreted as focused by virtue of focus projection of [Focus] from the PP (that dominates the default sentence stress position) up to the VP. At LF, the anaphoric DP ‘them’ raises to a position higher than the VP, escaping the focus domain (i.e. VP). This sort of movement is independently proposed in accounting for Dutch scrambling (Reinhart 1996, Neeleman and Reinhart 1998). The difference is that it takes place at LF in English and before spell-out in Dutch. The DP is prosodically a stressless pronoun, and perceived as deaccented relative to the neighboring elements at PF.

In (57), default sentence stress is realized as extra prosodic prominence. Therefore, whether default sentence stress is prosodically distinct or not, the same type of focus projection should be available for M-order as well, since C-order and M-order have the same right-branching structure. The availability of a VP-focus interpretation in the shifted sentence in (58B) supports this point:

(58) A: What did John do?
B: He [VP gave [PP to Mary] [DP all of the money in the SATCHEL]]. (Williams 2003:34)

The question sentence in (58A) requires an answer with a VP-focus. The felicitous answer in (58B) exhibits the marked V – PP – DP order, and the DP contains the prosodic word with extra prosodic prominence ‘SATCHEL’. Here, the entire VP inherits [Focus] by focus projection from the rightmost node DP dominating the N ‘SATCHEL’ in the default sentence stress position, and neither partial deaccenting nor LF scrambling takes place.

3.1.4.2. Judgment Task

The pronunciation task examined how the sentences were pronounced only, and did not address the issue of whether the sentence was acceptable or not. Let us now examine the results of the judgment task. The major findings of the task are summarized in (59):

(59) Major findings of the judgment task
i. Finding predicted by the Prosodic Phase Hypothesis
   M-order is acceptable for some speakers if the rightmost DP is (a part of) a narrow-focus (e.g. DP-focus, V+DP-focus) and ill-formed otherwise.
   ⇒ the generalization in (2)
ii. Findings unexpected by the Prosodic Phase Hypothesis
   a. M-order is not always acceptable even if the rightmost DP is (a part of) a narrow-focus.
   b. Non-adjacent focal stresses are better with M-order than with C-order.

First of all, the results of the judgment task show that M-order is acceptable for some speakers when the rightmost DP is (a part of) a narrow-focus (= finding in (59i)). If the DP is a narrow-focus, it carries focal stress and hence satisfies the Prosodic Weight Condition. Thus, the results of the judgment task support the central claim of the Prosodic Phase Hypothesis that the heavy NP shift alternation in English is prosodically conditioned.

However, the choice between C-order and M-order was sometimes not as clear as I expected it to be, and some subjects found C-order as natural as or more natural than the corresponding M-order, even in the case where the DP carries focal stress (= finding in (59i-a)). Look at the relevant DP-focus sentences repeated in (60) below. Four out of five subjects found the non-shifted sentence in (B1) better than the corresponding shifted sentence in (B2), and one subject had the opposite judgment.

(60) DP-focus [= (42)]
A: What did Kay donate to the Asian library?
C-order
B1: \[\text{She donated her collection of novels by MISHIMA} \text{ to the library].}\]

M-order
B2: \[\text{She donated the library her collection of novels by MISHIMA].}\]

Notice here that a definite description is used for the anaphoric PP (‘to the library’) in (60). When the definite description is replaced by the deaccented pronominal ‘them’, M-order is preferred over C-order, as is shown in the follow-up example in (61):

(61) DP-focus
A: What did Kay donate to the Asian library?
C-order
B1: \[\text{She donated her collection of novels by MISHIMA} \text{ to them].}\]
M-order
B2: \[\text{She} \text{ donated to them her collection of novels by MISHIMA}.\]

A native speaker of English mentioned that the non-shifted sentence in (61B1) was “weird because there are unstressed items, ‘to them’, at the end of the sentence”, which is reminiscent of the “widespread conspiracy against VP-final weak pronoun” (Erteschik-Shir and Strahov 2004:315-316) in English. I speculate that the degraded acceptability of the non-shifted sentence in (61B1) is due to a prosodic tension between the focal stress on ‘MISHIMA’ which induces an intonational break and the stressless DP ‘to them’ which needs to cliticize onto the preceding word ‘MISHIMA’.7 I defer further investigation of this issue for the moment until I have examined a larger number of examples.

Finally, let us consider another unexpected finding of the judgment task in (59ii-b): non-adjacent occurrences of prosodic prominence were judged to be more natural with M-order than with C-order. Look at the relevant shifted sentences in (62). In (B1), only the rightmost DP carries focal stress, whereas in (B2) the V as well as the DP carry focal stress.

(62) V + DP-focus
A: UBC Asian library is facing a serious financial problem. What did Kay do to save the library?
B1: She donated to the library her collection of novels by MISHIMA.
B2: She DONATED to the library her collection of novels by MISHIMA.

Although (62B1) reflects the prosodic pattern volunteered in the pronunciation task, native speakers showed no consistent preference for (B1) over (B2) in the judgment task. Furthermore, extra prosodic prominence on a V is not only felicitous but necessary if the V is contrastively focused as in (63) (cf. Henry Davis, Rose-Marie Déchaîne, p.c. 2003):

(63) A: UBC Asian library is facing a serious financial problem. I think that Kay sold her collection of videos by Kurosawa very cheap to save the library, right?
B: No. She #donated/DONATED to them her collection of novels by MISHIMA.

The point is that non-adjacent focal stresses are possible in English, which follows from the ability of English to tolerate partial deaccenting in the middle of a sentence (see the mechanism of partial deaccenting in (57)).

7 Two speakers noted that the preference for C-order over M-order in (60) may be due to the parallelism between the question sentence in (60A) and the non-shifted sentence in (60B1), in that both have a PP in sentence-final position. This, however, does not predict the opposite word order preference in (61).
Note further that in C-order as well as in M-order, the effect of deaccenting in (64) is vacuous, because a prosodic word with focal stress is always at the right edge of an IntP due to the SO-PO mapping constraint for [Focus] in (46b) in English:

(64) Deaccenting [= (58) in 2.4.3.1]
Any elements that follow a prosodic word with extra prosodic prominence are deaccented within an Intonational Phrase.
(65) Non-adjacent focal stresses

\[
\begin{align*}
\text{a. C-order: } & V - DP - PP \\
& \text{e.g. } [DONATED][them to the Asian LIBRARY] \\
\text{b. M-order: } & V - PP - DP \\
& \text{e.g. } [DONATED][to...][MISHIMA]
\end{align*}
\]

Since the deaccenting phenomenon is irrelevant to both C-order in (65a) and M-order in (65b), the question remains why non-adjacent focal stresses are better with M-order than with C-order. I speculate that the problem with C-order in (65a) is the intonational break that divides the V and the pronominal DP 'them'. Being prosodically weak, the pronoun 'them' needs to cliticize onto an adjacent word. In (65a), the pronoun cliticizes onto the following PP, forming the prosodic word 'them-to-the-Asian', because the intonational break prevents it from cliticizing onto the preceding V. I argue here that prosodically weak dependents in English can only cliticize onto the preceding word (if any), but not onto the following word, possibly due to the head-initiality of the language. This condition is violated in C-order, but not in M-order where the PP dependent 'to-them' does not cliticize onto any word. If correct, the unexpected findings in (59ii) are both analyzed prosodically, which provides further supporting evidence for the Prosodic Phase Hypothesis.

3.2. More on the Distinction between Canonical and Marked Orders

Section 3.1 showed that there is a prosodic markedness effect in the heavy NP shift alternation in English: the shifted \( V - PP - DP \) order is more marked than the non-shifted \( V - DP - PP \) order, in that the rightmost DP forms an independent Intonational Phrase (IntP) satisfying the Prosodic Weight Condition. This section considers the nature of the Prosodic Weight Condition, and speculates on the possibility that a right-edge heaviness effect reflects processing efficiency in the performance systems.

3.2.1. The Prosodic Weight Condition and Processing Efficiency

English makes use of different degrees of prosodic prominence in encoding disanaphoricity. In narrow-focus sentences, for example, anaphoric elements are often pronominalized and deaccented whereas focused elements receive extra prosodic prominence. Because of this mobility of prosodic prominence, English does not have to employ word order alternations in order to encode a particular disanaphoricity pattern in a sentence. The assumption behind this idea is that word order alternations are more costly than moving prosodic prominence in English.

I argued that the marked \( V - PP - DP \) order is available only when the rightmost DP is a prosodically legitimate Intonational Phrase (IntP), satisfying the Prosodic Weight Condition repeated below.

---

\[\text{The intuition behind this is that the prosodic word 'them-to-the-Asian', where the pronominal 'them' cliticized onto the following PP, is semantically incoherent and hence degraded. However, assuming that PF and LF do not directly interact with each other, it is not possible to resort to a semantic condition to rule out such prosodic words. (I thank Scott Shank for discussion on this.)}\]
(34) **The Prosodic Weight Condition**

PO₀ (= the rightmost Intonational Phrase) must be prosodically heavy by containing
a. a larger number of prosodic words than PO₀₋₁ (= the second rightmost
   Intonational Phrase), or
b. a prosodic word with extra prosodic prominence.

That is to say, English heavy NP shift is prosodically conditioned, and focus in itself only indirectly
affects the alternation, in that focused elements in English may affect the prosodic pattern of a sentence
when they carry focal stress.

A remaining question is why the Prosodic Weight Condition licenses the marked V – PP – DP
order. Following Hawkins (1990, 1994 and subsequent work), I speculate that in English, linearization
of verbal dependents reflects the PF interface requirement that elements in the VP must appear in an order of
increasing weight for processing efficiency, achieving right-edge heaviness (or “end-weight” in Quirk et
al. 1985:1398). In particular, a DP dependent cannot intervene between the V and the PP dependent when
the DP is prosodically heavier than the PP by a certain degree. This is because if a prosodically heavy DP
intervenes between the head V and the PP, the time elapsed between the head and the PP may become too
big to establish semantic association between them.

Let us look at the heavy NP shift alternation given in (66) below to see how this reasoning works.
In the shifted sentence in (66a), when the parser processes the prosodic word ‘five’, which signals the
theme DP of the V ‘donated’, the parser needs to retrieve the information of the associated head V, which
is two prosodic words back. In the corresponding non-shifted sentence in (66b), when the P ‘to’ which
signals the goal PP is processed, the parser needs to look back as many as eight prosodic words in looking
for the associated V.

(66) M-order
a. Kay donated \[PP \text{to-the-library}] [DP five hundred Canadian dollars and-her-
   \[\text{collection of-novels by-Mishima}]. \[= (15B1)]
>

C-order
b. Kay donated \[DP five hundred Canadian dollars and-her-collection of-novels
   \[\text{by-Mishima}][PP \text{to-the-library}]. \[= (15B2)]

Consideration of performance efficiency predicts that M-order is chosen over C-order in (66), because M-
order makes semantic association of verbal dependents with the V earlier and hence easier.

Hawkins’ efficiency-based explanation of linearization also makes a prediction that the DP must
be relatively heavier than the preceding IntP for M-order to be chosen (Hawkins 1990, 1994, Shiobara
2000, 2001, see also the table in (68) in 3.2.2). For example, C-order is preferred over M-order when the
goal PP contains the same number of prosodic words as the DP:

(67) C-order
a. Kay donated \[DP five hundred Canadian dollars and-her-collection of-novels
   \[by-Mishima][PP \text{to-the-UBC Asian library located near-the-Nitobe Japanese}
   \text{garden}].

M-order
b. #Kay donated \[PP \text{to-the-UBC Asian library located near-the-Nitobe Japanese}
   \text{garden}][DP five hundred Canadian dollars and-her-collection of-novels by-Mishima].

When the PP contains seven prosodic words as in the example (67), M-order is not acceptable any more
although the DP in the sentence is the same as that in (66). In the shifted sentence in (67b), the Prosodic
Weight Condition is violated because the DP ('five, hundred,... of-novels, by-Mishima'), which is the rightmost IntP, is lighter than the preceding IntP ('Kay, donated,... Japanese, garden') by two prosodic words. In this case, C-order is chosen as the elsewhere case, and the non-shifted sentence in (67a) is likely to be divided into smaller IntPs when pronounced. (See 2.4.2.2 and 2.5.1 for related examples.)

3.2.2. The Prosodic Markedness Effect and Performance Frequency

The shifted V – PP – DP order is prosodically more marked than the non-shifted V – DP – PP order, in that the shifted order must satisfy the Prosodic Weight Condition. The markedness effect in the heavy NP shift alternation in English is independently observed with regards to the performance data examined by Hawkins (1994). Hawkins counts the number of the sentences with V – DP – PP order and the sentences with V – PP – DP order in corpus texts, and compares them with respect to the relative heaviness of the DP and the PP. The result of Hawkins’ text counts is shown in the table (68):

(68) English heavy NP shift (based on Hawkins 1994:183, Table 4.22)

<table>
<thead>
<tr>
<th></th>
<th>V – DP – PP order (C-order)</th>
<th>V – PP – DP order (M-order)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>458</td>
<td>22</td>
</tr>
<tr>
<td>n=131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-order</td>
<td>68</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>M-order</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>PP&gt;DP: by 1-2 words</th>
<th>3-7</th>
<th>8+</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-order</td>
<td>209</td>
<td>103</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>M-order</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

The total number of the token is 480, among which C-order is 458 and M-order is 22. From this, Hawkins draws the first generalization that C-order is significantly more frequent than M-order in English VPs, which is consistent with the marked status of M-order.

Among 480 tokens, DP is heavier than or as heavy as PP (with respect to the number of syntactic words) in 131 cases, and PP is heavier than DP in 349 cases. C-order is still productively maintained even when DP is heavier than PP by up to four words. In particular, when DP is as heavy as or heavier than PP by one word, the sentences are all C-order (n = 68+25). When DP is heavier than PP by more than four words, the sentences are all M-order (n = 9). On the other hand, when PP is heavier than DP, only three sentences exhibit M-order and the others are all C-order (n = 209+103+34). Hawkins regards this observation as another evidence for the non-shifted V – DP – PP order as canonical.

Thus, the data on performance frequency of C-order and M-order is also consistent with the marked status of M-order.

3.3. Deriving the Properties of Marked Order

Previous studies on English heavy NP shift have explained various properties associated with the alternation, such as weight of the sentence-final DP, locality, the syntactic position of the sentence-final

---

9 Although Hawkins' research emphasis is on the effects of syntactic weight (i.e. the number of syntactic words) on word order, the main point of his parsing-based explanation still holds if we reinterpret syntactic weight as prosodic weight.

10 This observation leads Hawkins to claim that the canonical DP-PP order is an instance of diachronic grammaticalization of the observational data, where a heavy elements comes after a light element in head-initial languages. Hawkins bases this claim on the argument that PPs are heavier than DPs syntactically by containing a preposition, i.e. PP = P + DP. Hawkins' claim, however, is not necessarily valid in the context of the Prosodic Phase Hypothesis, where it is the number of prosodic words, not syntactic words, that determines the weight. In particular, a prosodically weak preposition does not form a prosodic word on its own, and hence DP and PP are of the same weight prosodically.
DP, binding effects, and so on (cf. Ross 1967, Larson 1989, Hawkins 1990, Rochemont and Culicover 1990 among others). This section shows that some properties of shifted sentences with the marked order are prosodic in nature and well accounted for by the Prosodic Weight Condition (3.3.1), and other syntactic properties can be explained in terms of the incremental structure-building mechanism that yields right-branching structure in the syntax (3.3.2).

3.3.1. Arguments for the Prosodic Weight Condition

In 3.3.1.1, I argue that the Prosodic Weight Condition is not specific to heavy NP shift, but applies to XP-shift in general and also to other alternations in the nominal domain. In particular, I show that a clausal (CP/IP) dependent always follows another dependent because it forms an independent Intonational Phrase (IntP). I also show that linear order between non-DP dependents is purely prosodically determined by the Prosodic Weight Condition. In 3.3.1.2, I argue that locality between a head verb and its dependent in part follows from the Prosodic Weight Condition.

3.3.1.1. More on the Right-Edge Heaviness Effect in XP-Shift

So far, our focus has been on the heavy NP shift alternation between the canonical V - DP - PP order and the marked V - PP - DP order. The Prosodic Phase Hypothesis explains linearization of the two dependents, DP and PP, with respect to the difference in the prosodic pattern between the canonical order (C-order) and the marked order (M-order): C-order is associated with the default intonation pattern, whereas in M-order, the rightmost DP must form a prosodically legitimate IntP on its own. This section examines the prosodic properties of XP-shift alternations in general, and shows that they provide supporting evidence for the Prosodic Phase Hypothesis.

Recall the constraints on the syntax-prosody mapping relevant to English heavy NP shift:

(69) a. SO-PO mapping for maximal prosodic objects (= Intonational Phrases) [= (5a)]
   English: Align (SO, R; PO, R) and Align (SO, L; PO, L),
   where SO = clause (CP/IP).
   b. The SO-PO Mapping Condition on English PPs [= (18)]
   Align (SO, R; PO, R), where SO = a PP dependent, and PO = Intonational Phrase.

The SO-PO mapping constraint for maximal prosodic objects in (69a) specifies that a non-shifted sentence as a whole corresponds to one IntP (unless it contains an embedded clause). The SO-PO Mapping Condition on English PPs in (69b) derives the prosodic property of shifted sentences in English, where the sentence-final DP forms an IntP on its own. The shifted order (i.e. M-order) is available when it satisfies the Prosodic Weight Condition:

(70) The Prosodic Weight Condition [= (34)]
   PO_n (= the rightmost Intonational Phrase) must be prosodically heavy by containing
   a. a larger number of prosodic words than PO_n+1 (= the second rightmost
      Intonational Phrase), or
   b. a prosodic word with extra prosodic prominence.

Based on the SO-PO mapping constraints in (69), I make the predictions given in (71) below, regarding the category of the phrases that participate in word order alternations. The first prediction in (71a) is that a clausal dependent, whether finite (CP) or infinite (IP), should always follow another dependent if any. This is because a clause corresponds to an independent IntP in the default case in English. (See also Hawkins 1994:294-295 for the positioning of clausal dependents.)

The second prediction in (71b) is based on the generalization captured by the Mapping Condition on English PPs in (69b) in the following sense. The canonical V - DP - PP order does not need a prosodic
motivation because it is canonical in that the Case Adjacency requirement between the V and the DP is satisfied. The flip side of this is that when XP is not a DP and does not require a Case, it is predicted that linear order between two dependents should be purely prosodically determined, and the two dependents appear in an order of increasing weight to achieve right-edge heaviness.11

(71) Predictions for XP-shift
   a. When XP = CP/IP, it follows another dependent.
   b. When XP = non-DP, linear order between two dependents is purely determined by the Prosodic Weight Condition.

In the following sections, I show the predictions in (71) are borne out. I also discuss DP-shift with preposition-stranding, and shows that it also supports a prosodically based analysis of XP-shift.

3.3.1.1.1. CP/IP-Shift

First, I show that CP/IP-shift examples bear out the prediction in (71a). In the examples in (72)-(73) below, the (a)-(c) examples were elicited by a linguist or cited from newspaper articles whereas the corresponding (a')-(c') examples were created for the purpose of comparison and tested with native speakers. A finite CP or a non-finite IP prefers appearing at the right edge of the whole sentence following another dependent as in (a)-(c), rather than appearing before another dependent as in (a')-(c'). (In the (a)-(c) examples, the position where a CP/IP would appear in the alternative order is indicated by underscore.)

(72) CP-shift in the verbal domain
   a. He said _ several times [CP that I would get the books I ordered by the end of the week].
      (Stucky 1987:384)
   >
   a'. He said [CP that I would get the books I ordered by the end of the week] several times.
   b. North Korea's ambassador to Russia said _ on Thursday [CP his country supports multilateral talks, including Russia, to ease tensions surrounding Pyongyang's nuclear program]...
      (Vancouver Sun, 1 August, 2003)
   >
   b'. North Korea's ambassador to Russia said [CP his country supports multilateral talks, including Russia, to ease tensions surrounding Pyongyang's nuclear program] on Thursday...
   c. Minnesota Public Radio noted _ in its state wide broadcast, [CP that in a city so badly hit by the SARS scare, the concert was "a recovery strategy that might well be described as 'party economics!'"] (ibid.)
   >
   c'. Minnesota Public Radio noted [CP that in a city so badly hit by the SARS scare, the concert was "a recovery strategy that might well be described as 'party economics!'"] in its state wide broadcast.

(73) IP-shift in the verbal domain
   a. They tried _ for many times [IP to force him to make a decision one way or the

11 The prosodically based analysis of the heavy NP shift alternation may be extended to apply to the dative alternation between V - DP1 - PP2 and V - DP2 - DP1 as in (i) (in V - DP2 - DP1 order, the sentence-final DP1 cannot be a weak pronoun, as is shown in (ii) (although there is a dialectal variation and (ii) may be acceptable in British English, cf. Henry Davis, p.c. 2004).

(i) a. Mary gave [DP1 the report on Dukakis] [PP2 to Sue].
   b. Mary gave [PP2 Sue] [DP1 the report on Dukakis].
(ii) *Mary gave [PP2 Sue] [DP1 it].

This seems to be the same right-edge heaviness effect found in shifted sentences. (See Hawkins 1994:214 for the analysis of the dative alternation as a "grammaticalized" weight effect, cf. fn.10.)
a'. They tried [IP to force him to make a decision one way or the other] for many times but to no avail.

b. Meanwhile, only 11 per cent of children in the survey were expected [CP to take out repayable loans from family members or financial institutions when they attend post-secondary education] by their parents.

(Vancouver Sun, 21 November, 2003)

b'. Meanwhile, only 11 per cent of children in the survey were expected [IP to take out repayable loans from family members or financial institutions when they attend post-secondary education] by their parents.

c. I asked the man who worked in the library [IP to go].

c'. *I asked [IP to go] the man who worked in the library.

In particular, in the sentence (72c) cited from a newspaper article, the CP is separated from the preceding adverbial PP by a comma, which indicates that the CP should start a new intonational domain according to English orthography. Incidentally, the CPs and IPs in (72)-(73) contain a fairly large number of prosodic words, conforming to the Prosodic Weight Condition in (70). For example, in (72b) the CP contains twelve prosodic words ('his-country, supports, multilateral, talks, including, Russia, to-ease, tensions, surrounding, Pyongyang’s, nuclear, program'), and the preceding part contains only six prosodic words ('North, Korea’s, ambassador, to-Russia, said, on-Thursday'). However, a clause does not have to satisfy the Prosodic Weight Condition, because it is an IntP by the SO-PO mapping condition for IntPs in (69a). This is exemplified in (73c), where the IP consists of only one prosodic word ('to-go'). The tendency for a clause to appear at the right-edge is also found in alternations within the nominal domain. The (a)-(b) examples in (74) were spontaneously produced, whereas the corresponding (a')-(b') examples were created for the purpose of comparison, and then tested with native speakers.

(74) CP-shift in the nominal domain

a. ... there are important differences [CP that fall below the threshold of perceptibility]. (Ross 1986)

a'. ... there are important differences [CP that fall below the threshold of perceptibility] for the study of language.

b. ... she was a Canadian citizen [CP who was killed while in Iranian custody]. (Vancouver Sun, 21 November, 2003)

b'. ... she was a Canadian citizen [CP who was killed while in Iranian custody] with dual citizenship.

In (74a) and (74b), the head noun is followed by a PP modifier and a CP modifier in PP – CP order, whereby the CP modifier is non-adjacent to its antecedent N(P) and at the right edge. The alternative CP – PP order in (74a') and (74b') is judged to be less natural than the corresponding PP – CP order.

Regarding the prosodic property of non-adjacent relative clauses (such as those in (74a) and (74b)), Croft (1995:851) observes that non-adjacent relative clauses are much more likely to be an independent intonational domain than adjacent relative clauses (such as those in (74a') and (74b')). According to Croft’s examination of a corpus of English oral narratives, 23 out of 25 (92%) instances of non-adjoined relative clauses are independent intonational domains, whereas only 5 out of 88 (6%) embedded relative clauses form their own intonational domains (Croft 1995:851, see also Frazier and Clifton 1996:105 for the intonational property of non-adjacent relative clauses). This observation is consistent with the prosodic condition on M-order that the rightmost phrase must form its own IntP.
Moreover, the SO-PO mapping constraint in (69a) specifies that clauses are always IntPs regardless of where they appear in a sentence.

3.3.1.1.2. Non-DP-Shift

Next, let us look at linearization of non-DP dependents and see that the prediction in (71b) is also borne out. Hawkins (1994) shows that when two dependents are both PPs, the order between them is determined by the relative weight of them. According to the result of Hawkins' text counts shown in the table (75) below, when the difference in weight becomes bigger, the ratio of the order where a heavy PP follows a light PP increases; when a PP is heavier than another PP by more than five prosodic words, all examples conform to an increasing weight order:

(75) English alternation between V – PP1 – PP2 and V – PP2 – PP1 (based on Hawkins 1994:129, Table 4.2)

<table>
<thead>
<tr>
<th></th>
<th>PP1=PP2</th>
<th>PP2&gt;PP1</th>
<th>by 1 word</th>
<th>2-3</th>
<th>4</th>
<th>5+</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP1-PP2</td>
<td>35</td>
<td>27</td>
<td>30</td>
<td>10</td>
<td>41</td>
<td>=108</td>
</tr>
<tr>
<td>PP2-PP1</td>
<td>12</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>=18</td>
<td></td>
</tr>
</tbody>
</table>

The examples from newspaper articles given in (76) below also exhibit a right-edge heaviness effect. In (76a), the second PP ('by-various, agencies including, Translink, and-the-provincial, transportation, ministry') is prosodically heavier than the first PP ('for-several, years') by five prosodic words. In (76b), the second PP ('of-injuries, received, during, her-prison, interrogation') is prosodically heavier than the first PP ('in-hospital') by four prosodic words.

(76) Alternation between V – PP1 – PP2 and V – PP2 – PP1
a. The proposed bypass has been under discussion [PP1 for several years] [PP2 by various agencies including Translink and the provincial transportation ministry].
   (Vancouver Sun, 15 August, 2003)
b. Kazemi died [PP2 in hospital] [PP1 of injuries received during her prison interrogation].
   (Vancouver Sun, 21 November, 2003)

The alternation between a PP and an AdvP also seems to exhibit a right-edge heaviness effect:

a. John speaks [AdvP well] [PP to his mother].
b. *John speaks [PP to his mother] [AdvP well].
c. John speaks [PP to his mother] [AdvP WELL].
d. John speaks [AdvP to his mother] [very well]. (Costa 1997:65)

The contrast in (77a) and (77b) shows that the prosodically light AdvP 'well' must appear before the PP 'to his mother'. Costa (1997) points out that when the AdvP is heavily stressed as in (77c) or contains another prosodic word as in (77d), it may appear after the PP.

Costa (1997) argues that the alternation that involves two AdvP dependents also shows a right-edge heaviness effect:

(78) Alternation between V – AdvP1 – AdvP2 and V – AdvP2 – AdvP1
a. John spoke to his mother [AdvP yesterday] [AdvP nicely and carefully].
b. *John spoke to his mother [AdvP yesterday and the day before] [AdvP nicely].
c. John spoke to his mother [AdvP yesterday and the day before] [AdvP nicely and carefully].
   (Costa 1997:66)
The contrast in (78a) and (78b) shows that the AdvPs must appear in an increasing weight order: in (a), the second AdvP ('nice, and-carefully') is prosodically heavier than the first AdvP ('yesterday') by one prosodic word, and in (b), the first AdvP ('yesterday, and-the-day, before') is prosodically heavier than the second AdvP ('nicely') by two prosodic words. This suggests that word order between two AdvP dependents is prosodically conditioned. Furthermore, when the second AdvP is preceded by an intonational break and explicitly rendered the domain of an intonation contour as in (78c), it can appear after the other AdvP despite that the number of prosodic words contained in the second AdvP ('nicely, and-carefully') is not bigger than that of the preceding AdvP ('yesterday, and-the-day, before').

The PP-shift out of a NP called “extraposition from NP” (Ross 1967, Guéron 1980) also exhibits a right-edge heaviness effect. Look at the following example:

(79) A: Which book came out recently?
B1: A book came out yesterday [p by the Nobel Prize laureate Mr. Tanaka and his co-workers at Shimazu company].
B2: A book came out yesterday [p by Mr. Tanaka].
B3: A book came out yesterday [p by HIM]. (pointing at Mr. Tanaka over there)
B4: *A book came out yesterday [p by him]. (pointing at Mr. Tanaka over there)

In (79), the PP modifier of the head N ‘book’ appears in sentence-final position. When the PP contains eight prosodic words as in (79B1) ('by-the-Nobel, Prize, laureate, Mr, Tanaka, and-his-co-workers, at-Shimazu, company') and is prosodically heavier than the preceding part ('a-book, came-out, yesterday') by five prosodic words, the shifted sentence is acceptable. In contrast, when the PP contains only two prosodic words as in (79B2) ('by-Mr, Tanaka') or one prosodic word as in (79B3) ('by-HIM'), the PP is prosodically lighter than the preceding part of the sentence and the shifted sentence is degraded. When the PP modifier consists only of unstressed function words ('by-him') as in (79B4), the shifted sentence is completely unacceptable.  

We have seen that non-DP dependents also participate in word order alternations, and linearization of non-DP dependents is prosodically conditioned. In particular, when one of the dependents is a clause (CP/IP), it follows another dependent because a clause always forms an IntP on its own. We have also seen that alternations in the nominal domain exhibit a right-edge heaviness effect. This is predicted by the prosodically based analysis of shifted sentences, because XPs that participate in the alternation in the nominal domain are all non-DPs, and hence their ordering is purely prosodically determined.

3.3.1.1.3. DP-Shift with Preposition-Stranding

Finally let us look at the shift of DPs out of PPs. Although DPs can participate in word order alternations

12 Chomsky (2001) analyzes extraposition from NP as an instance of heavy NP shift, postulating additional deaccenting of the second occurrence of the DP (indicated by small letters):

(i) a. We saw a painting yesterday, (that is), a_painting[one] [from the museum].
   b. I gave him a painting yesterday, (more precisely), a_painting[one] [from John’s collection].

(Chomsky 2001:19)

According to Chomsky’s analysis, the second occurrence of the DP “painting” is deaccented and can undergo deletion, yielding the shifted order, and the phrase after the intonational break introduces “qualifications or afterthoughts” (p. 19). (See also Wilder 1996 for a copy-and-deletion analysis of extraposition from NP.) This indicates that the sentence-final phrase is prosodically more prominent than the preceding deaccented (and deleted) DP, which is consistent with the gist of the Prosodic Weight Condition. (See Shiobara 1997 for more discussion on the weight effect in extraposition from NP.)

However, Chomsky’s analysis is inconsistent with the agreed observation that English does not usually use null DP anaphora. Furthermore, it predicts that extraposed phrases are always non-restrictive or appositive, which is incorrect (cf. Guéron and May 1984). (Thanks to Henry Davis for drawing my attention to this last point.)
as is exemplified by the heavy NP shift alternation, DP-shift seems to be impossible when the DP is an object of a preposition. In other words, rightward preposition-stranding is prohibited (Postal 1986:207, Larson 1989, Rochemont and Culicover 1990):

(80) DP-Shift out of a PP
   a. *May put the money on _ yesterday [DP a table that was sitting at the entrance to the hall].
   b. *John threw a look at _ as he was walking by [DP a man who was standing outside his office].
   c. *I mailed a letter to _ on my way to work [DP old friend from high school].

   (Rochemont and Culicover 1990:135)

   While the ill-formedness of the examples in (80) is severe, rightward preposition-stranding does not always give rise to a severe unacceptability:

(81) DP-Shift out of a PP
   a. ?We slept in _ when we were in Connecticut [DP a marvelous bed that had belonged to George Washington].
   b. *We slept under _ when we were in Connecticut [DP a marvelous bed that had belonged to George Washington].

(82) a. ?We looked at _ last night [DP a wonderful film about New York that had been made during the Depression].
   b. *We talked during _ last night [DP a wonderful film about New York that had been made during the Depression].

   (Rochemont and Culicover 1990:191, fn.34)

In (81)-(82), the (a) sentences are more acceptable than the corresponding (b) sentences. The difference is found in the prosodic property of the preposition: mono-syllabic prepositions (e.g. ‘in’, ‘at’) allow preposition-stranding more freely than bi-syllabic prepositions (e.g. ‘under’, ‘during’), when the PP immediately follows the verb. Rochemont and Culicover (1990) speculate that a mono-syllabic preposition is reanalyzed as a part of the verb, and hence (81a) and (82a) do not involve preposition-stranding. I agree with their speculation, and assume that a mono-syllabic preposition cliticizes onto the preceding verb, forming a prosodic word with the verb. Going back to the ill-formed examples in (80), I assume here that a preposition cannot cliticize onto the preceding noun, because the PPs are not dependents of the noun, but dependents of the non-adjacent verb. Therefore, I speculate that the severe ill-formedness of the examples in (80) is due to the prosodic property of mono-syllabic prepositions, which need to cliticize onto the following DP but cannot because the DP is shifted.

The prosodically based analysis of rightward preposition-stranding is further supported by the verb-particle constructions such as (83)-(84), where a verb and a preposition are regarded as a lexically specified unit, and preposition-stranding is possible:

(83) a. Max [v looked up] _ in the dictionary [the word that Bill had asked him about].
   b. Felix [v threw out] _ with the trash [a manuscript that Oscar had been working on for years].
   c. They [v sealed off] _ without incident [the Blob’s only remaining exit from the cave].

   (Larson 1989:12)

(84) a. ... that would threaten a landowner’s right to [v rely on] _, in good faith, [a city’s representation].
   b. I’ll [v go over] _ in my mind [all the things I did wrong].
   c. ... dinosaurs help children [v work through] _ in making believe [the issues they are grappling with daily].

   (Wasow 1995, as cited in Nishihara 1997:260)
In sum, I assume that preposition-stranding is possible when the preposition is prosodically light or lexically a part of the verb. In such cases, M-order should be available to the extent that the DP satisfies the Prosodic Weight Condition, as is the case with DP-shift in general.

3.3.1.2. Locality of XP-Shift

There is an asymmetry between leftward and rightward dependencies in English: while leftward dependencies such as wh-movement can in principle be unbounded, rightward dependencies such as heavy NP shift and extraposition from NP are subject to strict locality conditions (Ross 1967, Akmajian 1975, Baltin 1981, 1983). Recently, there are proposals that seek to derive the unavailability of long-distance rightward dependency from the nature of sentence processing in the performance systems (see e.g. Rochemont 1992, Alphonce and Davis 1997, Ackema and Neeleman 2000). This line of arguments can be taken as part of a larger research strategy that “it is a priori desirable to eliminate as much redundancy as possible between different components of the system” (Alphonce and Davis 1997:7). The Prosodic Phase Hypothesis is consistent with this spirit of the division of labor, if it is correct that the right-edge heaviness effect in M-order is a manifestation of performance considerations (cf. 3.2.1). I argue that the locality effect in shifted sentences in part follows from the prosodic properties of M-order, which are interface requirements reflecting performance efficiency.

Remember that M-order exhibits the relative right-edge heaviness effect: it is not only the prosodic weight of the rightmost phrase that matters, as is shown in (85B1) versus (85B2), but also the prosodic weight of an intervening phrase, as is shown by the degraded example in (85B3).

(85) A: What happened yesterday?
M-order
B1: Kay donated to the library [\(\text{intP}\) five hundred Canadian dollars and her collection of novels by Mishima]. [= (15B1)]
B2: #Kay donated to the library [\(\text{intP}\) her collection of novels by Mishima]. [= (12B)]
B3: #Kay donated to the UBC Asian library located near the Nitobe Japanese garden [\(\text{intP}\) five hundred Canadian dollars and her collection of novels by Mishima]. [= (67b)]

Therefore, the alleged locality effect in the heavy NP shift alternation, at least in part, derives from the relative weight effect: the rightmost phrase can be non-adjacent to the head verb to the extent that the rightmost phrase is prosodically heavier than the preceding part of the sentence, satisfying the Prosodic Weight Condition. However, it is not the case that the rightmost phrase may be arbitrarily far away from the head verb, and in this thesis, I limit the scope of the Prosodic Phase Hypothesis to linearization of dependents in the same domain, e.g. VP (cf. 1.3.2).

3.3.2. Arguments for Incremental Structure-Building in the Syntax

Assuming a strictly derivational model of the computational component, incremental structure-building in the syntax allows the most natural implementation of the Prosodic Phase Hypothesis: syntactic objects (SOs) are built incrementally from left to right, mapped to prosodic objects (POs), whereby SO-PO pairs are derivationally established and spelled out into the phonological component. In 3.3.2.1, I show that the results of syntactic constituency tests applied to shifted sentences support incremental structure-building in the syntax. In 3.3.2.2, I show that right-branching structure resulting from incremental structure-building is consistent with structure-sensitive LF effects (e.g. binding) exhibited by shifted sentences.

13 An important predecessor of this idea is Truckenbrodt (1994). He argues that the boundedness effect in German extraposition from NP is best captured by the position of intonation boundaries.
3.3.2.1. Syntactic Constituency Tests

Phillips (1996, 2003) convincingly shows that the conflicting results of different constituency tests are explained by assuming incremental structure-building: different constituents are visible to different syntactic processes (e.g. coordination, ellipsis, movement), according to which point of the derivation each process applies at. Phillips' main claim is that a syntactic process can only manipulate elements that are syntactic constituents at the point in the incremental derivation when the process applies (Phillips 2003:47). Since the Prosodic Phase Hypothesis assumes that not only non-shifted sentences but also shifted sentences are incrementally built in the syntax, shifted sentences should show the same results as non-shifted sentences with respect to syntactic constituency tests. Following Phillips, I make the specific prediction for coordination and ellipsis tests:

\[(86)\] Constituency tests such as (i) coordination and (ii) ellipsis, which involve the coordination of the same type of strings \(\alpha_1\) and \(\alpha_2\), may refer to the string \(\alpha_1\) that is a syntactic constituent at the point in the incremental derivation when the conjunction ‘and’ is merged.

In this section, I show that the prediction in (86) is borne out by shifted sentences involving coordination or ellipsis.

3.3.2.1.1. Coordination Test

First, let us look at the coordination test. It is well known that coordination is a liberal diagnostic of constituency, allowing coordination of many strings that do not seem to form constituents under traditional phrase structure analyses. Under incremental structure-building, the liberality of the coordination test is predicted: since the two conjuncts in coordinated structures are almost string adjacent (separated only by the conjunction ‘and’), the first conjunct does not lose its constituency at the point when the conjunction is merged (Phillips 2003:47-51). Consider the following example:

\[(87)\] [These cats [saw [these rats]]].

\[\begin{array}{c}
\text{step 1} \Rightarrow \text{step 2} \Rightarrow \text{step 3} \Rightarrow \text{step 4} \\
\\
\text{these cats} \quad \text{these cats} \quad \text{these cats saw} \quad \text{these cats saw these rats}
\end{array}\]

\[(88)\] step 1 [these cats]

a. \([\alpha_1 \text{ These cats}] \text{ and } [\alpha_2 \text{ those dogs}] \text{ saw these rats.}\]

b. \(\text{These } [\alpha_1 \text{ cats}] \text{ and } [\alpha_2 \text{ dogs}] \text{ saw these rats.}\)

\[(89)\] step 2 [these cats saw]

a. \([\alpha_1 \text{ These cats saw}] \text{ and } [\alpha_2 \text{ those dogs ate}] \text{ these rats.}\]

b. \(\text{These cats } [\alpha_1 \text{ saw}] \text{ but } [\alpha_2 \text{ ignored}] \text{ these rats.}\)

c. *\(\text{These } [\alpha_1 \text{ cats saw}] \text{ but } [\alpha_2 \text{ dogs ignored}] \text{ these rats.}\)

\[(90)\] step 3 [these cats saw these]

a. \([\alpha_1 \text{ These cats saw these}] \text{ and } [\alpha_2 \text{ those dogs ignored those}] \text{ rats.}\]

b. \(\text{These cats } [\alpha_1 \text{ saw these}] \text{ and } [\alpha_2 \text{ ignored those}] \text{ rats.}\)

c. *\(\text{These cats saw } [\alpha_1 \text{ these}] \text{ and } [\alpha_2 \text{ those}] \text{ rats.}\)

Let us take the step 2 in (89), for example. At this point, the second conjunct can target either the subject DP – V sequence ‘these cats saw’ as in (89a), or just the V ‘saw’ as in (89b), because both of them are
syntactic constituents. However, the N – V sequence ‘cats saw’ cannot be coordinated because the N and the V do not form a constituent. Notice that the DP – V sequence in (89a) is no longer a constituent in the final structure (at the step 4). This is not a problem, because it is a constituent at the step 2 when the conjunction ‘and’ is merged.

Shifted sentences exhibit the same coordination paradigm, and hence bear out the prediction in (86). Firstly, in the marked V – AdvP – DP order, coordinated sentences such as V ‘and’ V – AdvP – DP are possible as is shown in (91).

(91)  [V₁ ‘and’ [V₂ – AdvP] – DP
a. I [v₁ presented] before [v₂ discussing] [AdvP thoroughly] [DP an argument that you told me about].
b. I wanted to [v₁ present] and [v₂ discuss] [AdvP very badly] [DP an argument that you told me about].  

In (91) the sentence-final DP is semantically associated with the two Vs, ‘present’ and ‘discuss’. Examples such as (91a) are referred to as parasitic gap constructions, which involve coordinated structure under certain analyses (e.g. Williams 1989/90). Examples such as (91b) are referred to as Right Node Raising. In these examples, the conjunction ‘before’ or ‘and’ coordinates adjacent Vs, which is predicted to be possible by (86). Likewise, coordination of V – AdvP strings, to the exclusion of a following DP, is also possible.

(92)  [V₁ – AdvP₁] ‘and’ [V₂ – AdvP₂] – DP
a. I [v₁ will presented] [AdvP₁ briefly] before [v₂ discussing] [AdvP₂ more thoroughly] [DP an argument that you told me about].
b. I want to [v₁ present] [AdvP₁ briefly] and [v₂ discuss] [AdvP₂ thoroughly] [DP an argument that you told me about].

Secondly, the PP – DP string in the marked V – PP – DP order can be coordinated as well.

(93)  V – [PP₁ – DP₁] ‘and’ [PP₂ – DP₂]
a. John put [p₁ into the wastebasket] [DP₁ the long letter from his ex-wife complaining about her financial situation], and [p₂ into the drawer] [DP₂ the photos of their five-year-old daughter].  

b. John sent [pp₁ to Mary] [DP₁ a letter] and [pp₂ to Sue] [DP₂ a book].
c. I gave [pp₁ to Maxwell] [DP₁ five dollars] and [pp₂ to Chris] [DP₂ three dollars].  

In the examples in (93), the conjunction ‘and’ coordinates the adjacent PP – DP strings. This is possible because the PP – DP string forms a constituent at the point when the conjunction is merged. This is schematized in (94).

(94)  Derivation of coordinated shifted sentences
The conjunct $\alpha_1 (= PP_1 - DP_1)$ is a constituent at the step 2 when the conjunction ‘and’ is merged. The conjunct $\alpha_1$ and the conjunct $\alpha_2 (= PP_2 - DP_2)$ are strings of the same type (i.e. PP - DP). Notice here that conjuncts $\alpha_1$ and $\alpha_2$ do not correspond to Intonational Phrases. This is not a problem, because the prediction in (86) only concerns syntactic constituency, and is not sensitive to which syntactic object is already spelled out as an IntP.

Likewise, the sentence-final DPs can be coordinated:

\[
(95) \quad V - PP - [DP,] \text{ ‘and’ [DP}_2]
\]

What the Kazemi case did was to highlight for Canadian people $\text{[DP}_1$, the situation of journalists in Iran] and $\text{[DP}_2$, the absence of freedom of expression].

(Vancouver Sun, 21 November, 2003)

This is also compatible with the prediction in (86), because the coordinated DPs are string adjacent and hence the first conjunct DP forms a constituent at the point when the conjunction ‘and’ is merged.

### 3.3.2.1.2. Ellipsis Test

Next, let us look at the VP-ellipsis test. In shifted sentences, the elided element in the second conjunct can be anaphoric to the whole $V - PP - DP$ string in the first conjunct as in (96), but not to the partial $V - PP$ string, stranding a DP behind as is shown in (97):

\[
(96) \quad [\alpha_1 V - PP - DP,] \text{ ‘and’ } [\alpha_2 \phi]
\]

\(\text{a. John } [\alpha_1 \text{ gave to Mary a picture of Lyndon Johnson}], \text{ and Bill } [\alpha_2 \text{ did}] \text{ too.}
\)

\(\text{b. John } [\alpha_1 \text{ read in The Times a scathing review of his new book}], \text{ and Sally } [\alpha_2 \text{ did}] \text{ too.}
\)

\(\text{c. Sally } [\alpha_1 \text{ noticed in the foyer a famous portrait by Rembrandt}], \text{ and Bill } [\alpha_2 \text{ did}] \text{ too.}
\)

(Rochemont and Culicover 1990:118)

\[
(97) \quad [V - PP] - DP \text{ ‘and’ } [\phi] - DP
\]

\(\text{*John [bought for Mary] a picture of her father, and Sally [did] every book she could find.}
\)

(ibid:120)

The contrast between (96) and (97) bears out the prediction in (86): in (96), the antecedent of the elided element forms a constituent $\alpha_1$ at the point when the conjunction ‘and’ is merged, whereas in (97), the antecedent of the elided element does not form a constituent at the same point because of the existence of the following DP.

\[
(98) \quad \text{The point when the conjunction ‘and’ is merged}
\]
In this section, we saw that syntactic constituencies in shifted sentences involving coordination and VP-ellipsis are well captured by incremental structure-building in the syntax. In the next section, I will show that the LF effects exhibited by shifted sentences are consistent with their right-branching structure resulting from incremental structure-building.

### 3.3.2.2. Base-Generation and LF Effects

#### 3.3.2.2.1. Interpretation of the Shifted XP

The Prosodic Phase Hypothesis argues that shifted sentences as well as non-shifted sentences are base-generated incrementally via Merge and do not involve syntactic movement. The legitimate derivations of non-shifted and shifted sentences in the verbal domain are schematized as follows:

(99) **Derivation of non-shifted and shifted sentences in the verbal domain**

**a. Non-shifted sentence: C-order**

\[
\text{e.g. } [\text{IntP} \ldots V - \text{DP} - \text{PP}]
\]

Numeration

\[\downarrow\]

Incremental structure-building

**b. Shifted sentence: M-order**

\[
\text{e.g. } [\text{IntP} \ldots V - \text{PP}][\text{IntP} \text{DP}]
\]

Numeration

\[\downarrow\]

Incremental structure-building

The issue of whether shifted sentences are derived by syntactic (rightward) movement or not is independent of incremental structure-building in the syntax. In other words, nothing prevents postulating a trace of a DP immediately after the V in the shifted sentence in (99b), and deriving M-order from C-order via downward movement (e.g. \(V - t_{DG} - \text{PP} - \text{DP}\)).

Under a base-generation approach to XP-shift alternations, both C-order and M-order are derived in the syntax freely, and illegitimate orderings are ruled out by correspondence conditions on the syntax-prosody mapping. On the other hand, under a movement approach, illegitimate shifted sentences are not motivated, and hence not derived in the syntax.

(100) **Base-generation vs. movement approaches**
a. Base-generation of non-shifted and shifted sentences

\[
\begin{array}{c}
\text{Non-shifted sentence} \\
(\text{e.g. } V - DP - PP)
\end{array}
\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad
(103) Condition C effects
   a. *We gave _ to [him], on Friday [DP [John]'s brand-new toy].
   b. *[He], was willing to discuss _ yesterday [DP the claim that [Phil], made].
      (Shiobara 2002b)

In the acceptable shifted sentences in (101), the reflexive `herself` is bound by the preceding DP `Mary` in (a), and the reciprocal `each other` is bound by the preceding DP `them` in (b). In the acceptable shifted sentence in (102), the pronoun `his` is bound by the quantificational phrase `every student`, having a bound variable interpretation. In the unacceptable shifted sentences in (103), names `John` and `Phil` are bound by the preceding indirect object `him` in (a), and the subject `he` in (b), respectively, resulting in Condition C violations. Although binding effects may show variations depending on discourse context, there is an agreement that the sentence-final DP in shifted sentences is structurally lower than the preceding subject DP or object DP (see Rochemont and Culicover 1990 and the references in (101)-(103) from which I cite the examples).

Furthermore, Williams (1994) points out that binding possibilities differ between non-shifted and shifted sentences:

(104) Condition A effects
   a. *I recommended [himself], to [Bill].
   b. ?I recommended _ to [Bill] [DP [himself]].

(105) Condition B effects
   a. *I gave a copy of [his], grades to [every boy].
   b. I gave _ to [every boy] [DP a copy of [his], grades]. (Williams 1994:154-155)

In (104)-(105), the (a) sentences exhibit the canonical V – DP[theme] – PP[goal] order, and the (b) sentences exhibit the marked V – PP[goal] – DP[theme] order. The (a) sentences are unacceptable, suggesting that the preceding theme DP should not be bound by the following goal DP. The (b) sentences are acceptable, suggesting that the preceding goal DP should bind the following theme DP. The generalization is that the preceding DP binds and hence c-commands the following DP, regardless of which DP is theme or goal. This is consistent with right-branching structure assigned to non-shifted as well as shifted sentences.

Assuming that binding possibilities are determined at LF, the different binding possibilities in non-shifted and shifted sentences argue against the traditional view that shifted sentences are stylistically marked sentences derived in the phonological component (as in e.g. Ross 1967, Rochemont 1978, see 3.4.1.1) An incremental approach to structure-building in the syntax solves this problem.17

3.4. Situating the Prosodic Phase Hypothesis in Previous Analyses of XP-Shift

There are two types of analyses involved in the derivation of shifted sentences in English: one is to derive the shifted/marked order (M-order) from the corresponding non-shifted/canonical order (C-order) via movement, either in the phonology or in the syntax; the other is to base-generate M-order in the syntax. The Prosodic Phase Hypothesis is a base-generation analysis: M-order as well as C-order are base-generated in the syntax, and illegitimate orderings are filtered out by correspondence conditions on the syntax-prosody mapping. This section reviews some representative analyses of XP-shift in English to highlight their differences from the Prosodic Phase Hypothesis.

---

16 I need to assume that the rightmost DP in the shifted (b) sentences in (104)-(105) is prosodically heavy, as long as these sentences are acceptable. Since these DPs do not contain more prosodic words than the preceding part of the sentence, it should be the case that the DPs carry extra prosodic prominence in an appropriate context although it is not indicated in Williams (1994) (from which I cited these examples).

17 See Shiobara (2003) for a linear order-based analysis of binding effects exhibited by extraposition from NP.
3.4.1. Movement Approaches

3.4.1.1. Phonological Movement

3.4.1.1.1. Stylistic Rule: Ross (1967/86) and Rochemont (1978/85)

A traditional analysis of shifted sentences in English, involving rightward dependencies such as heavy NP shift (HNPS) and extraposition from NP (EXNP), is to derive them via rightward movement. In the pre-Principles and Parameters era of generative grammar, Ross (1967/86) formulated HNPS (or “Complex NP Shift” in his term) and EXNP in the following way:

(106)  
\[ X - \text{NP} - Y \rightarrow \text{OPT} \]

b. Extraposition from NP (Ross 1986:4)
\[ X - [\text{NP} - S] - Y \rightarrow \text{OPT} \]

In (106), a shifted sentence involving HNPS or EXNP is optionall derived from its canonical counterpart, via an “OPT” stipulation in the description of the structural change. Ross regards HNPS and EXNP as post-cyclic rules, suggesting that they occur in the phonological component and do not feed any syntactic rules. Although Ross acknowledges that shifted sentences are “stylistic” in some sense, the right-edge heaviness effect exhibited by shifted sentences is not discussed.

The term “stylistic” is based on the norm of “stylistic rule”, which was introduced in Chomsky and Lasnik (1977) and further discussed in Rochemont (1978/1985):

[PP extraposition from NP] constructions will be motivated as stylistic, in the sense that they apply to surface structures [...]. Evidence will be advanced to show that stylistic constructions of this sort do not interact with syntactic transformations and are in no way relevant to the determination of semantic properties of sentences.

(Rochemont 1985:2-3)

Rochemont (1978/1985) argues that stylistic rules, including HNPS and EXNP operations, are movement operations in the phonological component, and yet contribute to pragmatic representations. Rochemont proposes the model of the grammar as in (107):

(107)  
\[ \text{PS Rules} \rightarrow \text{Transformations} \rightarrow \text{Surface Structure} \]
\[ \text{Stylistic rules} \rightarrow \text{Interpretive rules} \]
\[ \text{Derived Surface Structures} \rightarrow \text{LF} \]
\[ \text{Pragmatic Representations} \rightarrow \text{Other cognitive representations} \]

(cf. Rochemont 1985:20,106)
In this model, optionality is incorporated into the way stylistic rules apply: stylistic rules apply freely to Surface Structure representations.

The advantage of a phonological movement approach to shifted sentences is that the different behavior of shifted sentences (involving e.g. HNPS or EXNP) and sentences involving leftward syntactic movement (e.g. wh-movement, DP-movement) follows from the assumption that shifted sentences are derived in the phonological component, and therefore do not have to obey syntactic constraints. The weakness of this approach is that, as Rochemont (1985:109) notices, the stylistic rules permit the derivations of many sentences which seem to be systematically excluded (e.g. HNPS with preposition-stranding, as we saw in 3.3.1.1.3). That is to say, stylistic rules are too unrestricted by themselves.


Among phonological movement analyses of shifted sentences, Zubizarreta's (1998) analysis of HNPS is closer to the Prosodic Phase Hypothesis than Ross' (1967/86) or Rochemont's (1978/85), in that Zubizarreta regards HNPS as a "prosodically-motivated" movement (p-movement) in the syntax. Adopting the predicate raising analysis proposed by Larson (1989) (see 3.4.1.2.2 below), Zubizarreta (1998) assumes that a shifted sentence is derived "only when the rightmost constituent is analyzed as metrically 'heavy', owing to its syntactically complex structure or to the heavy accent it carries", in order to achieve an intonationally balanced sentence (p.149).

Based on data from Romance languages, which suggest that p-movement has an impact on LF (Zubizarreta 1998:ch.3), and couched theoretically in the Minimalist Program, Zubizarreta proposes the following model of the grammar:

\[
\begin{align*}
\text{(108)} & \quad \downarrow \\
\Sigma\text{-Structure} & \quad \text{(sets of phrase markers, feature checking)} \\
\downarrow & \quad \text{(unique phrase marker)} \\
\text{LF} & \quad \text{(p-movement)} \\
\text{PF} & \quad \text{Assertion Structure} & \quad \text{(Zubizarreta 1998:32)}
\end{align*}
\]

This proposal departs from the standard view of the grammar, in that PF is directly mapped from LF. Within this model of the grammar, shifted sentences are prosodically-motivated and not optional. Zubizarreta's proposal is in part theoretically driven by the Minimalist assumption that every movement must be motivated and hence obligatory. However, it remains unclear exactly how the syntactic component looks ahead to check prosodic properties.

3.4.1.2. Syntactic Movement

3.4.1.2.1. Rightward Movement: Guérøn (1980) and Rochemont and Culicover (1990)

The Principles and Parameters theory (Chomsky 1981) dispenses with construction-specific rules, and derives them from the interaction of more general principles. In the Principles and Parameters framework, HNPS and EXNP fall under the principle of Move α. For example, Guérøn (1980) formulates extraposition of PP from NP as a general principle of Move α in (109):

\[
\text{(109)} \quad \text{PP Extrapolation (Guérøn 1980:639)} \quad \text{Move PP (or Move category)}
\]

89
With this formulation of EXNP, unacceptable EXNP sentences are ruled out by independently motivated constraints on rules of Semantic Interpretation (SI-I and SI-II in her terms, following Chomsky 1976:336). That is to say, any shifted sentences involving EXNP are optionally derived via Move α in the syntax.

Rochemont and Culicover (1990) also analyze HNPS as syntactic movement (in particular, rightward adjunction). In syntactic movement analyses of shifted sentences, prosodic markedness associated with shifted sentences is not given any discussion.

3.4.1.2.2. Leftward Movement: Larson (1989) and Kayne (1994)

The HNPS alternation in English may also be analyzed as leftward movement (see e.g. Larson 1989, Kayne 1994, Rochemont 1998, and Takano 1998). For example, Larson (1989) proposes a predicate raising account of HNPS. In his analysis, the shifted sentence in (110a) is derived from the canonical counterpart via leftward movement of a verbal category as in (110b), not by rightward movement of an NP as in (110c):

(110) a. Mary gave to John everything that he demanded.
   
   b. Mary [v gave to John] everything that he demanded.
   
   Not
   
   c. Mary gave to John [NP everything that he demanded].

The structure of (110b) is shown in (111):

(111)

Larson’s analysis of HNPS assumes that (i) VPs have shell structures (Larson 1988), and (ii) V’ may be reanalyzed as V when it is thematically monotransitive (i.e. determines exactly two thematic roles). If V’-reanalysis does not apply, head-to-head movement of V proceeds, yielding the non-shifted order ('Mary [v gave] everything that he demanded to John'). In Larson’s predicate raising analysis, the derivation of shifted sentences depends on the occurrence of V’-reanalysis, although it is not clear under what condition V’-reanalysis applies.

Kayne (1994) also proposes a leftward movement analysis of HNPS. In Kayne’s Antisymmetry theory of syntax, the Linear Correspondence Axiom (LCA) prohibits rightward adjunction in general, and hence apparent rightward displacement such as HNPS and EXNP need to be reanalyzed as a result of leftward movement. Since movement of a complex constituent as in (111) yields a structure violating the Antisymmetry requirement by having two complex constituents as sisters, Kayne’s analysis of HNPS differs from Larson’s predicate raising analysis in what exactly is moved leftward: the PP moves leftward to the specifier position of a functional head X, independently of V-movement.
In the Antisymmetry theory, the HNPS alternation in English is regarded as an instance of scrambling of the sort found in German, the difference being that "in English the verb ends up to the left of both complements, whereas in the corresponding German sentences, the verb ends up to their right" (Kayne 1994:72). The LCA forces the leftward movement analysis of HNPS.

Kayne (1994) also provides a sketch of how shifted sentences involving EXNP are derived via leftward movement:

\[(113) \text{a. } [\text{NP Something that you should know about}] \text{ just happened.}\]
\[\text{b. } [\text{N Something}] \text{ just happened [t} \text{ that you should know about}].\]

Subject NPs are assumed to be base-generated within the VP. The non-shifted sentence is derived when the whole NP moves leftward as in (113a), whereas the shifted sentence involving EXNP is derived when only the head N moves leftward leaving the post-nominal modifier behind as in (113b). The motivation for this movement is a need for Case, and Kayne notes that there is optionality as to whether (113a) or (113b) is derived (Kayne 1994:120-121).

Thus, in Kayne's (1994) analysis of shifted sentences, a shifted sentence involving HNPS is derived when the PP undergoes leftward scrambling, and a shifted sentence involving EXNP is derived when the N, not the whole NP, moves leftward. In general, a leftward movement approach to shifted sentences is theoretically motivated by particular assumptions about phrase structure (e.g. Larsonian shell structure for VPs, Antisymmetry theory of syntax), and is concerned with how shifted sentences are derived, rather than why they are derived and when. As a result, prosodic properties associated with shifted sentences are simply ignored, and shifted sentences are derived via particular syntactic mechanisms, such as V'-reanalysis in Larson's analysis of HNPS.

Takano (1998) adopts Kayne's Antisymmetry theory of phrase structure. However, Takano's structure differs from Larson's (1989) or Kayne's (1994) in that the base position of Theme is lower than that of non-Theme (e.g. Goal) (following Pesetsky 1995). As a consequence, Takano analyzes the shifted order as base-generated, and proposes that an object NP obligatorily undergoes partial object shift, as in other grammatical instances of NP-PP order: 18,19

\[\text{18 Takano (1998) argues that an NP raises for a Case reason and hence partial object shift conforms to Last Resort in the Minimalist Program. It follows from his analysis that raising is not obligatory in PP-PP order, because a PP does not require Case (p.844):}\]
\[\text{i.}\]
\[\text{a. John talked [to Bill] [about Mary].}\]
\[\text{b. John talked [pp about Mary] [to Bill].}\]

In other words, it remains to be explained why the PP 'about Mary' raises at all in (i-b). Furthermore, Takano's analysis does not account for the increasing weight effect found in non-DP-shift examples (cf. 3.3.1.1.2).

\[\text{19 Pesetsky (1995) proposes a Dual System of phrase structure: Layered Syntax and Cascade Syntax. The shifted sentence in (i), for}\]

\[\text{91}\]
In this analysis, the derivation of shifted sentences is dependent on the possibility of deleting the head of the chain (NP, NP') in the phonological component (assuming the copy theory of movement, as in the Minimalist Program): in the unmarked case, the tail deletes and the canonical NP-PP order is derived, whereas the marked option of deleting the head is allowed only when the head of the chain is in a non-checking position (e.g. a VP-adjoined position in (114)) in the Minimalist sense.20,21

Regarding the markedness of the shifted NP-PP order, Takano (1998) states that shifted sentences involving HNPS bear "some kind of marked stress/intonation at the PF interface and also provide some kind of new information at the LF interface" (p.864); "the marked option of deleting the head of the chains results in a focus stress/intonation of the lower NP at PF; and new information provided by the lower NP at LF" (ibid). Although Takano acknowledges certain prosodic markedness effects associated with shifted sentences, the specific prosodic condition that the rightmost phrase must be an independent Intonational Phrase in shifted sentences is not accounted for in his analysis.

3.4.2. Base-Generation Approaches

Heavy NP shift (HNPS) and extraposition from NP (EXNP) are known to exhibit peculiar properties

Pesetsky proposes that the binding theory applies to Cascade Syntax in (i-b), where the Goal DP 'John' c-commands the Theme DP 'a brand new toy' that has undergone downward movement (see 3.3.2.2.2 for binding effects in shifted sentences). For other purposes such as the c-command condition on traces, the shifted phrase must be higher than the rest of the VP as in (i-a). Takano's (1998) structure for shifted sentences involving HNPS differs from Pesetsky's Cascade Syntax, in that HNPS involves leftward (and upward) movement, rather than a rightward (and downward) movement.

There are two more logical possibilities: to realize both copies (as in certain analyses of partial wh-movement, e.g. McDaniel 1989 for German, and Thornton 1990 for child English), and to realize neither of them (as in null operator movement). These are not observed in the English and Japanese word order alternations at issue, and hence are irrelevant to the current discussion.

20 There are two more logical possibilities: to realize both copies (as in certain analyses of partial wh-movement, e.g. McDaniel 1989 for German, and Thornton 1990 for child English), and to realize neither of them (as in null operator movement). These are not observed in the English and Japanese word order alternations at issue, and hence are irrelevant to the current discussion.

21 See also Wilder (1996) for a "copy-and-deletion" analysis of EXNP.
compared to standard syntactic movement (e.g. wh-movement, DP-movement). Notoriously, HNPS and EXNP are both less and more restricted than wh-movement, and do not obey the same syntactic condition as leftward syntactic movement (see e.g. Ross 1967/86, Guéron 1980, Baltin 1981, 1983, Culicover and Rochemont 1990, Rochemont and Culicover 1990). This is why HNPS and EXNP are sometimes regarded as stylistic rules in the phonological component (see 3.4.1.1.1) or prosodically-motivated movement (see 3.4.1.1.2), rather than purely syntactic movement.

Alternatively, shifted sentences involving HNPS and EXNP are analyzed as base-generated, and therefore they do not have to obey the conditions on syntactic movement. Culicover and Rochemont (1990) and Rochemont and Culicover (1990) provide a base-generation analysis of shifted sentences involving EXNP. They claim that shifted sentences involving EXNP (but not HNPS, as we noted in 3.4.1.2.1) are base-generated, not derived via movement. Note that this is not a necessary argument and there actually are various attempts to explain the exceptional behavior of EXNP, for example, by distinguishing leftward and rightward movements and their respective susceptibility to syntactic constraints. For example, Baltin (1981, 1983) proposes "generalized subjacency", which explicitly distinguishes leftward and rightward movement with respect to the type of bounding nodes that count.

It is to a certain extent a matter of explanatory elegance whether to take a base-generation approach, thereby complicating syntactic constraints, as in Culicover and Rochemont (1990), or to keep a movement approach by modifying syntactic constraints as in Baltin (1981, 1983). While keeping syntactic constraints uniform, Culicover and Rochemont are forced to introduce another principle to ensure that a shifted phrase is related to its antecedent NP. This is referred to as the "Complement Principle":

\[(115) \beta \text{ is a potential complement of } \alpha \text{ only if } \beta \text{ is a constituent governed by } \alpha.\]

(NB: The "complement" is a phrase that holds an adjunct or argument relation to the head of the phrase it takes as antecedent.)

(Culicover and Rochemont 1990:26)

In the case of EXNP, the Complement Principle concerns the interpretation of an extraposed phrase (= \(\beta\)) and its antecedent NP (= \(\alpha\)).

The Prosodic Phase Hypothesis pursued in this thesis applies a base-generation analysis to shifted sentences involving HNPS as well. In 3.3.2.2.1, I relaxed a theta-locality condition on the head V and its arguments by stipulation, which plays a role equivalent to Culicover and Rochemont's Complement Principle in (115), in the sense that they serve to guarantee the semantic association between the head and its non-adjacent argument.

Therefore, movement and base-generation approaches share certain interpretive mechanisms, specifically those of construal relating non-adjacent elements. They differ, however, in their treatment of such relations. In a movement approach, these relations are established in the syntax via the trace left by movement, whereas in a base-generation approach, an additional interpretive rule such as the Complement Principle in (115) relates non-adjacent elements.

A base-generation analysis of shifted sentences shares the same problem of overgeneration with an analysis of shifted sentences as stylistic sentences derived in the phonological component (cf. 3.4.1.1.1): if shifted sentences are freely base-generated, there may arise many sentences which should be systematically excluded. There need to be constraints that would exclude those unacceptable sentences, and the Complement Principle in (115) is such an instance. In the Prosodic Phase Hypothesis, correspondence conditions on the syntax-prosody mapping does this job. In the rest of this section, I review a strictly representational approach to natural language, "Representation Theory" proposed by Williams (2003), and look at how shifted sentences are treated in the Representational Theory.

---

The copy theory of movement in the Minimalist Program also makes it hard to distinguish between the two types of approaches. (See Chomsky 1981:90-92 for discussion on "virtual interchangeability" of movement approaches and base-generation approaches.)
Williams' (2003) Representation Theory analyzes displacement phenomena in strictly representational terms. In his model, the syntax of a sentence is "a collection of structures" or sublanguages, e.g. Theta Structure (TS), Case Structure (CS), Surface Structure (SS), Quantificational Structure (QS), Focus Structure (FS), Prosodic Structure (PS), and "a set of shape-conserving mappings among them" (Williams 2003:2).

Williams argues that a shifted sentence involving HNPS does not result from movement, but arises from the "mismapping" between CS and SS for the sake of the proper mapping between SS and FS. Given the contrast between the two shifted sentences in (116c) and (116d) below for example, Williams argues that "HNPS must take place only to aid and abet canonical FS representation, in which focused elements are final" (Williams 2003:34). The background assumption of his argument is that sentence-final position is the canonical focus position in English. On the other hand, shifted sentences involving HNPS deviate from canonical CS representation in that the DP is not adjacent to the V that assigns Case.

\[ \begin{array}{c|c|c}
\text{C-order} & \text{CS-SS mapping} & \text{SS-FS mapping} \\
\hline
\text{a. John gave [DP all of the money in the satchel] to MARY.} & \checkmark & \checkmark \\
\text{b. John gave [DP all of the money in the SATCHEL] to Mary.} & \checkmark & \checkmark \\
\text{M-order/shifted order} & & \text{*} \\
\text{c. John gave to Mary [DP all of the money in the SATCHEL].} & \text{*} & \checkmark \\
\text{d. *John gave to MARY [DP all of the money in the satchel].} & \text{*} & \text{*} \\
\end{array} \]

(Williams 2003:34)

According to Williams, shifted sentences are acceptable only if the sentence-final DP contains a focus (indicated by upper case letters following Williams' notation), whereas canonical sentences are acceptable regardless of the position of a focus. In the acceptable shifted sentence in (116c), the mismatch between CS and SS is tolerated because of the SS-FS match. In (116d), however, both CS-SS and SS-FS are mismatched and the sentence is ruled out.

What does Representation Theory tell us about a (prosodic) markedness effect in the HNPS alternation? Nothing much. Given the acceptability of the sentence with the canonical V-DP-PP order in (116b), Williams (2003) stipulates that HNPS is "optional" (p.35-37). The (prosodically) marked status of the shifted order is not predicted unless we assume that the mapping between SS and FS is more marked than the mapping between CS and SS in English, which needs to have independent motivation. Although the Prosodic Phase Hypothesis also takes a base-generation approach to shifted sentences, the empirical emphasis regarding shifted sentences involving HNPS differs between the two theories. In the Representation Theory, M-order is characterized as having a focus in the sentence-final DP; in the Prosodic Phase Hypothesis, M-order is characterized by the prosodic property of the sentence-final DP as an Intonational Phrase, and focus in itself only indirectly characterizes the shifted order via its prosodic realization as increased prosodic prominence.

3.4.3. Summary

The Prosodic Phase Hypothesis is a base-generation approach to English XP-shift: the shifted order as well as the non-shifted order are base-generated in the syntax, and illegitimate orderings are filtered out by correspondence conditions on the syntax-prosody mapping. The Prosodic Phase Hypothesis is based on the observation that the shifted order is prosodically more marked than the non-shifted order. It differs from any movement or base-generated approaches we have reviewed in this section, but shares some insight with Zubizarreta (1998) and Williams (2003) in the following respects.

Zubizarreta's (1998) p-movement approach to XP-shift (reviewed in 3.4.1.1.2) is similar to the
Prosodic Phase Hypothesis in that it aims to account for the prosodic property of the shifted order. However, under the p-movement analysis, the shifted order is derived from its non-shifted counterpart via p-movement in the syntax, and not base-generated.

In William's (2003) Representation Theory (reviewed in 3.4.2.1), both non-shifted and shifted sentences are base-generated as in the Prosodic Phase Hypothesis, and illegitimate orderings are filtered out when they are not motivated for any mapping between structural representations. Williams' analysis, however, centers on the focus effect in shifted sentences and their prosodic markedness is not discussed.

Appendix. Details of the Elicitation

A1. Purpose

A systematic elicitation was designed as a test of how a native speaker of English pronounces sentences with the canonical V – DP – PP order and sentences with the marked V – PP – DP order (i) with certain focus interpretations (pronunciation task), and how natural or unnatural these sentences sound to native speakers of English (i) with certain focus interpretations and (ii) with certain prosodic patterns (judgment task). The sentences produced in the pronunciation task were used as the stimuli sentences for the judgment task.

A2. Procedure and Methodology

A2.1. Focus Patterns

A question-answer format was used to construct sentences with certain focus interpretations. Two sets of alternations were examined: the first alternation pattern involves the V ‘donated’, which selects two internal arguments, a theme DP and a goal PP; the second alternation pattern involves the V ‘ate’, which selects one internal argument, a theme DP, followed by an optional adverbial PP. For each set, sentences with five types of focus interpretations were examined: a broad-focus (i.e. IP-focus) interpretation, two single narrow-focus (DP-focus and PP-focus) interpretations, and two double narrow-focus (V+DP-focus and V+PP-focus) interpretations.

Stimuli sentences

The (A1) sentences represent the results of the pronunciation task.

The following notational conventions are used:

default       lexical item in bolding carries default sentence stress
HEAVY          lexical item in upper case letters carries extra prosodic prominence
focused       intended narrow-focus is underlined

Alternation pattern 1: V = ‘donated’
C-order: V – DP[theme] – PP[goal]
M-order: V – PP[goal] – DP[theme]

(1) IP-focus
C-order
Q: What happened yesterday?
A1: Kay donated her collection of novels by Mishima to the Asian library.
M-order
Q: What happened yesterday?
A1: Kay donated to the Asian library her collection of novels by Mishima.
(2) DP-focus
C-order
Q: What did Kay donate to the Asian library?
A1: She donated her collection of novels by MISHIMA to the library.
M-order
Q: What did Kay donate to the Asian library?
A1: She donated to the library her collection of novels by MISHIMA.
A2: She donated to the library her collection of novels by MISHIMA. (*1)

(3) V+DP-focus
C-order
Q: UBC Asian library is facing a serious financial problem. What did Kay do to save the library?
A1: She donated her collection of novels by MISHIMA to the library.
M-order
Q: UBC Asian library is facing a serious financial problem. What did Kay do to save the library?
A1: She donated to the library her collection of novels by MISHIMA.
A2: She DONATED to the library her collection of novels by MISHIMA. (*2)

(4) PP-focus
C-order
Q: Where did Kay donate her collection of novels by Mishima?
A1: She donated them to the Asian LIBRARY.
Q: Where did Kay donate her collection of novels by Mishima?
A2: She donated them to the ASIAN LIBRARY.
M-order
Q: Where did Kay donate her collection of novels by Mishima?
A1: She donated them to the Asian LIBRARY.
A2: She DONATED them to the Asian LIBRARY. (▲1)

(5) V+PP-focus
C-order
Q: Kay will go back to Japan soon, but she had too many books to ship to Japan. Do you know what she did with her collection of novels by Mishima?
A1: She donated them to the Asian LIBRARY.
Q: Kay will go back to Japan soon, but she had too many books to ship to Japan. Do you know what she did with her collection of novels by Mishima?
A2: She DONATED them to the Asian LIBRARY.
M-order
Q: Kay will go back to Japan soon, but she had too many books to ship to Japan. Do you know what she did with her collection of novels by Mishima?
A1: She donated them to the Asian LIBRARY.

Alternation pattern 2: V = 'ate'
C-order: V – DP[theme] – PP[time]

(6) IP-focus
C-order
Q: What happened yesterday?
A1: Ken ate more than ten California rolls in five minutes.
M-order
Q: What happened yesterday?
A1: Ken ate in five minutes more than ten California rolls.

(7) DP-focus
C-order
Q: I heard that Ken was at the party only for five minutes or so because he had to catch the last ferry at 9:00pm. But he said he thoroughly enjoyed the food there. What did Ken eat so rapidly?
A1: He ate more than ten CALIFORNIA rolls in five minutes.

M-order
Q: I heard that Ken was at the party only for five minutes or so because he had to catch the last ferry at 9:00pm. But he said he thoroughly enjoyed the food there. What did Ken eat so rapidly?
A1: He ate in five minutes more than ten CALIFORNIA rolls.

Q: I heard that Ken was at the party only for five minutes or so because he had to catch the last ferry at 9:00pm. But he said he thoroughly enjoyed the food there. What did Ken eat so rapidly?
A2: He ate in five minutes MORE than TEN CALIFORNIA rolls.

V+DP-focus
C-order
Q: I heard that Ken was at the party only for five minutes or so because he had to catch the last ferry at 9:00pm. But he said he thoroughly enjoyed the party. What did Ken do in such a short time?
A1: He ate more than ten CALIFORNIA rolls in five minutes.

M-order
Q: I heard that Ken was at the party only for five minutes or so because he had to catch the last ferry at 9:00pm. But he said he thoroughly enjoyed the party. What did Ken do in such a short time?
A1: He ate in five minutes more than ten CALIFORNIA rolls.

Q: I heard that Ken was at the party only for five minutes or so because he had to catch the last ferry at 9:00pm. But he said he thoroughly enjoyed the party. What did Ken do in such a short time?
A2: He ATE in five minutes more than TEN CALIFORNIA rolls.

PP-focus
C-order
Q: How rapidly did Ken eat more than ten California rolls?
A1: He ate them in FIVE minutes.

Q: How rapidly did Ken eat more than ten California rolls?
A2: He ate them in FIVE MINUTES.

M-order
Q: How rapidly did Ken eat more than ten California rolls?
A1: He ate in FIVE minutes them. (*)

(9) V+PP-focus
C-order
Q: Ken was talking about ten or so California rolls. What happened to them?
A1: He ate them in FIVE minutes.

Q: Ken was talking about ten or so California rolls. What happened to them?
A2: He ATE them in FIVE minutes.

M-order
Q: Ken was talking about ten or so California rolls. What happened to them?
A1: He ate in FIVE minutes them.
A2.2. Prosodic Patterns

A2.2.1. Procedure for the Pronunciation Task

Subjects

I had two native speakers of English pronounce the stimuli sentences: the speaker TP (male) pronounced the (Q) sentences and the speaker CG (female) pronounced the (A) sentences. Both were graduate linguistics students at UBC.

Procedure

The (A1) sentences in the stimuli sentences above represent how the speaker CG pronounced the sentences without any instruction given. The (A2) sentences are prosodically controlled sentences: the speaker CG was told to focus on what was being asked by the question. For example, in the V+NP-focus sentences with M-order in (3) (repeated below), CG was told to focus on the fact that what Kay did to save the library was to “donate her collection of novels by Mishima”. Given this instruction, CG said that she would put extra prosodic prominence on the V ‘donated’ and produced the (A2) sentence.

(3) V+NP-focus
M-order
Q: UBC Asian library is facing a serious financial problem. What did Kay do to save the library?
A1: She donated to the library her collection of novels by Mishima.
A2: She DONATED to the library her collection of novels by Mishima.

I elicited the controlled (A2) sentences in relevant cases only, in order to include them in the stimuli sentences for the judgment task.

Equipment

The stimuli sentences were recorded on a portable Marantz cassette recorder, sent to a PC computer, and saved as sound files. Recorded files were digitized and analyzed using the Praat program.

A2.2.2. Procedure for the Judgment Task

Subjects

I had five native speakers of English listen to the stimuli sentences and evaluate the naturalness of (A) sentences as answers to the corresponding (Q) sentences. All the subjects were graduate linguistics students at UBC, and were not told the purpose of the study.

Procedure

The subjects were asked to evaluate the naturalness of (A) sentences as answers to the corresponding (Q) sentences in six scales: 6 = completely natural; 1= completely unnatural. The sound files were presented as speaker icons on Microsoft Powerpoint program, with each Q-A pair corresponding to each icon. A subject put on a headset connected to the computer and heard a Q-A pair by clicking on an icon. Subjects were allowed to listen to any pair of sentences anytime and as many times as they wanted.
A3. Specific Predictions

A3.1. C-order

First, the table in (11) summarizes the specific predictions of the Prosodic Phase Hypothesis for the prosodic properties of C-order.

(11) Predictions for C-order

<table>
<thead>
<tr>
<th>C-order: V – DP – PP</th>
<th>A. Pronunciation task</th>
<th>B. Judgment task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Broad-focus</td>
<td>Default (i.e. rightmost) stress on PP</td>
<td>6</td>
</tr>
<tr>
<td>Single narrow-focus</td>
<td>DP</td>
<td>Focal stress on DP (optionality arises)</td>
</tr>
<tr>
<td></td>
<td>PP</td>
<td>Focal stress on PP or Normal stress on PP</td>
</tr>
<tr>
<td>Double narrow-focus</td>
<td>V+DP (adjacent)</td>
<td>Focal stresses on V and DP (optionality arises)</td>
</tr>
<tr>
<td></td>
<td>V+PP (non-adjacent)</td>
<td>Focal stresses on V and PP</td>
</tr>
</tbody>
</table>

A3.1.1. Pronunciation Task

The Default Sentence Stress Assignment in (6a) predicts that default sentence stress (i.e. rightmost stress in a sentence) should fall on the rightmost prosodic word in the PP.

(12) The locus of default sentence stress in the canonical V – DP – PP order

For DP-focus, the prediction is that the DP should carry focal stress, because the projection of [Focus]_so onto the DP node is not possible by focus projection. On the other hand, for PP-focus, it is predicted that focal stress on the PP should not be necessary, because the PP is marked by [Focus]_so and dominates the default sentence stress position.

For V+DP-focus, the prediction is that both the V and the DP should carry focal stress, because neither of them dominates the default stress position. Likewise, for V+PP-focus, it is predicted that not only the V but also the PP should carry focal stress. Note that the PP needs focal stress, not just default sentence stress, because the sentence would encode only V-focus otherwise.

A3.1.2. Judgment Task
Since C-order in English is not subject to any prosodic condition, C-order is predicted to be always natural (= 6 in the scale) unless the availability of M-order degrades C-order in DP-focus and V+DP-focus.

A3.2. M-order

The table in (13) summarizes the specific predictions for the prosodic properties of M-order.

(13) Predictions for M-order

<table>
<thead>
<tr>
<th>M-order: V - PP - DP</th>
<th>A. Pronunciation task</th>
<th>B. Judgment task</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. donate [pp to the library]</td>
<td>Default (i.e. rightmost) stress on DP</td>
<td>1-6 (C-order is chosen unless M-order conforms to right-edge heaviness)</td>
</tr>
<tr>
<td>[DP the novels by Mishima]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broad-focus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single narrow-focus</td>
<td>DP</td>
<td>Focal stress on DP or Normal stress on DP</td>
</tr>
<tr>
<td></td>
<td>PP</td>
<td>Focal stress on PP</td>
</tr>
<tr>
<td>Double narrow-focus</td>
<td>V+DP (non-adjacent)</td>
<td>Focal stresses on V and DP</td>
</tr>
<tr>
<td></td>
<td>V+PP (adjacent)</td>
<td>Focal stresses on V and PP</td>
</tr>
</tbody>
</table>

A3.2.1. Pronunciation Task

In M-order, the rightmost prosodic word is contained in the DP, and hence the DP should carry default sentence stress, which allows focus projection.

For DP-focus, it is predicted that focal stress on the DP should not be necessary, because the DP is marked by [Focus]\textsubscript{30} and dominates the default sentence stress position. On the other hand, for PP-focus, the prediction is that the PP should carry focal stress, because the projection of [Focus]\textsubscript{30} onto the PP node is not possible by focus projection.

For V+DP-focus, it is predicted that not only the V but also the DP should carry focal stress. Note that the DP needs focal stress, not just default sentence stress, because the sentence would encode only V-focus otherwise. Likewise, for V+PP-focus, the prediction is that both the V and the PP should carry focal stress, because neither of them dominates the default sentence stress position.

A3.2.2. Judgment Task

In broad-focus sentences where no focal stress is involved, M-order is predicted to be natural if the rightmost DP forms an Intonational Phrase by containing a large number of prosodic words.

In narrow-focus sentences, the naturalness of M-order should vary depending on whether the DP is (a part of) a narrow-focus (e.g. DP-focus and V+DP-focus), or not (e.g. PP-focus and V+PP-focus). If the DP is narrowly focused, M-order is predicted to be natural to the extent that the rightmost DP carries focal stress and is prosodically heavy. Otherwise, M-order is predicted to be unnatural compared with the corresponding C-order.

A4. Results
A4.1. Pronunciation task

The (A1) sentences in the stimuli sentences represent how the sentences are pronounced by Speaker CG without any instruction, and the (A2) sentences represent how the sentences are pronounced by CG when she was told to intentionally focus on the intended focus. The table in (14) summarizes the results of the pronunciation task.

(14) Results of the pronunciation task

<table>
<thead>
<tr>
<th>Focus pattern</th>
<th>Word order</th>
<th>C-order (V – DP – PP)</th>
<th>M-order (V – PP – DP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad-focus</td>
<td>DP</td>
<td>Rightmost stress on PP</td>
<td>Rightmost stress on DP</td>
</tr>
<tr>
<td>Single narrow-focus</td>
<td>PP</td>
<td>Focal stress on DP</td>
<td>Focal stress on DP</td>
</tr>
<tr>
<td>Double narrow-focus</td>
<td>V+DP</td>
<td>Focal stress on DP</td>
<td>Focal stress on PP</td>
</tr>
<tr>
<td></td>
<td>V+PP</td>
<td>Focal stress on PP</td>
<td>Focal stress on PP</td>
</tr>
</tbody>
</table>

Regardless of word order, focal stress is employed to mark a narrow-focus.

More details

IP-focus in (1) and (6)

Since IP-focus is broad-focus by definition, the rightmost lexical stress in the sentence is taken as default sentence stress. In the alternation pattern 1, the rightmost stress is found on the rightmost prosodic word, the N ‘library’ in C-order, and the N ‘Mishima’ in M-order. In the alternation 2, the rightmost stress is found on the rightmost prosodic word, the N ‘minutes’ in C-order, and the compound N ‘California-rolls’ in M-order.

The native speakers, both TP (who pronounced the (Q) sentences) and CG (who pronounced the (A) sentences), found C-order completely natural, and M-order less natural than C-order, for the IP-focus sentences.

DP-focus in (2) and (7)

In C-order, focal stress is found in the DP, specifically on the N ‘Mishima’ in (2) and on the compound N ‘California-rolls’ in (7). In (2), the following PP is deaccented.

In M-order, focal stress is found in the rightmost DP, specifically on the N ‘Mishima’ in (2), and on the compound N ‘California-rolls’ in (7). The locus of focal stress coincides with the locus of default sentence stress, but the pitch peak of focal stress is higher than that of default sentence stress.

Speaker CG produced the (A2) sentences when she was given the instruction to intentionally focus on the DP. In (2), CG put extra prosodic prominence on the N ‘Mishima’, and the produced sentence exhibited the same prosodic pattern as the corresponding (A1) sentence (see *1 in the stimuli sentences). In (7), the prosodic pattern of (A2) differs from (A1) in that more than one prosodic word in the DP receives focal stress. In particular, the prosodic words, ‘more’ and ‘ten’, were pronounced with increased prosodic prominence and more slowly.

Both TP and CG found C-order completely natural, and M-order less natural than C-order in the DP-focus sentences.

PP-focus in (4) and (9)
In C-order, focal stress is found in the rightmost PP, specifically on the N ‘library’ in (4), and on the second last prosodic word ‘five’ in (9), presumably because it gives more information about how rapidly Ken (=he) ate California rolls (=them) than the rightmost prosodic word ‘minutes’.

Speaker CG had a hard time in producing the sentences with M-order, because she found the sentences completely unnatural and almost impossible to pronounce (see *2 in the stimuli sentences). When CG managed to pronounce them, extra prosodic prominence was found on the same prosodic word as in the corresponding C-sentences, and the rightmost pronoun ‘them’ was deaccented.

CG produced the (A2) sentences when she was given the instruction to intentionally focus on the PP. Given this instruction, CG put another extra prominence on another prosodic word in the PP, ‘Asian’ in (4) and ‘minutes’ in (9).

Both TP and CG found C-order completely natural, and M-order completely unnatural.

V+DP-focus in (3) and (8)

In both C-order and M-order, the prosodic pattern is the same as in DP-focus sentences: focal stress is found on the prosodic word in the default sentence stress position, and the V does not carry focal stress.

Speaker CG produced the (A2) sentences when she was given the instruction to intentionally focus on V+NP. Given this instruction, CG put extra prominence on the V ‘donated’ in (3) and ‘ate’ in (8).

Both TP and CG found C-order completely natural, the (A1) sentences with M-order less natural than C-order, and the (A2) sentences with M-order the least natural but not completely unnatural. Comparing (A1) and (A2), (A1) was volunteered by CG and (A2) was prosodically controlled. In this sense, when (A1) and (A2) show different prosodic patterns, (A2) was always found less natural than (A1) by CG, because (A2) contains additional prosodic manipulation she does not expect.

V+PP-focus in (5) and (10)

In both C-order and M-order, the prosodic pattern is the same as in PP-focus sentences.

Speaker CG had a hard time in producing the sentences with M-order, because she found the sentences completely unnatural and almost impossible to pronounce (see *2 in the stimuli sentences). CG produced the (A2) sentences when she was given the instruction to intentionally focus on the V+PP. Given this instruction, CG put extra prominence on the V ‘donated’ in (5) and ‘ate’ in (10).

Both TP and CG found C-order completely natural and M-order completely unnatural.

A4.2. Judgment Task

The results of naturalness judgments of the (A1) and (A2) sentences by five native speakers of English (FC, SG, IW, RW and LB: all graduate linguistics students at UBC) are presented in a six-point scale (6 = completely natural; 1 = completely unnatural). The table in (15) summarizes the results of the judgment task:

(15) Results of the judgment task
    (the ‘=’ sign indicates ‘as natural as’; the ‘≥’ sign indicates ‘as natural as or more natural than’)

102
Overall, C-order is preferred over the corresponding M-order.

A consistent preference for C-order over M-order is observed in broad-focus sentences. The contrast is the clearest when the DP is not (a part of) a narrow-focus (i.e. PP-focus and V+PP-focus): M-order is completely unnatural (= 1 in the scale) in these cases. The naturalness of C-order, on the other hand, shows a certain variety, but at least three out of five subjects judge each C-order as completely natural (= 6). In DP-focus and V+DP-focus sentences, C-order is still preferred over M-order when comparing the overall average, although M-order is not completely unnatural, unlike in PP-focus and V+PP-focus sentences.

Regarding the prosodic pattern, the volunteered pattern (A1) is judged to be as natural as or more natural than the prosodically controlled (A2) counterpart in general, except that adding extra focal stress on the PP in the canonical V–DP–PP order in PP-focus sentences improves their naturalness.

Non-adjacent occurrences of focal stress are not particularly problematic with the marked V–PP–DP order in V+DP-focus sentences, whereas they are degraded with the canonical V–DP–PP order in V+PP-focus sentences.

More details

(16) Results of the judgment task
### Alternation pattern 1

<table>
<thead>
<tr>
<th>Focus-pattern</th>
<th>Subject</th>
<th>FC</th>
<th>SG</th>
<th>IW</th>
<th>RW</th>
<th>LB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) IP-focus</td>
<td>C-order A1: V - DP - PP</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>M-order A1: V - PP - DP</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(2) DP-focus</td>
<td>C A1: V - DP - PP</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>M A1: V - PP - DP</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>A2: V - PP - DP (=A1, see *1)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>(3) V+DP-focus</td>
<td>C A1: V - DP - PP</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>M A1: V - PP - DP</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>A2: V - PP - DP (focal stress on V)</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>(4) PP-focus</td>
<td>C A1: V - DP - PP</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>A2: V - DP - PP (extra stress on PP)</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>(5) V+PP-focus</td>
<td>C A1: V - DP - PP</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>A2: V - DP - PP (focal stress on V)</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>M A1: V - PP - DP</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

### Alternation pattern 2

<table>
<thead>
<tr>
<th>Focus-pattern</th>
<th>Subject</th>
<th>FC</th>
<th>SG</th>
<th>IW</th>
<th>RW</th>
<th>LB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6) IP-focus</td>
<td>C-order A1: V - DP - PP</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>M-order A1: V - PP - DP</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>(7) DP-focus</td>
<td>C A1: V - DP - PP</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>M A1: V - PP - DP</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>A2: V - PP - DP (extra stress on NP)</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>(8) V+DP-focus</td>
<td>C A1: V - DP - PP</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>M A1: V - PP - DP</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>A2: V - PP - DP (focal stress on V)</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>(9) PP-focus</td>
<td>C A1: V - DP - PP</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>A2: V - DP - PP (extra stress on PP)</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>M A1: V - PP - DP</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(10) V+PP-focus</td>
<td>C A1: V - DP - PP</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>A2: V - DP - PP (focal stress on V)</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>M A1: V - PP - DP</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
A5. Issues for Future Research

The major contribution of this elicitation is that it systematically controlled for the discourse context of the stimuli sentences, and presented them as sound files. Hence, it succeeded in taking the prosodic pattern of stimuli sentences into consideration. However, in my elicitation, the stimuli Q-A pairs were not volunteered sentences. This could have been a potential experimental flaw, in that some unexpected semantic/pragmatic infelicity may be included in the Q-A pairs. In a future experiment, this flaw should be improved by presenting only (Q) sentences and having a native speaker volunteer the corresponding (A) sentences, in constructing stimuli sentences.

Another issue has to do with how to control prosodic patterns. One of the findings of the study is that a verb does not receive focal stress in double narrow-focus sentences (e.g. V+DP-focus and V+PP-focus). As I noted with respect to the example in (62B) in 3.1.4.2, however, the most natural context where a verb receives focal stress is when it is contrastively focused, not merely the target of a wh-question. Having the stimuli sentences as statement-correction pairs rather than Q-A pairs would be useful in looking into the possibility of more than one occurrence of extra prosodic prominence.

The most obvious problem is that the number of stimuli sentences and subjects is extremely small, which does not even allow for a statistical analysis of the results. Given the gradient acceptability judgments provided for the two set of stimuli sentences in (1)-(10), which also vary depending on speakers, the number of stimuli sentences should be systematically expanded to include different types of verbs.
CHAPTER IV
Linearization of Verbal Dependents in Japanese

This chapter shows how the Prosodic Phase Hypothesis explains a cluster of properties associated with linearization in the verbal domain in Japanese. First, I show that the short-scrambling alternation in Japanese maintains the same intonation pattern: the left edge of a syntactic phrase corresponds to the left edge of the domain of a downstep. I argue that nevertheless the short-scrambling alternation is prosodically conditioned: it takes place in order to make maximal use of default sentence stress in encoding a particular disanaphoricity pattern. Section 4.1 shows that constraints on the syntax-prosody mapping derive the generalization that the short-scrambling alternation in Japanese is prosodically conditioned, but neither of the two orders is more marked than the other. In 4.2, I consider the possibility that the prosodic condition on Japanese short-scrambling reflects efficiency of sentence production, and discuss the lack of prosodic markedness with respect to Case morphology in Japanese. Section 4.3 shows that the prosodic and syntactic properties associated with the alternation provide supporting evidence for the Prosodic Phase Hypothesis and for the claim that structure is built incrementally in the syntax. Lastly in 4.4, I give a brief review of previous analyses of scrambled sentences, and compare them with the Prosodic Phase Hypothesis.

4.1. Prosodic Properties of Linearization in the Verbal Domain

4.1.1. Overview

This section shows that the Prosodic Phase Hypothesis in (1), as a general syntax-prosody mapping algorithm, accounts for the prosodic properties of the short-scrambling alternation in Japanese.

(1) The Prosodic Phase Hypothesis
A syntactic object is spelled out as a prosodic object.

In the short-scrambling alternation, the non-scrambled PP - NP - V order and the scrambled NP - PP - V order exhibit the same intonation pattern: the left edge of a syntactic phrase corresponds to the left edge of an Intonational Phrase (IntP). With respect to the linear precedence relation between the head V and a dependent, the non-scrambled PP - NP - V order forms mirror images with the canonical V - DP - PP order (C-order or non-shifted order) in English, and the scrambled NP - PP - V order, with the marked V - PP - DP order (M-order or shifted order) in English. I call the non-scrambled order in Japanese “Order-I”, and the scrambled order, “Order-II”, because there is no prosodic markedness associated with Japanese short-scrambling.

Although neither Order-I nor Order-II is more marked than the other, the choice of word order is not entirely optional, and determined by the locus of focus. If there is no narrow-focus involved, word order is optional, whereas if there is a narrow-focus, the order where the narrow-focus is left-adjacent to the V is chosen.

(2) Linearization in the verbal domain in Japanese
Let the Order-I = PP - NP - V with the prosodic pattern [\textit{IntP} PP] [\textit{IntP} NP - V], and the Order-II = NP - PP - V with the prosodic pattern [\textit{IntP} NP] [\textit{IntP} PP - V],

a. select Order-I or Order-II if there is no narrow-focus,
b. select Order-I if NP is (a part of) a narrow-focus, and
c. select Order-II if PP is (a part of) a narrow-focus.

The methodology that I employed in reaching this generalization consists of acceptability judgments of pronounced stimuli sentences. In order to examine interface properties of the short-
scrambling alternation, I designed and conducted a systematic elicitation that includes a judgment task as a test of how natural or unnatural stimuli sentences sound to native speaker of Japanese with particular focus interpretations, and with particular prosodic patterns. (See appendix to this chapter for the details of the pilot study.)

4.1.2. Broad-Focus Sentences

This section shows that broad-focus sentences with two different word orders, PP – NP – V and NP – PP – V, exhibit the same intonation pattern in Japanese.

4.1.2.1. Default Sentence Stress

By setting up a question that requires an answer with the whole sentence (IP) as a focus, we can elicit a broad-focus sentence with the default word order and with the default intonation pattern. In broad-focus sentences in Japanese, two dependents in the verbal domain may appear in PP – NP order or in NP – PP order. For example, in (3B), the two internal arguments of the V moratta ‘received’, the source –ni ‘from’ NP and the theme –o ‘Acc’ NP, appear in PP-NP order. In (4B), the two internal arguments appear in the opposite order.1

(3) A: Kinou nani-ga atta no? yesterday what-Nom happened Q ‘What happened yesterday?’

F₀ contour:

B: Mayumi-ga [PP Mamoru-ni] [NP nuigurumi-o] moratta yo.2
Mayumi-Nom Mamoru-from doll-Acc received yo
‘Mayumi received a doll from Mamoru’

(4) A: Kinou nani-ga atta tte?
‘What happened yesterday?’

F₀ contour:

B: Mayumi-ga [NP nuigurumi-o] [PP Mamoru-ni] moratta yo.
Mayumi-Nom doll-Acc Mamoru-from received yo
‘Mayumi received a doll from Mamoru’

Both (3B) and (4B) are felicitous broad-focus sentences, and exhibit the same intonation pattern. In particular, the left edge of a syntactic phrase is marked with a Low tone, and aligned with the left edge of the domain of downstep. This contrasts with English, where the whole sentence corresponds to the domain of an intonation contour in the default case. The constraint on the mapping of maximal prosodic objects in Japanese in (5b) formulates the default intonation pattern in Japanese:

1 It is controversial whether the particle –ni is a postposition projecting a PP, or a dative case-marker. I call nouns with source –ni “PPs” in order to distinguish it from “NPs” with the accusative case-marker –o. (See 4.2.2 for more discussion on the Case morphology in Japanese.) Note also that I call nominal arguments in Japanese NPs, not DPs, assuming that Japanese lacks determiners (cf. Fukui 1986).
2 See fn. 9 in chapter 2 for variable endings used in Japanese examples.
SO-PO mapping for maximal prosodic objects (= Intonational Phrases) [= (36) in 2.4.2.1]

a. English: Align (SO, R; PO, R) and Align (SO, L; PO, L)
where SO = clause (CP/IP)
b. Japanese: Align (SO, L; PO, L), where SO = phrase (XP)

In broad-focus sentences, default sentence stress is found in the syntactic phrase immediately left-adjacent to the verb in Japanese (Selkirk and Tateishi 1991, Nagahara 1994, Ishihara 2000). In particular, in the non-scrambled PP – NP – V order in (3B), default sentence stress falls on the N nuigurumi ‘doll’ in the theme NP, and in the scrambled NP – PP – V order in (4B), default stress falls on the N Mamoru in the source PP (as indicated by bolding). This conforms to the Default Sentence Stress Assignment in (6b), because the N that carries default sentence stress is the leftmost prosodic word in the rightmost IntP, which corresponds to the rightmost syntactic phrase, i.e. VP, in Japanese.

(6) Default Sentence Stress Assignment [= (41) in 2.4.2.1]
Default sentence stress is assigned to a prosodic word within the rightmost Intonational Phrase (= POₙ).
a. In English, it is assigned to the rightmost prosodic word in the POₙ.
b. In Japanese, it is assigned to the leftmost prosodic word in the POₙ.

The core part of the rule in (6) does not prevent a verb from carrying default sentence stress, and this is exactly what happens in some head-final languages, such as German, Dutch, Bengali, and Turkish (cf. chapter 5). At this point, I leave the observation in Japanese as a stipulation that default sentence stress is assigned to the leftmost prosodic word in the rightmost IntP, which is the PO that is spelled out finally (= POₙ).³

The focus projection mechanism in (7) predicts that not only IP-focus interpretations but also VP-focus interpretations are possible for a sentence with the default intonation pattern. The prediction is borne out, both with Order-I and Order-II. The VP₁-focus interpretation is possible with Order-I as is shown in (8B₁)-(8B₂), and with Order-II as is shown in (9B₁)-(9B₂).

(7) Focus projection [= (29a) in 2.4.1.2]
[Focus]ₙ can project to any node that dominates it, when [Focus] dominates the default sentence stress position.

(8) VP₁-focus
A: ‘What do you know about Mayumi?’
Order-I
B1: [VP₁ Mamoru-ni nuigurumi-o moratta yo].
     Mamoru-from doll-Acc received yo
     ‘(she) received a doll from Mamoru’

B2: Kanojo-wa [VP₁ Mamoru-ni nuigurumi-o moratta yo].
     she-Top Mamoru-from doll-Acc received yo
     ‘She received a doll from Mamoru’

(9) VP₁-focus
A: ‘What do you know about Mayumi?’
Order-II

³ According to Ladd (1996:187-195), some languages (e.g. Italian and Catalan) but not others (e.g. English, German and Dutch) are sensitive to the distinction between “arguments” (e.g. DP/NP) and “predicates” (e.g. V), as to the locus of default sentence stress. Japanese seems to be sensitive to this distinction because default sentence stress never falls on the verb.
As is illustrated in the structures in (10) below, the VP1 node dominates the NP with [Focus]s0, which in turn dominates the prosodic word in the default sentence stress position. Therefore, the VP1 can inherit [Focus]s0 by focus projection. (We will look at VP2-focus examples in 4.1.3.1.)

(10) a. Order-I

\[
\begin{array}{c}
\text{IP[F]s0} \Rightarrow (3B) \\
\text{(kanojo wa)} \\
\text{VP1[F]s0} \Rightarrow (8B) \\
\text{Mamoru ni} \\
\text{NP[F]s0} \Rightarrow (9B) \\
\text{nuigurumi o} \\
\text{N} \\
\end{array}
\]

b. Order-II

\[
\begin{array}{c}
\text{IP[F]s0} \Rightarrow (4B) \\
\text{(kanojo wa)} \\
\text{VP1[F]s0} \Rightarrow (9B) \\
\text{nuigurumi o} \\
\text{PP[F]s0} \Rightarrow (9B) \\
\text{Mamoru ni} \\
\text{N} \\
\end{array}
\]

Comparing (B1) and (B2) in (8)-(9), the (B1) sentences, where the anaphoric argument is deleted, is preferred over the (B2) sentences, where it is retained as a pronominal, kanojo-wa ('she-Top'). This contrasts with English, where anaphoric arguments are pronominalized rather than deleted.

The pitch peak of default sentence stress in Japanese is usually slightly higher than the lexical stress found in the preceding phrase, despite the general downstep of the F0 contour throughout an utterance (Selkirk and Tateishi 1991). This suggests that default sentence stress in Japanese is not just the rightmost lexical stress in a sentence, but the main sentence stress prosodically. Remember that in English, default sentence stress is perceived as only relatively more prominent in partial broad-focus sentences. This is because anaphoric elements are not deleted but only deaccented in English, which yields a prosodic contrast with the focused element that contains default sentence stress. See the table in (11), which summarizes the way English encodes (partial) broad-focus prosodically:

(11) Prosodic encoding of (partial) broad-focus in English [= (10) in 3.1.2.1]

<table>
<thead>
<tr>
<th>Disanaphoricity</th>
<th>Prosodic process</th>
<th>Deaccenting</th>
<th>Normal prominence</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Broad-focus</td>
<td>[IP[F]s0]</td>
<td>\checkmark</td>
<td>\checkmark</td>
</tr>
<tr>
<td>b. Partial broad-focus</td>
<td>[IP [XP[F]s0] ]</td>
<td>\checkmark</td>
<td>\checkmark</td>
</tr>
</tbody>
</table>

In Japanese, anaphoric arguments are deleted whenever their informational content is recoverable from discourse context as in (8B1) and (9B1). Even in such cases, default sentence stress in Japanese is
always prosodically prominent. When anaphoric arguments are retained as pronominals as in (8B2) and (9B2), they are not deaccented and carry lexical stress. The way Japanese encodes (partial) broad-focus prosodically is summarized in (12), and the prosodic processes employed in encoding disanaphoricity in Japanese and English are shown in the Disanaphoricity Scale in (13).

(12) Prosodic encoding of (partial) broad-focus in Japanese

<table>
<thead>
<tr>
<th>Disanaphoricity pattern</th>
<th>Prosodic process</th>
<th>Deletion</th>
<th>Deaccenting</th>
<th>Normal prominence</th>
<th>Extra prominence</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Broad-focus (e.g. (3)-(4))</td>
<td>[imfio]</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>b. Partial broad-focus w/ deletion (e.g. (8B1)-(9B1))</td>
<td>[p{xfio}</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>b'. partial broad-focus w/o deletion (e.g. (8B2)-(9B2))</td>
<td>[p [xfio]</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

(13) Disanaphoricity Scale

<table>
<thead>
<tr>
<th>Anaphoric</th>
<th>(i) deletion</th>
<th>(ii) deaccenting</th>
<th>(iii) normal prominence (i.e. lexical stress)</th>
<th>(iv) extra prominence</th>
</tr>
</thead>
<tbody>
<tr>
<td>focused/disanaphoric</td>
<td>(i) null NP</td>
<td>(ii) ---</td>
<td>(iii) pronominal</td>
<td>(iv) default sentence stress</td>
</tr>
</tbody>
</table>

Thus, Japanese differs from English in how it encodes disanaphoricity prosodically. Japanese makes extensive use of null-NPs for anaphoric arguments (= (13i)), whereas English normally uses deaccenting (including pronominalization) for them (= (13ii)).

4.1.2.2. Absence of the Distinction between Canonical and Marked Orders

In Japanese, broad-focus sentences can be either Order-I or Order-II, and the preference for one order over the other is rather mild (if present at all). For example, with a six-point scale (6 = completely natural; 1 = completely unnatural), five out of five native speakers of Japanese judge the non-scrambled sentence in (3B), *Mayumi-ga Mamoru-ni nuigurumi-o moratta yo* (‘Mayumi-Nom Mamoru-from doll-Acc received’), as 6, and the scrambled sentence in (4B), *Mayumi-ga nuigurumi-o Mamoru-ni moratta yo*, as two 6s, two 5s, and one 4 (see appendix to this chapter for the details of acceptability judgments).4

Notice that there is yet another phrase in the domain to the left of the V: subject NP. However, a broad-focus sentence where the theme NP precedes the subject NP is infelicitous.

(a) A: ‘What happened yesterday?’
(b) B: #nuigurumi-o Mayumi-ga Mamoru-ni moratta yo.

Among three native speakers of Japanese I consulted, one judged the (i-B) sentence as 4, one 2, and one 1, while all of them judged the corresponding non-scrambled sentence in (3B) to be completely natural (= 6). I limit the scope of my investigation of word order

---

4 Notice that there is yet another phrase in the domain to the left of the V: subject NP. However, a broad-focus sentence where the theme NP precedes the subject NP is infelicitous.

(i) A: ‘What happened yesterday?’
(ii) B: #nuigurumi-o Mayumi-ga Mamoru-ni moratta yo.

Among three native speakers of Japanese I consulted, one judged the (i-B) sentence as 4, one 2, and one 1, while all of them judged the corresponding non-scrambled sentence in (3B) to be completely natural (= 6). I limit the scope of my investigation of word order.
Let us look at more examples with different verbs. In (14)-(17), the (B1) sentences represent Order-I, where a theme accusative NP immediately precedes the V, and the (B2) sentences represent Order-II, where another argument or adjunct phrase intervenes between the accusative NP and the V.

(14) A: Kinou nani-ga atta no?
   'What happened yesterday?'
   
   Order-I
   B1: Mamoru-ga Mayumi-ni nuigurumi-o ageta rasii yo.
   'They say that Mamoru gave a doll to Mayumi'

(15) A: 'What happened yesterday?'
   Order-I
   B1: Ken-ga go-fun-de jukko-no California roll-o tabeta
   Ken-Nom five-minute-in ten-of California roll-Acc ate
   'They say that Ken ate ten California rolls in five minutes'

(16) A: 'What happened yesterday?'
   Order-I
   B1: Megu-ga titioya-ni karesi-o syookaisita rasii yo.
   Meg-Nom father-to BF-Acc introduced they.say yo
   'They say that Meg introduced her boyfriend to her father'

(17) A: Saku-ban-wa nani-o sita no?
   'What did you do last night?'
   Order-I
   B1: Jeriko-biiti-de supein-hanabi-o mita yo.
   Jericho-beach-at Spanish-firework-Acc saw yo
   '(I) saw Spanish fireworks at Jericho beach'

The naturalness judgments of the (B) sentences in (14)-(17) by four native speakers of Japanese are shown in the table (18):

(18) Naturalness judgment (6 = completely natural; 1 = completely unnatural)

alternations within the verbal domain (i.e. V and its dependents), and exclude the word order in (i-B) from the main discussion.
Unlike English, Order-II may be only slightly less acceptable than the corresponding Order-I, although there is an implicational relation that Order-II is never more acceptable than Order-I.

Furthermore, the difference in the frequency of Order-I versus Order-II in Japanese is not as big as that of C-order versus M-order in English. According to Hawkins (1994), the number of PP – NP – V sequences in a corpus of 153 was 109 (73%), and the number of NP – PP – V sequences was 44. Comparing this with Hawkins’ corpus research in English, where the number of V – NP – PP sequences was 458 (95%), and V – PP – NP, 22 out of 480 (see 3.2.2), we notice that there is a greater difference in the frequency of C-order versus M-order in English than that of Order-I versus Order-II in Japanese:

\[
\begin{array}{c|cccc}
\text{Order-I} & \text{AT} & \text{TT} & \text{NH} & \text{MK} \\
\hline
(14) & 6 & 6 & 4 & 6 \\
(15) & 6 & 6 & 6 & 5 \\
(16) & 6 & 6 & 6 & 5 \\
(17) & 6 & 6 & 6 & 5 \\
\hline
\end{array}
\]

(19) Frequency of two orders (based on Hawkins 1994:153)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>458 (95%)</td>
<td>22</td>
<td>480</td>
</tr>
<tr>
<td></td>
<td>109 (73%)</td>
<td>44</td>
<td>153</td>
</tr>
</tbody>
</table>

In Japanese short-scrambling, the change of word order does not affect the prosodic pattern either (see e.g. (3)-(4)). Therefore, neither order is prosodically more marked than the other. Furthermore, the prosodic weight effect observed in English shifted sentences is absent in Japanese. For example, Order-II in (20B1)-(20B3) is all equally acceptable regardless of the number of prosodic words contained in the NP:

(20) A: Kinou rani-ga atta no?
‘What happened yesterday?’

B1: Megu-ga \[ NP go-hyaku gojuu kanada-doru ijou mo-no taikin-o \] Linguistlist-ni kifusita rasii yo.
Meg-Nom five-hundred fifty Canadian-dollar more-than even-of big-money-Acc Linguistlist-to donated they.say yo
‘They say that Meg donated big money such as more than five hundred and fifty Canadian dollars to Linguistlist’

B2: Megu-ga \[ NP go-hyaku kanada-doru-o \] Linguistlist-ni
Meg-Nom five-hundred Canadian-dollar-Acc Linguistlist-to
kifusita rasii yo.
donated they.say yo
‘They say that Meg donated five hundred Canadian dollars to Linguistlist’

B3: Megu-ga \[ NP gojuu-doru-o \] Linguistlist-ni kifusita rasii yo.
Meg-Nom fifty-dollar-Acc Linguistlist-to donated they.say yo
‘They say that Meg donated fifty dollars to Linguistlist’
A large number of prosodic words in an NP does not affect word order either:

(21) A: Kinou nani-ga atta no?
   ‘What happened yesterday?’

Order-I
B1: Megu-ga Midori-ni [NP suu-jitu mae-ni PTA-no menbaa-ga
    Meg-Nom Midori-to several-day ago-at PTA-of member-Nom
    sakuseisita Taro-nituite-no houkokusyo-o] watasita rasi yo.
    completed Taro-about-of report-Acc gave they.say yo
   ‘They say that Meg gave the report on Taro which the PTA members completed
    several days ago to Midori’

Order-II
B2: Megu-ga [NP suu-jitu mae-ni PTA-no menbaa-ga sakuseisita Taro-nituite-no

In the non-scrambled sentence in (21B1), the theme NP that contains seven prosodic words intervenes between the goal PP, *Midori-ni* ‘Midori-to’, and the V *watasita* ‘gave’. The sentence is still as acceptable as the scrambled sentence in (21B2), where the NP appears to the left of the PP.\(^5\)

The absence of a prosodic weight effect in Japanese broad-focus sentences contrasts with the presence of a right-edge heaviness effect in English broad-focus sentences: in English, the shifted V – PP – DP order is clearly more marked than the non-shifted V – DP – PP order, in that the rightmost DP must be an independent IntP. Remember that I formulated the prosodic property of English M-order in terms of the SO-PO Mapping Condition on specific to English PPs:

(22) The SO-PO Mapping Condition on English PPs \([= (18) in 3.1.2.2]\)
    Align (SO, R; PO, R), where SO = a PP dependent, and PO = Intonational Phrase.

The Mapping Condition in (22) derives the generalization that the rightmost DP in the marked V – PP – DP order in English forms an IntP on its own. The condition in (22) subsumes an adjacency requirement on Case assignment in English: a DP dependent canonically appears adjacent to the head V that assigns Case to it, although the adjacency requirement is violable and overridden when M-order is licensed by the Prosodic Weight Condition.

(23) Case Adjacency \([= (14) in 3.1.2.2]\)
    DPs must be adjacent to the Case-assigning verb.

(24) The Prosodic Weight Condition \([= (34) in 3.1.2.3]\)
    PO\(_n\) (= the rightmost Intonational Phrase) must be prosodically heavy by containing
    a. a larger number of prosodic words than PO\(_{n-1}\) (= the second rightmost
      Intonational Phrase), or
    b. a prosodic word with extra prosodic prominence.

In Japanese, Case is assigned by morphologically overt case-markers or postpositions (see fn.1), and therefore NP dependents need not be adjacent to the head V.\(^6\) The lack of an adjacency requirement

---
\(^5\) The absence of a prosodic weight effect in (20)-(21) is confirmed by three native speakers of Japanese I consulted. However, Hawkins (1994:137-154) and Yamashita and Chang (2001) report a “left-edge heaviness effect” in Japanese in performance data such as text counts and production experiments, respectively. According to their observation, a heavy phrase with a large number of lexical items is placed in front of a light phrase in Japanese. Such an effect is absent in acceptability judgments by native speakers for unknown reasons (as is pointed out by e.g. Shiotara 2001, to appear, Tokizaki 2001, Ernst 2002:251), and the weight effect in performance data in Japanese remains to be explained.

\(^6\) See Miyagawa (1997:8-9) for related discussion on a base-generation approach to Japanese short-scrambling and a Case requirement on NPs.
between V and NPs in Japanese is consistent with the observation that the short-scrambling alternation
between PP – NP – V (Order-I) and NP – PP – V (Order-II) in Japanese is freer than the heavy NP shift
 alternation between V – DP – PP (C-order) and V – PP – DP (M-order) in English, and as free as CP-shift, PP-shift and AdvP-shift alternations in English where no DP dependent is involved (see 3.3.1.1 for
English examples).

In 3.1.2.2, I suggested that in English the two conditions in (22) and (23) compete with each other
in an optimality theoretic way, and the satisfaction of the highly ranked Prosodic Mapping Condition in
(22) is given priority over the satisfaction of the Case Adjacency condition in (23). This is what is
happening in M-order in English. If we apply the optimality theoretic reasoning to Case assignment in
Japanese, the Case Adjacency condition should be ranked lower than Case assignment by
morphologically overt case-markers, and hence the adjacency effect is not visible in Japanese. (See 4.2.2
for further discussion on morphological Case in Japanese.)

Thus, the short-scrambling alternation in broad-focus sentences in Japanese differs from the heavy
NP shift alternation in English, in that the short-scrambling alternation is not prosodically conditioned:
the scrambled NP – PP – V order (Order-II) is as available as the non-scrambled PP – NP – V order (Order-I),
and the two orders exhibit the same prosodic pattern.

4.1.2.3. Derivations

Since Order-I and Order-II exhibit the same intonation pattern in broad-focus sentences, the Prosodic
Phase Hypothesis predicts that they should go through the same derivational steps. Let us look at how
broad-focus sentences with Order-I or Order-II are derived.

(25) A: Kinou nani-ga atta no?
yesterday what-Nom happened Q
‘What happened yesterday?’

Order-I
B1: [imp Mayumi-ga] [imp Mamoru-ni] [imp nuigurumi-o] moratta yo.
Mayumi-Nom Mamoru-from doll-Acc received yo
‘Mayumi received a doll from Mamoru’ [= (3B)]

Order-II
B2: [imp Mayumi-ga] [imp nuigurumi-o] [imp Mamoru-ni] moratta yo.
Mayumi-Nom doll-Acc Mamoru-from received yo
‘Mayumi received a doll from Mamoru’ [= (4B)]

The derivation of the non-scrambled sentence in (25B1) is illustrated in (26) below. Syntactic
structure is incrementally built from left to right via Merge, and spelled out into the phonological
component. In broad-focus sentences where no focal stress is involved, the left edge of a syntactic phrase
(XP) is aligned with the left edge of an Intonational Phrase (IntP) in Japanese, and spell-out applies three
times accordingly. The Default Sentence Stress Assignment in (6b) assigns default sentence stress to the
leftmost prosodic word, the N nuigurumi ‘doll’, in the rightmost IntP (= POa).


Numeration {Mayumi, ga, Mamoru, ni, nuigurumi, o, moratta}

syntactic structure
The derivation of the scrambled sentence in (25B2) proceeds in the same way as the corresponding non-scrambled sentence. In this case, default sentence stress is assigned to the N, *Mamoru*, because this is the leftmost prosodic word in the PO.

(27) Derivation of the example (25B2): \[\text{Mayumi-ga} \text{ nuigurumi-o} \text{ Mamoru-ni moratta yo}\] Numeration \{Mayumi, ga, Mamoru, ni, nuigurumi, o, moratta\} \[\text{syntactic structure}\] Prosodic Weight Condition in (24), which explains the prosodic property of word order alternations in English, is satisfied in both sentences, because the PO is prosodically heavier than the preceding IntP (= PO,.) by containing a verb.

In both (25B1) and (25B2), the IP-focus interpretation is not the only possible interpretation. The NP (theme), VP1 (= PP+NP+V), and VP2 (= NP+V) focus interpretations are also available in the non-scrambled sentence in (25B1), and PP (source), VP1 (= NP+PP+V), and VP2 (= PP+V) focus interpretations are also available in the scrambled sentence in (25B2). (See (8)-(9) for VP1-focus examples.) This is guaranteed by the syntactic feature [Focus] that can be freely introduced in the Numeration and assigned to the N in the default sentence stress position. By focus projection (defined as in (7)), all the syntactic nodes that dominate the feature [Focus] can be interpreted as focused.

Thus, the non-scrambled sentence in (25B1) and the scrambled sentence in (25B2) start with the same Numeration and go through the same derivational steps: syntactic structure is built from left to right via Merge, paired with POs, and spelled out into the phonological component by IntPs. Therefore, no prosodic markedness is associated with Order-I or Order-II in broad-focus sentences in Japanese.

Order-I and Order-II, however, can be differentiated with respect to which element can be interpreted as focused. In other words, the factor that may differentiate between Order-I and Order-II is the possible focus interpretations associated with them (which is called the "focus set"): 115
The focus set associated with broad-focus sentences

a. Order-I

\[
\begin{align*}
\text{Focus set:} & \quad \{\text{IP, VP1, VP2, NP}\} \\
& \quad \text{(default sentence stress)}
\end{align*}
\]

b. Order-II

\[
\begin{align*}
\text{Focus set:} & \quad \{\text{IP, VP1, VP2, PP}\} \\
& \quad \text{(default sentence stress)}
\end{align*}
\]

In particular, NP+V (=VP2 in (28a)) and NP-focus interpretations should be available only with Order-I, and PP+V (=VP2 in (28b)) and PP-focus interpretations should be available only with Order-II. This is because focus projection is possible only from the feature \([\text{Focus}]_S\) which dominates the default sentence position. With this in mind, let us look at the word order alternation in narrow-focus sentences, where a focus does not necessarily occupy the default sentence stress position.

4.1.3. Narrow-Focus Sentences

In this section, I will show that the word order alternation in Japanese is not associated with a prosodic markedness effect in narrow-focus sentences either. Unlike in broad-focus sentences, however, the presence of a narrow-focus affects linearization of verbal dependents in narrow-focus sentences, and the choice between Order-I and Order-II is prosodically conditioned. In particular, the order where (a part of) a narrow-focus sits in the default sentence stress position is preferred over the other order. I will formulate this effect as a “prosodic economy condition”.

4.1.3.1. Avoidance of Focal Stress and the Prosodic Economy Effect

Since Japanese uses deletion of anaphoric elements when possible, narrow-focus sentences usually consist only of focused elements. This is exemplified in the (B1) sentences given in (29) below. (A narrow-focus is indicated by underscore.) For the purpose of comparison with English, which does not delete anaphoric DPs but pronominizes them instead, we will also look at the prosodic pattern of narrow-focus sentences which retain anaphoric elements as pronominals. This is exemplified in the (B2) sentences in (29).

(29) a. NP[subject]-focus

A: ‘Who received a doll from Mamoru?’
B1: \[\text{hap Mayumi}.\]
>
B2: \[\text{hap MAYUMI-ga kare-ni sore-o moratta yo}.\]
‘Mayumi received it from him’

b. PP[source]-focus
A: ‘From whom did Mayumi receive a doll?’
B1: [imp_Mamoru].

B2: [imp Kanojo-wa][imp MAMORU-ni sore-о moratta yo].
    ‘She received it from Mamoru’

A: ‘What did Mayumi receive from Mamom?’
B1: [imp Nuigurumi].
    ‘A doll’

B2: [imp Kanojo-wa][imp kare-ni][imp nuigurumi-o moratta yo].
    ‘She received a doll from him’

B3: [imp Kanojo-wa][imp kare-ni][imp MUGURUMI-o moratta yo].

c. NP[theme]-focus
A: ‘What did Mayumi receive from Mamoru?’
B1: [imp Nuigurumi].
‘A doll’

B2: [imp Kanojo-wa][imp kare-ni][imp nuigurumi-o moratta yo].
‘She received a doll from him’

When all anaphoric elements are deleted as in (B1), the remaining focused element is pronounced only with normal lexical stress and without any extra prosodic prominence (e.g. focal stress).

In the (B2) sentences, where anaphoric elements are retained as pronominals, the prosodic pattern of a sentence varies depending on the locus of focal stress (indicated by upper case letters). The presence of a narrow-focus on a given phrase boosts the pitch peak of the lexical High tone (indicated by an upward arrow), and the pitch peak of focal stress is higher than that of default sentence stress (Pierrehumbert and Beckman 1988, Ishihara 2000:148). In addition, the elements following a prosodic word with focal stress are all deaccented to the right edge of the sentence (indicated by small letters), with the following domains of downstep being incorporated into one intonational domain. In such a case, default sentence stress is suppressed (Ishihara 2000:149). Therefore, a narrow-focus, when it is realized as focal stress, has the effect of eliminating all following intonational breaks within a sentence (Ishihara 1994, Ishihara 2000). Following Ladd (1996), I call this change in prosodic structure “prosodic dephrasing” (p.196).

In the NP[theme]-focus sentences (29c-B2) and (29c-B3), the narrowly focused NP occupies the default sentence position. In (B2), the NP carries only default sentence stress, and therefore the prosodic pattern of the sentence is the same as that of a broad-focus sentence. On the other hand, in (B3), the
narrowly focused NP carries focal stress, whose pitch peak is higher than that of default sentence stress. The intonation pattern of the two sentences is the same, but the two sentences differ in the height of the pitch peak on the N *nuigurumi* ‘doll’. In these cases, the prosodic pattern in (B2) is preferred over the pattern in (B3). I argue that this is because default sentence stress in Japanese is already prosodically distinct, independently of focus (cf. 4.1.2.1), and hence adding an extra prosodic prominence is disfavored. This effect is formulated as the Prosodic Economy Condition specific to Japanese:

\[(30)\] **The Prosodic Economy Condition for Japanese**

Japanese employs default sentence stress rather than focal stress whenever possible, in order to prosodically encode disanaphoricity.

The Prosodic Economy Condition chooses the (B2) pattern over the (B3) pattern in (29). I assume that a violation of the Prosodic Economy Condition in (30) degrades, but not totally rules out, the sentence.

The Prosodic Economy Condition for Japanese is evaluated by a pairwise comparison between two syntactic objects in the same way as the Prosodic Weight Condition (cf. 3.1.2.3): once a syntactic object with [Focus]\_so is merged, then the V must be merged immediately after it. This way, it is guaranteed that a focus is immediately left-adjacent to the V occupying the default sentence stress position, and the application of the Prosodic Economy Condition is localized to a pair of syntactic objects, i.e. XP with [Focus]\_so and the V.\(^7\)

The prosodic economy effect in Japanese contrasts with the prosodic pattern in English, which employs focal stress to differentiate a narrow-focus from a partial broad-focus.

\[(31)\] **Prosodic encoding of disanaphoricity in English when narrow-focus = partial broad-focus \([= (54) \text{ in } 3.1.4.1]\)**

When a narrow-focus is in the default sentence stress position, English employs focal stress rather than default sentence stress, in order to prosodically encode disanaphoricity.

Therefore, what is meant by “economy” in the Prosodic Economy Condition for Japanese in (30) should be economy from a production perspective: prosodic prominence with a lower pitch peak demands less effort in production than prosodic prominence with a higher pitch peak. In contrast, the statement in (31) is regarded as economy from a perception perspective. For the moment, I regard the Prosodic Economy Condition in (30) as specific to Japanese relative to English.\(^8\) (I will revisit the Prosodic Economy Condition for Japanese and consider it from a performance perspective in 4.2.1.)

In sum, when a narrow-focus can be reanalyzed as a partial broad-focus, Japanese uses the default intonation pattern whereas English employs focal stress in encoding disanaphoricity. This difference seems to be correlated with the prosodic difference in default sentence stress between the two languages: unlike Japanese, the default sentence stress in English is not always distinct from normal lexical stress (cf. 3.1.2.1), and hence the default sentence stress in Japanese and focal stress in English belong to the same position of the Disanaphoricity Scale in (32) \((= \text{iv})\). In order to accommodate the assignment of an extra higher prosodic prominence, I accordingly expand the Disanaphoricity Scale we saw in (13). In the expanded scale in (32) below, the process in (v) indicates the prosodic process of extra-higher prominence, which is exemplified by focal stress in Japanese.

\[(32)\] **Disanaphoricity Scale (expanded from (13))**

---

\(^7\) Thanks to Henry Davis for insightful suggestions and discussion on this.

\(^8\) As I noted in 3.1.4.1, Neeleman and Reinhart (1998) propose a production-based economy condition in their analysis of Dutch scrambling. See also Choi (1999:101-103) for related discussion in her analysis of Korean scrambling.
e.g. In English,
(i) ---
(ii) pronominal
(iii) default sentence stress
(iv) focal stress
(v) ---

Thus, it is when a narrow-focus is not in the default sentence stress position (as in (29a), (29b), and (29d)) that the presence of a narrow-focus may create a different prosodic pattern in Japanese. In these cases, the left edge of a prosodic word with focal stress starts a new domain of downstep, and the domain of downstep spreads toward the right edge of the sentence. This effect is formulated as the SO-PO mapping constraint in (33a), and the Japanese-specific Prosodic Condition in (33b):^9

\[
\begin{align*}
&\text{(33a)} \quad \text{SO-PO mapping for maximal prosodic objects (} = \text{Intonational Phrases) with [Focus]} \\
&\text{in Japanese, Align (SO, L; PO, L), where SO is with [Focus]_{SO}.} \\
&\text{(33b)} \quad \text{The Prosodic Condition on Japanese [Focus]_{PO} (} = \text{in 2.4.3.1)} \\
&\text{In Japanese, [Focus]_{PO} must be in the rightmost Intonational Phrase (} = \text{PO}_{n}).
\end{align*}
\]

The SO-PO mapping constraint in (33a) specifies that the left edge of a prosodic word with focal stress starts a new Intonational Phrase (IntP), and the Prosodic Condition in (33b) specifies that the IntP extends up to the end of the sentence in Japanese.

In Japanese, word order does not affect the intonation pattern of narrow-focus sentences, and Order-II exhibits the same prosodic pattern as Order-I. Look at Order-II examples:

(34) a. NP[subject]-focus
A: ‘Who received a doll from Mamoru?’
B1: [nlp Mayumi].

\[
\begin{align*}
&\text{(i) Taro-ga \ [cp DARE-ga \ tatao \ kata \ ka] \ Jiro-ni \ itta \ no?} \\
&\text{Taro-Nom \ who-Nom \ book-Ace \ bought \ that \ Jiro-Dat \ said \ Q} \\
&\text{‘Who did Taro say to Jiro bought a book?’}
\end{align*}
\]

\[
\begin{align*}
&\text{(ii) Taro-ga \ [cp DARE-ga \ tatao \ kata \ ka] \ Jiro-ni \ itta \ no?} \\
&\text{Taro-Nom \ who-Nom \ book-Ace \ bought \ Q \ Jiro-Dat \ said \ Q} \\
&\text{‘Did Taro say to Jiro who bought a book?’}
\end{align*}
\]

This suggests that there is an intonational break after the interrogative complementizer in (ii), but not after the declarative complementizer in (i). Then, the prediction is that it is possible to assign focal stress to Jiro-ni ‘Jiro-Dat’ in (ii), but not in (i), which seems correct to me.

Kitagawa and Fodor (2002) point out that deaccenting after the interrogative complementizer is disfavored, but possible when the wh-phrase takes a scope over the matrix clause. I leave further investigation of this issue for future research.

---

^9 In addition to focus, the domain of deaccenting seems to be sensitive to other semantic factors such as the type of complementizers. According to Ishihara (2000:155, fn 17), in Japanese the declarative complementizer to ‘that’ does not delimit deaccenting and the elements following the wh-phrase dare-ga ‘who-Nom’ with focal stress are all deaccented (indicated by small letters), as in (i). In contrast, the interrogative complementizer ka ‘Q’ defines the endpoint of deaccenting and the elements following the interrogative complementizer are not deaccented, as in (ii):

\[
\begin{align*}
&\text{(i) Taro-ga \ [cp DARE-ga \ tatao \ kata \ ka] \ Jiro-ni \ itta \ no?} \\
&\text{Taro-Nom \ who-Nom \ book-Ace \ bought \ that \ Jiro-Dat \ said \ Q} \\
&\text{‘Who did Taro say to Jiro bought a book?’}
\end{align*}
\]

\[
\begin{align*}
&\text{(ii) Taro-ga \ [cp DARE-ga \ tatao \ kata \ ka] \ Jiro-ni \ itta \ no?} \\
&\text{Taro-Nom \ who-Nom \ book-Ace \ bought \ Q \ Jiro-Dat \ said \ Q} \\
&\text{‘Did Taro say to Jiro who bought a book?’}
\end{align*}
\]

This suggests that there is an intonational break after the interrogative complementizer in (ii), but not after the declarative complementizer in (i). Then, the prediction is that it is possible to assign focal stress to Jiro-ni ‘Jiro-Dat’ in (ii), but not in (i), which seems correct to me.
The (B1) examples in (34) are exactly the same as the (B1) examples in (29). In these cases, the anaphoric elements are all deleted and there arises no issue of word order, because there is no distinction between Order-I and Order-II.

The scrambled (B2) sentences in (34) exhibit the same prosodic pattern as the corresponding non-scrambled (B2) sentences in (29). When a narrow-focus occupies the default sentence position as in (34c), the prosodic pattern in (34c-B2), where the narrow-focus PP carries only default sentence stress, is preferred over the pattern in (34c-B3), where the PP carries focal stress. This was also the case with Order-I (see (29c-B2) versus (29c-B3)), and formalized by the Prosodic Economy Condition for Japanese
To sum up so far, Order-I and Order-II exhibit the same intonation pattern in narrow-focus sentences in Japanese, and no prosodic markedness is associated with Order-I or Order-H. Unlike in broad-focus sentences, however, the Prosodic Economy Condition for Japanese determines word order: when the theme NP is a narrow-focus, the non-scrambled PP – NP – V order (Order-I) is preferred over the scrambled NP – PP – V order (Order-II), because the former does not have to employ focal stress and can encode an NP-focus interpretation as a partial broad-focus case. This is illustrated in (35) below. When the source PP is a narrow-focus, Order-II is preferred over Order-I, because in this case, Order-II does not have to employ focal stress in order to encode a PP-focus interpretation. This is illustrated in (36).

(35) NP[theme]-focus  
Order-I [= (29c-B3)]  
a. [[intP Kanojo-wa][intP kare-ni] nuigurumi-o moratta yo].  
she-Top he-from doll-Acc received yo  
'She received a doll from him'  
>  
Order-II [= (34b-B2)]  
b. [[intP Kanojo-wa][intP NUIGURUMI-o kare-ni] moratta yo].  
she-Top doll-Acc he-from received yo  
'She received a doll from him'

(36) PP[source]-focus  
Order-II [= (34c-B3)]  
a. [[intP Kanojo-wa][intP sore-o] Mamoru-ni moratta yo].  
she-Top it-Acc Mamoru-from received yo  
>  
Order-I [= (29b-B2)]  
b. [[intP Kanojo-wa][intP MAMORU-ni sore-o] moratta yo].  
she-Top Mamoru-from it-Acc received yo  
'She received it from Mamoru'

Let us move on to double narrow-focus sentences. The Prosodic Condition on Japanese [Focus]PO in (33b) predicts that occurrences of prosodic prominence must be adjacent in Japanese. If they are non-adjacent, all elements following the first focal stress should be deaccented and incorporated in one IntP, and therefore the second focal stress cannot be realized.

An independent piece of evidence for the deaccenting phenomenon in Japanese is found in wh-sentences. Ishihara (2000) shows that wh-words in Japanese, which always carry extra prosodic prominence as a lexical property, “prefer to be adjacent to each other within a clause” (p.165). Look at the following examples:

(37) Order-II  
a. DARE-ga NANI-o Hanako-ni ageta no?  
who-Nom what-Acc Hanako-to gave no?  
'Who gave what to Hanako?'

Notice that the prosodic dephrasing phenomenon captured by the Prosodic Condition on Japanese [Focus]PO in (33b) has the effect of conforming to the Prosodic Weight Condition in (24): since a prosodic word with focal stress ends up being in the rightmost IntP (= PO), the PO is always heavier than the preceding IntP (= PO1), if any. However, the Prosodic Weight Condition should not be a motivation for dephrasing in Japanese, because the Prosodic Weight Condition is also satisfied in sentences with the default intonation pattern in Japanese (cf. 4.1.2.3). (I am indebted to Scott Shank for drawing my attention to the relation between the Prosodic Weight Condition and the dephrasing phenomenon in Japanese.)
In (37a), the wh-words, the subject NP DARE-ga ‘who-Nom’ and the theme NP NANI-o ‘what-Acc’ are adjacent to each other. In (37b), the goal PP Hanako-ni ‘Hanako-to’ intervenes between the two wh-phrases. Native speakers prefer the sentence in (37a) over (37b). The example in (37c) exhibits the deaccenting phenomenon in that the elements following the first wh-phrase with extra prosodic prominence are deaccented. However, this is not acceptable because the second wh-phrase nani-o must be realized with extra prosodic prominence, but it is not. In a nutshell, when there is more than one element with extra prosodic prominence, a particular word order must be chosen so that two prosodic words with prosodic prominence can become adjacent to each other.

The same observation holds for double, non-wh, narrow-focus sentences. When there are two prosodic words with focal stress within a sentence, they must be adjacent to each other (Ishihara 2000). For example, when the theme NP and the V are narrowly focused and carry focal stress, the non-scrambled PP - NP - V order (Order-I) is chosen over the scrambled NP - PP - V order (Order-II). This is shown in the contrast between (B2) and (B3) in (38):

(38) NP[theme]+V-focus
A: ‘What happened to Mamoru and Mayumi?’

Order-I

B1: [Mayumi-ga] [Mamoru-ni] nuigurumi-o moratta rasii].
They say that Mayumi received a doll from Mamoru

Order-II

B2: [Mayumi-ga] [Mamoru-ni] MUIGURUMI-o MORATTA rasii].
They say that Mayumi received a doll from Mamoru

B3: [#Mayumi-ga] [MUIGURUMI-o Mamoru-ni MORATTA rasii].
They say that Mayumi received a doll from Mamoru

---

11 Bear in mind that the term “double” narrow-focus is used for expository purposes only, referring to sentences where the focused elements are distributed on more than one syntactic object. In fact, when two narrow-foci, XP and V, are adjacent to each other, it can be a single VP-focus as an instance of focus projection.

12 Related to this point, Matsumoto (2000), based on an experimental study of intonation and discourse, argues that Japanese tends to have only one informationally new element per intonational unit. Grouping two occurrences of focal stress together as in (38B2) is consistent with Matsumoto’s argument, because it has the effect of achieving one string of narrow-focus within one IntP. See fn. 15 in chapter 2 for Donati and Nespor’s (2003) prosodic view of the same phenomenon.
Note, however, that the most acceptable prosodic pattern is the default intonation pattern in (38B1), which does not employ any focal stress. In this case, the NP+V-focus interpretation is possible due to focus projection of [Focus]$_{S}$ from the N *nuigurumi ‘doll’ in the default sentence stress position, and neither the NP nor the V has to carry focal stress. The Prosodic Economy Condition in (30) correctly captures the preference of (38B1) over (38B2) or (38B3).

When the source PP and the V both carry focal stress, Order-II in (39B2) is chosen over Order-I in (39B3):

(39) PP[source]+V-focus
A: ‘Mayumi is holding a doll carefully. Do you know why?’

B1: [IntP Mamoru-ni moratta rasii].
   ‘They say that (she) received (it) from Mamoru’
> Order-II

B2: [Inf Kanojo-wa [Inf are-o] [Inf MAMORU-ni MORATTA rasii].
   she-Top that-Acc Mamoru-from received they.say
   ‘They say that she received it from Mamoru’

Order-I

B3: #[Inf Kanojo-wa [Inf MAMORU-ni are-o MORATTA rasii].

Notice here that the sentence fragment in (39B1), where the anaphoric arguments are deleted, is the most acceptable. When the anaphoric elements are deleted, the sentence does not involve focal stress and the whole sentence inherits the syntactic feature [Focus]$_{S}$ from the PP *Mamoru-ni ‘Mamoru-from’ which dominates the N *Mamoru in the default sentence stress position. The way Japanese encodes (double) narrow-focus prosodically is summarized in (40):
Prosodic encoding of (double) narrow-focus in Japanese

<table>
<thead>
<tr>
<th>Prosodic process</th>
<th>(i) Deletion</th>
<th>(ii) De-accenting</th>
<th>(iii) Normal prominence (e.g. Jap pronominals)</th>
<th>(iv) Extra prominence (e.g. Jap default sentence stress)</th>
<th>(v) Extra-extra prominence (e.g. Jap focal stress)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Single narrow-focus w/ deletion (e.g. (B1) in (29),(34))</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a'. Single narrow-focus in the default stress position (e.g. (29c-B2),(34c-B2))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a''. Single narrow-focus w/o deletion (e.g. (29c-B3),(34c-B3))</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Double narrow-focus w/ deletion (e.g. (39B1))</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b'. Double narrow-focus, one in the default stress position (e.g. (38B1))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b''. Double narrow-focus w/o deletion (e.g. (B2) in (38)-(39))</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Japanese, deleting anaphoric elements (= (i)) as in (40a) and (40b) is the most preferable option to encode disanaphoricity. When there is only one narrow-focus as in (40a), the sentence ends up consisting of focused elements only, which carry normal lexical stress (= (iii)). When there are two narrow-foci as in (40b), the intended focus interpretation is achieved by focus projection from the element in the default sentence stress position (= (iv)).

When deletion is not possible for pragmatic reasons of recoverability, Japanese prefers using default sentence stress (= (iv)) as in (40a') and (40b'), rather than invoking focal stress (= (v)) as in (40a'') and (40b''). When focal stress is employed at all, the left edge of a prosodic word with focal stress starts a new IntP (in accordance with the SO-PO mapping constraint for [Focus] in (33a)), and the following elements are deaccented (= (ii)) (in accordance with the Prosodic Condition on Japanese [Focus]PO in (33b)).

In double narrow-focus sentences, the change in the prosodic pattern forces a change in word order, so that two occurrences of focal stress can be adjacent to each other and form a single IntP. In other words, Japanese “dephrases” prosodic structure to encode disanaphoricity in such cases. The Prosodic Economy Condition in (30) captures the generalization that in Japanese, the dephrasing strategy is avoided and the word order alternation takes place instead, in order to achieve the default intonation pattern.

Remember that English does not normally use deletion for anaphoric arguments, and relies on different degrees of prosodic prominence, e.g. deaccenting anaphoric elements (= (ii)) and accenting focused elements (= (iv)), in encoding disanaphoricity. The mobility of prosodic prominence is correlated with less flexible word order in English: word order alternations do not have to be employed in order to encode the intended focus pattern, because the language can use mobile prosodic prominence for that purpose. In other words, Japanese and English resort to different types of economy: movement of dependents (i.e. word order alternations) is more economical than movement of prosodic prominence in Japanese, whereas movement of prosodic prominence is more economical than movement of dependents.
(i.e. word order alternations) in English. In chapter 5, I will attempt to reduce this difference to the difference in the general prosodic properties of the two languages, in particular, the distinctiveness of lexically specified pitch features.

Thus, we have reached the generalization we saw in (2):

(2) Linearization in the verbal domain in Japanese
Let the Order-I = PP – NP – V with the prosodic pattern \[\text{In}P \text{PP}\] \[\text{In}P \text{NP} - V\], and
the Order-II = NP – PP – V with the prosodic pattern \[\text{In}P \text{NP}\] \[\text{In}P \text{PP} - V\],

a. select Order-I or Order-II if there is no narrow-focus,

b. select Order-I if NP is (a part of) a narrow-focus, and

c. select Order-II if PP is (a part of) a narrow-focus.

The driving force for the word order alternation in Japanese is the Prosodic Economy Condition in (30): in Japanese short-scrambling, a narrow-focus is placed in the default sentence stress position so that focal stress need not be employed. Thus, linearization of verbal dependents in Japanese is also prosodically conditioned, despite that neither order is prosodically more marked than the other.

The generalization in (2) is primarily based on the results of a systematic elicitation I conducted in order to examine the prosodic properties associated with the alternation between Order-I and Order-II. (The details of the elicitation are presented in appendix to this chapter.) The table in (41) summarizes the acceptability of Order-I and Order-II with five different focus interpretations. The sample sentences are presented in (42)-(46). I constructed stimuli sentences which do not delete anaphoric elements but retain them as pronouns or lexical words. This was necessary in order to examine the prosodic properties of alternate word orders in Japanese, and compare them directly with English.

(41) Summary of acceptability

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad</td>
<td>(42)IP</td>
<td>√ [\text{In}P \text{PP} \text{NP}^f - V]</td>
<td>√# [\text{In}P \text{NP} \text{PP}^f - V]</td>
</tr>
<tr>
<td>Single narrow</td>
<td>(43)NP</td>
<td>√ [\text{In}P \text{PP} \text{NP}^f - V]</td>
<td># [\text{In}P \text{NP} \text{PP}^f - V]</td>
</tr>
<tr>
<td>Double narrow</td>
<td>(44)PP</td>
<td># [\text{In}P \text{PP} \text{NP}^f - NP – V]</td>
<td>√ [\text{In}P \text{NP} \text{PP}^f - V]</td>
</tr>
<tr>
<td>Double narrow</td>
<td>(45)NP+V</td>
<td>√ [\text{In}P \text{PP} \text{NP}^f - V]</td>
<td># [\text{In}P \text{NP} \text{PP}^f - V]</td>
</tr>
<tr>
<td></td>
<td>(46)PP+V</td>
<td># [\text{In}P \text{PP} \text{NP}^f - NP – V]</td>
<td>√ [\text{In}P \text{NP} \text{PP}^f - V]</td>
</tr>
</tbody>
</table>

(42) Broad-focus (IP-focus)
[‘What happened yesterday?’]

a. [\text{In}P \text{Mayumi-ga}] [\text{In}P \text{Mamoru-ni}] [\text{In}P \text{nuigurumi-o moratta yo}].

‘Mayumi received a doll from Mamoru’

b. \# [\text{In}P \text{Mayumi-ga}][\text{In}P \text{nuigurumi-o}] [\text{In}P \text{Mamoru-ni moratta yo}].

(43) Single narrow-focus (NP[theme]-focus)
[‘Mamoru and Mayumi are holding something. What are they doing?’]

a. [\text{In}P \text{Mayumi-ga}] [\text{In}P \text{Mamoru-ni}] [\text{In}P \text{nuigurumi-o} moratta]

Mayumi-Nom Mamoru-from doll.Acc received yo

mitai yo.

‘It seems that Mayumi received a doll from Mamoru’

b. \# [\text{In}P \text{Mayumi-ga}][\text{In}P \text{NUIGURUMI-o} Mamoru-ni moratta mitai yo].

(44) Single narrow-focus (PP[source]-focus)
[‘Mayumi is holding a doll carefully. Did she receive it from someone?’]
a. #\[\text{IntP Kanojo-wa} \quad \text{[IntP MAMORU-ni are-o moratta rassii yo].} \quad \text{she-Top Mamoru-from that-Acc received they.say yo} \]

‘They say that she received it from Mamoru’

b. \[\text{[IntP Kanojo-wa]} \quad \text{[IntP are-o]} \quad \text{[IntP MAMORU-ni moratta rassii yo].} \]

(45) Double narrow-focus (NP+V-focus)

[‘Mamoru and Mayumi look intimate. What happened to them?’]

a. \[\text{[IntP Mayumi-ga]} \quad \text{[IntP Mamoru-ni]} \quad \text{[IntP nuigurumi-o moratta Mayumi-Nom Mamoru-from doll-Acc received mitai yo].} \]

‘It seems that Mayumi received a doll from Mamoru’

b. \[\text{[IntP Mayumi-ga]} \quad \text{[IntP NUGURUMI-o Mamoru-ni MORATTA mitai yo].} \]

(46) Double narrow-focus (PP+V-focus)

[‘Mayumi was holding a doll carefully. Do you know why?’]

a. \[\text{[IntP Kanojo-wa]} \quad \text{[IntP MAMORU-ni are-o MORATTA rassii yo].} \quad \text{she-Top Mamoru-from that-Acc received they.say yo} \]

‘They say that she received it from Mamoru’

b. \[\text{[IntP Kanojo-wa]} \quad \text{[IntP are-o]} \quad \text{[IntP MAMORU-ni moratta rassii yo].} \]

The examples in (42)-(46) contain the V moratta ‘received’ which takes two internal arguments, a theme NP and a source PP. The results were: in broad-focus sentences, Order-I is perfectly acceptable, whereas Order-II shows variable acceptability but can be perfectly acceptable for some speakers, as is seen in (42); in narrow-focus sentences where the NP is (a part of) a narrow-focus, Order-I is preferred over Order-II, as is seen in (43) and (45); in narrow-focus sentences where the PP is (a part of) a narrow-focus, Order-II is preferred over Order-I, as is seen in (44) and (46). The results were clearer in the written test where the stimuli sentences were presented in a written form, rather than in the listening test where they were presented orally.

In the next section, I will illustrate how narrow-focus sentences with Order-I or Order-II are derived, and how the Prosodic Economy Condition in (30) explains the acceptability of narrow-focus sentences. After that, I will discuss the unexpected findings of the listening test of the elicitation.

4.1.3.2. Derivations

Unlike in English, two orders in Japanese are derived through the same derivational steps because they exhibit the same prosodic pattern. The preference for one order over the other depends on the locus of a narrow-focus. We already saw the derivation of broad-focus sentences in 4.1.2.3. In broad-focus sentences, both orders are available and exhibit the same intonation pattern. In particular, the left edge of a syntactic phrase is aligned with the left edge of an Intonational Phrase (IntP) by the SO-PO mapping constraint for maximal prosodic objects in (5b).

In single narrow-focus sentences, the order where a narrow-focus occupies the default sentence stress position is preferred over the other order. I formulated this generalization by the Prosodic Economy Condition specific to Japanese in (30) (repeated below):

(30) The Prosodic Economy Condition for Japanese

Japanese employs default sentence stress rather than focal stress whenever possible, in order to prosodically encode disanaphoricity.

First, look at the derivations of NP-focus sentences given in (47) below. The non-scrambled sentence in (43a) shows the default intonation pattern: it consists of three IntPs, each of which corresponds to a syntactic phrase, the subject NP, the source PP, and the VP. Since the feature [Focus]$_{SO}$ is on the NP nuigurumi-o ‘doll-Acc’ which dominates the N in the default sentence stress position, the prosodic
pattern does not differ from that of a broad-focus sentence.

On the other hand, in the scrambled sentence in (43b), the narrow-focus NP does not occupy the default sentence stress position. In this case, the NP *NUIGURUMI-o 'doll-Acc'* is marked with [Focus]$_{SO}$, and the left edge of the N *NUIGURUMI 'doll'* with [Focus]$_{PO}$ starts a new IntP, and the IntP becomes the rightmost IntP (= PO$_n$). This dephrasing effect is captured by the SO-PO mapping constraints for [Focus] in (33a) and the Prosodic Condition on Japanese [Focus]$_{PO}$ in (33b) (repeated below):

(33) a. SO-PO mapping for maximal prosodic objects (= Intonational Phrases) with [Focus] in Japanese, Align (SO, L; PO, L), where SO is with [Focus]$_{SO}$.

b. The Prosodic Condition on Japanese [Focus]$_{PO}$
   In Japanese, [Focus]$_{PO}$ must be in the rightmost Intonational Phrase (= PO$_n$).

Therefore, the scrambled sentence in (43b) consists of two IntPs. The feature [Focus]$_{PO}$ is realized as prosodic prominence on the prosodic word *NUIGURUMI 'doll'*; and the following elements are all deaccented.

(47) Derivation of the examples in (43): single NP[theme]-focus

a. Order-I (43a) : [IntP Mayumi-ga][IntP Mamoru-ni][IntP nuigurumi-o moratta]

   Numeration  {Mayumi, ga, Mamoru, ni, nuigurumi, o, [F]$_{SO}$, moratta}  

   syntactic structure

   \[
   \begin{array}{c}
   \text{[NP Mayumi ga]} \\
   \text{[Mg] [NP Mamoru ni]} \\
   \text{[Mg] [Mg] [VP[NP[F]SO, nuigurumi-o] moratta]} \\
   \text{PO$_{n-1}$} \\
   \text{[IntP Mayumi-ga]} \\
   \text{[IntP Mamoru-ni]} \\
   \text{[IntP nuigurumi[F]$_{PO}$-o moratta]} \\
   \text{PO$_n$} \\
   \end{array}
   \]

   prosodic structure

b. Order-II #(43b) : [IntP Mayumi-ga][IntP nuigurumi-o Mamoru-ni moratta].

   Numeration  {Mayumi, ga, Mamoru, ni, nuigurumi, o, [F]$_{SO}$, moratta}  

   syntactic structure

   \[
   \begin{array}{c}
   \text{[NP Mayumi ga]} \\
   \text{[Mg] [Mg][F]$_{SO}$, nuigurumi o]} \\
   \text{[Mg][F]$_{SO}$} [VP[Mamoru ni] moratta]} \\
   \text{PO$_{n-1}$} \\
   \text{[IntP Mayumi-ga]} \\
   \text{PO$_n$} \\
   \end{array}
   \]

   prosodic structure

   \(\text{(# SO[F]$_{SO}$ is not left-adjacent to the V.)}\)
The Prosodic Economy Condition for Japanese in (30) correctly predicts that Order-I is preferred over Order-II in this case: Order-I does not have to employ focal stress in order to encode the intended NP-focus interpretation, whereas Order-II does because the NP with [Focus] is not left-adjacent to the V and hence does not dominate the default sentence stress position. I assume that a violation of the Prosodic Economy Condition makes the sentence degraded, but not totally ill-formed.

Note that the Prosodic Weight Condition that formulates the right-edge heaviness effect in English is satisfied in both Order-I and Order-II, and hence irrelevant to word order alternations in Japanese. In the non-scrambled sentence in (43a), the PO₉ (nuigurumi-o, moratta) is prosodically heavier than the PO₉₋₁ (Mamoru-ni) by one prosodic word, and also contains a prosodic word with extra prosodic prominence (nuigurumi), because default sentence stress in Japanese is prosodically distinct. In the scrambled sentence in (43b), the PO₉ (NUIGURUMI-o, Mamoru-ni, moratta) is heavier than the PO₉₋₁ (Mayumi-ga) by two prosodic words, and again also contains a prosodic word with extra-extra prosodic prominence, i.e. focal stress in Japanese (NUIGURUMI).

The derivations of the PP-focus sentences in (44) go through the same steps as NP-focus sentences. The difference is that it is Order-II, not Order-I that is preferred in this case, as is predicted by the Prosodic Economy Condition.

Let us move on to double narrow-focus cases. In double narrow-focus sentences, the order where two prosodic words with extra prosodic prominence are adjacent to each other is chosen. For example, in the NP+V-focus sentences in (45), Order-I must be chosen, and in the PP+V-focus sentences in (46), Order-II must be chosen. If they are non-adjacent, the second occurrence of extra prosodic prominence leads to a violation of the Prosodic Condition on Japanese [FOCUS]PO in (33b).

Look at the NP+V-focus sentences in (45), for example. In NP+V-focus sentences, PP – NP – V order (Order-I) in (45a) is preferred over the alternative NP – PP – V order (Order-II) in (45b). The non-scrambled sentence in (45a) can encode the NP+V-focus interpretation by focus projection of the feature [Focus]SO on the NP (indicated by the dotted arrow), because the NP dominates the default sentence stress position. The derivation is illustrated in (48a) below.

On the other hand, in the scrambled sentence in (45b), the two narrow-foci, the NP and the V, are non-adjacent and hence must be independently marked for [Focus]SO. In this case, the Prosodic Condition on Japanese [FOCUS]PO in (33b) is violated, because the left edge of the first prosodic word with [Focus]PO, NUIGURUMI 'doll' is only aligned with the left edge of the second IntP, not with the rightmost IntP (= PO₉).

(48) Derivation of the examples in (45): double NP+V-focus

a. Order-I (45a): [ɪw Mayumi]-ga[[ɪw  Mamoru]-ni][ɪw  nuigurumi-o  moratta]

Numeration {Mayumi, ga, Mamoru, ni, nuigurumi, o, [F]SO, moratta}

syntactic structure

\[
\begin{array}{c}
\text{[NP Mayumi ga]} \\
\Rightarrow \\
\text{[Mayumi ga] [NP Mamoru ni]} \\
\Rightarrow \\
\text{[Mayumi ga, Mamoru ni] [nuigurumi-o moratta]} \\
\end{array}
\]

prosodic structure

\[
\begin{array}{c}
\text{[ɪw Mayumi]-ga} \\
\text{SO-PO mapping} \\
\text{spell-out} \\
\text{PO₉₋₁} \\
\end{array}
\]

\[
\begin{array}{c}
\text{[ɪw Mamoru]-ni} \\
\text{PO₉} \\
\end{array}
\]

\[
\begin{array}{c}
\text{[ɪw nuigurumi]-F[F]SO-o moratta} \\
\end{array}
\]
4.1.4. Residue: Unexpected Results of the Elicitation

The key data dealt with in this chapter are results of a systematic elicitation I conducted in order to examine the interface properties of the short-scrambling alternation in Japanese. I have already shown the stimuli sentences used in the study in (42)-(46), and illustrated in (47)-(48) how the Prosodic Phase Hypothesis accounts for their acceptability in terms of the Prosodic Economy Condition for Japanese in (30) and the relevant SO-PO mapping constraints in (33). In this section, I summarize the findings of the elicitation, and then discuss an unexpected result of the elicitation. (More details of the elicitation are found in appendix to this chapter.)

A systematic elicitation was designed as a test of how natural or unnatural sentences with two different word orders sound to native speakers of Japanese. The major findings of the judgment task are summarized in (49):

(49) Major findings of the judgment task

i. Findings predicted by the Prosodic Phase Hypothesis

a. In narrow-focus sentences (whether single or double), default sentence stress on a narrow-focus is preferred over extra higher prosodic prominence (e.g. focal stress) on it.

⇒ The Prosodic Economy Condition for Japanese in (30).

This listening test was followed by a written test which does not control for the prosodic pattern of stimuli sentences. The results of the written test are clearer than, and hence support, the findings of the listening test. See appendix for the details of the written test.
b. In narrow-focus sentences (whether single or double), the order where
(a part of) a narrow-focus occupies the default sentence stress position
is preferred over the other order.
⇒ The Prosodic Economy Condition for Japanese in (30).

c. In double narrow-focus sentences, the order where two occurrences of
focal stress are adjacent to each other is preferred over the order where
they are non-adjacent.
⇒ The Prosodic Condition on Japanese [Focus] in (33b).

ii. Finding unexpected by the Prosodic Phase Hypothesis
In double narrow-focus sentences, employment of prosodic prominence is
preferred over non-adjacent focal stresses, even in the cases where the
prosodic prominence is not on any narrow-focus.

First of all, a consistent preference was found for default sentence stress over focal stress (=
finding in (49i-a)). For example, in the NP-focus sentence with Order-I as in (50a), the NP with default
sentence stress is judged as natural as or more natural than the NP with focal stress. Likewise, in the PP-
focus sentence with Order-II as in (50b), the PP with default sentence stress is judged as natural as or more
natural than the PP with focal stress.

(50) Single narrow-focus: NP-focus in Order-I and PP-focus in Order-II

a. Order-I
\[
\begin{array}{c}
{[\text{NP}]} - {[\text{PP} - \text{V}]}
\end{array}
\]
default stress, not focal stress
e.g.
\[\text{[Mamoru-ni][migurumi-o moratte]}\]
(‘Mamoru-from doll-Acc received’)

b. Order-II
\[
\begin{array}{c}
{[\text{NP}] - [\text{PP} - \text{V}]}
\end{array}
\]
default stress, not focal stress
e.g.
\[\text{[are-o][Mamoru-ni moratte]}\]
(‘that-Acc Mamoru-from received’)

The double narrow-focus cases also support the same point. When one of the narrow-foci sits in the
default sentence stress position, it carries only default sentence stress, not focal stress with an extra higher
prominence.

(51) Adjacent double narrow-focus: NP+V-focus in Order-I and PP+V-focus in Order-II

a. Order-I
\[
\begin{array}{c}
{[\text{NP}]} - {[\text{PP} - \text{V}]}
\end{array}
\]
default stress, not focal stress
e.g.
\[\text{[Mamoru-ni][migurumi-o moratte]}\]

b. Order-II
\[
\begin{array}{c}
{[\text{NP}] - [\text{PP} - \text{V}]}
\end{array}
\]
default stress, not focal stress
e.g.
\[\text{[are-o][Mamoru-ni moratte]}\]

The preference for default sentence stress over focal stress in Japanese was formulated as the Prosodic
Economy Condition for Japanese in (30). Recall that the preference for default sentence stress over focal
stress in Japanese contrasts with the opposite preference in English: in English, default sentence stress is
not necessarily distinct prosodically, and focal stress needs to be employed in encoding a narrow focus
(see (31)).

The preference for default sentence stress over focal stress is the driving force for Japanese short-
scrambling: in single narrow-focus sentences, the order where (a part of) a narrow-focus occupies the
default sentence stress position is preferred over the other order (= finding in (49i-b)). For example, PP –
NP – V order (Order-I) is preferred over Order-II when the NP is (a part of) a narrow-focus, whereas NP –
PP-V order (Order-II) is preferred over Order-I when the PP is (a part of) a narrow-focus. The order that forces the employment of focal stress contrary to the Prosodic Economy Condition makes the sentence degraded.

Moving on to non-adjacent focus cases, native speakers of Japanese consistently preferred the prosodic pattern with adjacent occurrences of focal stress over the pattern with non-adjacent occurrences of focal stress (= finding in (49i-c)). This effect was formulated by the Prosodic Condition on Japanese [Focus] in (33b), which requires the elements after focal stress to be incorporated into one Intonational Phrase (IntP). A violation of the Prosodic Condition on Japanese [Focus] makes the sentence degraded.

An unexpected finding was that when two foci are disjoint from each other, the prosodic pattern in (52A1) below, where there is only default sentence stress, is judged as natural as or more natural than the prosodic pattern in (52A2), where there are non-adjacent occurrences of focal stress (= finding in (49ii)). This suggests that the primary source of unacceptability of narrow-focus sentences should be non-adjacent occurrences of extra prosodic prominence (e.g. focal stress), rather than non-adjacent occurrences of semantic/pragmatic focus.

(52) Non-adjacent double narrow-focus: PP+V-focus in Order-I and NP+V-focus in Order-II

-a. Order-I

(A1) \(\text{[Imp PP]} - \text{[Imp NP} - \text{V]}\)

\(\text{default stress}\)

e.g. \([Mamoru-}\text{ni} \text{are-o moratta}\]

('Mamoru-from that-Acc received')

(A2) \# \(\text{[Imp PP} - \text{NP} - \text{V]}\)

\(\text{focal stress focal stress}\)

e.g. \([\text{Nui} \text{gurumi-o} \text{Mamoru-}\text{ni}\text{ moratta}\]

('doll-Acc Mamoru-from received')

-b. Order-II

(A1) \(\text{[Imp NP]} - \text{[Imp PP} - \text{V]}\)

\(\text{default stress}\)

e.g. \([\text{Nui} \text{gurumi-o} \text{Mamoru-}\text{ni}\text{ moratta}\]

('doll-Acc Mamoru-from received')

(A2) \# \(\text{[Imp NP} - \text{PP} - \text{V]}\)

\(\text{focal stress focal stress}\)

e.g. \([\text{Nui} \text{gurumi-o M-}\text{ni}\text{MORATTA}\]

Although the Prosodic Condition on Japanese [focus] in (33b) correctly rules out the prosodic pattern in (A2), it remains to be explained why the prosodic pattern in (A1) is possible at all: in (A1), default sentence stress falls on the phrase intervening between the two foci, and it is not explained how the non-adjacent narrow-focus interpretation is achieved.

It is worth noting here that the (A1) examples are not perfectly natural by themselves, in that they retain anaphoric arguments. Remember that Japanese normally deletes anaphoric elements when their informational content is recoverable from discourse context. For example, with the PP+V-focus interpretation as in (53) below, the most natural answer consists only of PP and V. In (53B), the prosodic word Mamoru carries default sentence stress but not focal stress.

(53) PP+V-focus

A: ‘Mayumi was holding a doll carefully. Do you know why?’
B: \(\text{[Imp Mamoru-}\text{ni}\text{moratta mitai yo]}\)

‘Mamoru-from received it-seems yo’

‘It seems that (she) received (it) from Mamoru’

When anaphoric elements are deleted, the issue of word order disappears, because Order-I and Order-II are not distinguishable.

Going back to the remaining problem in (49ii-a), I argue that Japanese prosody is subject to a general restriction on the employment of extra higher prosodic prominence (e.g. focal stress): since Japanese normally uses deletion or a word order alternation in order to encode disanaphoricity, it
accordingly avoids assignment of extra prosodic prominence, which requires the boosting of the peak of lexical High tone. Although this predicts the preference for the prosodic pattern (52A1) over (52A2), how the intended focus-interpretation is achieved with the pattern (A1) still remains to be explained. Given the limited scope of the elicitation, however, I would like to defer further speculation on this issue until I have looked into the prosodic properties of a larger set of examples.

4.2. More on the Absence of the Distinction between Canonical and Marked Orders

Section 4.1 showed that no prosodic markedness is associated with the short-scrambling alternation in Japanese: in broad-focus sentences, the non-scrambled PP – NP – V order (Order-I) and the scrambled NP – PP – V order (Order-II) are available; in narrow-focus sentences, the order where (a part of) a narrow-focus occupies the default sentence stress position is preferred over the other, conforming to the Prosodic Economy Condition for Japanese. This section considers why Japanese short-scrambling is motivated by the Prosodic Economy Condition, and is not associated with prosodic markedness. In section 4.2.1, I propose that the Prosodic Economy Condition for Japanese reflects production efficiency in the performance systems. In section 4.2.2, I argue that morphologically overt Case in Japanese provides an independent argument for the lack of markedness in short-scrambling.

4.2.1. The Prosodic Economy Condition and Production Efficiency

The Prosodic Economy Condition for Japanese formalizes the prosodic properties that trigger short-scrambling in Japanese:

(54) **The Prosodic Economy Condition for Japanese [≡ (30)]**

Japanese employs default sentence stress rather than focal stress whenever possible, in order to prosodically encode disanaphoricity.

The term “economy” in the Prosodic Economy Condition specifically means economy consideration from a production perspective: prosodic prominence with a lower pitch peak demands less effort in production than prosodic prominence with a higher pitch peak (cf. 4.1.3.1). The Prosodic Economy Condition cannot be a universal economy condition, because the opposite effect is found in narrow-focus sentences in English: English employs focal stress to differentiate a narrow-focus from a partial broad-focus.

(55) **Prosodic encoding of disanaphoricity in English when narrow-focus = partial broad-focus [≡ (31)]**

When a narrow-focus is in the default sentence stress position, English employs focal stress rather than default sentence stress, in order to prosodically encode disanaphoricity.

Therefore, the Prosodic Economy Condition in (54) is a Japanese-specific condition, which triggers word order alternations in Japanese. On the other hand, the statement in (55) is a descriptive generalization regarding the production of narrow-focus sentences in English, and does not constrain linearization of verbal dependents in English.

A remaining question is why there is such a difference between Japanese and English. I speculate that the difference is due to the prosodic property of default sentence stress: default sentence stress in Japanese is prosodically distinct in that its pitch peak is higher than that of normal lexical stress, whereas default sentence stress in English is normal lexical stress and can be perceived as prominent only relative to neighboring deaccented elements. Therefore, focal stress in Japanese requires extra boosting of a pitch peak. This is illustrated in the Disanaphoricity Scale:
### 4.2.2. Absence of the Prosodic Markedness Effect and Case Morphology

While Japanese short-scrambling is prosodically constrained by the Prosodic Economy Condition, neither PP – NP – V order (Order-I) nor NP – PP – V order (Order-II) is prosodically more marked than the other. This contrasts with English heavy NP shift, where there is a prosodic markedness effect and the marked order must satisfy the Proodic Weight Condition. Regarding this difference between Japanese and English, I made an assumption in 4.1.2.2 that Japanese NPs receive Case from morphologically overt case-markers within NPs, and therefore do not have to be adjacent to the case-assigning V, whereas English does not have overt case-markers and DPs need to be adjacent to the case-assigning V, as was formulated by the Case Adjacency condition (see (23)). An additional assumption was that the Case Adjacency requirement is violable and overridden by the Prosodic Weight Condition in the shifted sentences in English.

Other things being equal, Order-I and Order-II in Japanese should be equally available in principle. However, there is an implicational relation that in broad-focus sentences, Order-II is never more acceptable than Order-I (cf. the table (18) in 4.1.2.2), and Order-II is lower in frequency than Order-I (cf. the table in (19) in 4.1.2.2). In the literature, Order-I is sometimes considered as canonical and Order-II is analyzed as involving a short-scrambling operation, which derives the scrambled NP – PP – V order (Order-II) from its canonical PP – NP – V counterpart (Order-I) via movement of the NP (e.g. Saito 1985, Hoji 1985). Based on acquisition data, Sugisaki and Isobe (2001) also argue for a movement approach to scrambled sentences: Japanese kids have more difficulty in comprehension of scrambled sentences than non-scrambled sentences, which suggests that Order-II is more marked than Order-I.

Another potential argument for the marked status of Order-II is found in the usage of

---

<table>
<thead>
<tr>
<th>(i) deletion</th>
<th>(ii) deaccensing</th>
<th>(iii) normal prominence (i.e. lexical stress)</th>
<th>(iv) extra prominence</th>
<th><em>(v)</em> extra-extra prominence</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. In English,</td>
<td>e.g. In Japanese,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) ---</td>
<td>(i) null NP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ii) pronominal</td>
<td>(ii) ---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iii) default sentence stress</td>
<td>(iii) pronominal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iv) focal stress</td>
<td>(iv) default sentence stress</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(v) ---</td>
<td>(v) focal stress</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notice that the prosodic processes employed in Japanese are one scale higher than those in English (as indicated by bi-directional arrows). For example, English pronominals are lexically stressless (= (iii)) whereas Japanese pronominals carry lexical stress (= (iii)). Japanese differentiates default sentence stress from lexical stress by assigning extra prominence to it (= (iv)), whereas English marks default sentence stress only with normal lexical stress (= (iii)) by contrasting it with deaccented pronominals. (See chapter 5 for further comparison from a broader typological perspective.)

Since the default sentence stress in Japanese already carries “extra” prosodic prominence, Japanese needs to differentiate focal stress from default sentence stress by assigning an “extra higher” prominence to it (= (v)). I assume that there is a universal restriction, motivated for production efficiency, against the employment of extra-extra prominence in (v), from which the Prosodic Economy Condition in Japanese derives ultimately. That is to say, the process of boosting an already increased prominence, in fact, is not available in natural languages (as indicated by “*” in (56v)). The Prosodic Economy Condition is irrelevant to English, because English does not use extra-extra prosodic prominence from the beginning.
morphologically overt case-markers. The accusative case-marker -o can drop in colloquial Japanese when the accusative NP immediately precedes the V (e.g. Order-I) as is shown in (57), whereas it does not when the NP is not adjacent to the V (e.g. Order-II) as is shown in (58) (Kuroda 1992:322-323):

(57) Order-I
Kay-Nom library-to Mishima-by novel-Acc donated yo
‘Kay donated novels by Mishima to the library’
b. Kei-ga [pp tosyokan-ni] [np Mishima-no syoosetu] kizousita yo.

(58) Order-II
Kay-Nom Mishima-by novel library-to donated yo
‘Kay donated novels by Mishima to the library’
b. ?? [np Mishima-no syoosetu*()]Kei-ga [pp tosyokan-ni] kizousita yo.14

In (58a), the accusative NP Mishima-no syoosetu-o ‘novels by Mishima’ appears before the goal PP, and in (58b), the accusative NP appears before the subject NP. In these cases, the accusative case-marker cannot drop. The contrast between (57) and (58) may suggest that Order-II is more marked than Order-I from a perspective of Case morphology, in that only Order-I allows the accusative case-marker to drop. Notice, however, that this is independent of the issue of whether Order-II is based-generated or derived from Order-I via movement. In fact, based on experimental data, Yamashita (1997) argues that the contrast between the two orders has to do with the type of morphological Case itself.

The point I would like to emphasize is that neither order is more marked than the other with respect to prosody. This contrasts with the heavy NP shift alternation in English, where the shifted order is more marked than the non-shifted order with respect to prosody, in addition to the Case Adjacency requirement.

4.3. Deriving the Properties of Order-II

Previous studies of Japanese short-scrambling have explained various properties associated with the alternation, such as relatively free ordering of dependents, locality, the syntactic position of a scrambled phrase, binding effects, scope effects, and so on (cf. Ross 1967, Saito 1985, Hoji 1985, Miyagawa 1997, Ishihara 2000 among others). While some of these properties are syntactic in nature, not every property can be explained solely in syntactic terms. This section shows that some properties of scrambled sentences with Order-II are well captured by the Prosodic Economy Condition for Japanese (4.3.1), and that syntactic properties can be explained in terms of the incremental structure-building mechanism that yields right-branching structure in the syntax (4.3.2).

14 The sentence in (58b) sounds slightly better than (58a) with an intonational break after the preposed NP. This suggests that the preposed NP acts as a sentential topic in this case. Other case-markers or postpositions cannot be dropped no matter where they appear.
4.3.1. Arguments for the Prosodic Economy Condition for Japanese

The Prosodic Economy Condition for Japanese predicts that no prosodic markedness is associated with Order-I or Order-II. This contrasts with the heavy NP shift alternation in English, where the shifted V-PP-DP order must satisfy the Prosodic Weight Condition. In 4.3.1.1, I show that a prosodic weight effect is absent in short XP-scrambling in general in Japanese. In 4.3.1.2, I argue that the locality effect between a head verb and its dependent in part follows from the Prosodic Economy Condition, in that a dependent can appear anywhere within the verbal domain as long as its positional pattern conforms to the Prosodic Economy Condition.

4.3.1.1. More on the Absence of a Prosodic Weight Effect in XP-Scrambling

So far, our focus has been on the short-scrambling alternation between the non-scrambled PP-NP-V order (Order-I) and the scrambled NP-PP-V order (Order-II). In contrast with English heavy NP shift, neither order is prosodically more marked than the other, and the two orders exhibit the same intonation pattern. In particular, the left edge of a syntactic phrase (XP) corresponds to the left edge of an Intonational Phrase (IntP), as is required by the SO-PO mapping constraint for maximal prosodic objects in Japanese in (5b), repeated below:

\[(59)\]  
Japanese: Align (SO, L; PO, L), where SO = phrase (XP)

Although Order-II forms mirror images with the marked V-PP-DP order (M-order) in English, scrambled sentences fail to exhibit a prosodic weight effect that characterizes shifted sentences in English. This section shows the absence of a prosodic weight effect holds for not only short NP-scrambling but also short XP-scrambling in general.

Accusative -o NPs and source -ni PPs are not the only dependents that participate in word order alternations in the verbal domain in Japanese. Dative -ni NPs, source -kara PPs, CPs, and AdvPs also participate in alternations, as is shown in (60)-(61):

\[(60)\]  
A: ‘What happened?’
B1: Jennifer-ga [NP Dianne-ni] [CP suguni uti-ni kaeru to] Jennifer-Nom Dianne-Dat soon home-to return Comp yakusokusita mitai. promised it.seems ‘It seems that Jennifer promised Dianne to go back home soon’
B2: Jennifer-ga [CP suguni utini kaeru to] [NP Dianne-ni] yakusokusita mitai.

\[(61)\]  
A: ‘What happened?’

In (60), the V ‘yakusokusita ‘promised’ takes two arguments, the NP ‘Dianne-Nom ‘Dianne-Dat’ and the CP ‘suguni uti-ni kaeru to ‘that (she will) go back home soon’, and both orders in (B1) and (B2) are possible. In (61), the VP contains the AdvP ‘kinou ‘yesterday’ and the PP ‘bikutoria-kara ‘from Victoria’, and again, both orders in (B1) and (B2) are possible. In broad-focus sentences where no focal stress is involved, all sentences show the same intonation pattern regardless of the syntactic category of dependent phrases: the left edge of a syntactic phrase corresponds to the left edge of an IntP.

A prosodic weight effect is absent in short XP-scrambling in Japanese. Remember that in English, non-DP dependents (e.g. CPs/IPs, PP, AdvPs), which do not require Case, participate in the XP-shift
alternation more freely than DPs, because the Case Adjacency condition is not at work in these cases. Moreover, CPs/IPs must appear at the right edge of the sentence, following another dependent (if any), because they are always IntPs on their own in English, due to the SO-PO mapping constraint for IntPs in English in (5a) (cf. 3.3.1.1). In contrast, the syntactic category of dependents does not affect the availability of a particular order in Japanese. This is because the Case Adjacency condition is not in effect in Japanese, and the SO-PO mapping constraint in (59) applies to XPs in general. For example, NPs and CPs equally freely participate in the short-scrambling alternation:

(62) NP-scrambling
   ‘Jennifer promised Dianne to go back home soon’

(63) CP-scrambling
   ‘Jennifer promised Dianne to go back home soon’

In (62), the proposition ‘that Jennifer would return home soon’ is expressed as an accusative NP with the aid of the noun koto ‘fact/thing’; in (63), the same proposition is expressed as a CP introduced by the complementizer to. Whether the NP follows (e.g. (62a)) or precedes a dative phrase (e.g. (62b)), the sentences are as acceptable as the corresponding CP-scrambling examples in (63). Furthermore, the absence of preference for Order-I over Order-II in NP-scrambling in (62) supports the argument that there is no adjacency requirement for Case-assignment on NPs in Japanese (cf. 4.2.2).

In English, a clausal (CP/IP) dependent always appears in VP-final position, due to their prosodic status as IntPs. In contrast, the examples in (63) show that CPs may appear either before or after another dependent in Japanese, which suggests that there is no weight effect in Japanese short-scrambling.

The number of prosodic words contained in the CP does not affect linearization either, at least as far as acceptability judgments by native speakers are concerned (cf. fn. 5).

(64) CP-scrambling
   owatta-ra suguni takusii-de uti-ni kaeru to] yakuokusuta.
   end-when soon taxi-by home-to return Comp promised
   ‘Jennifer promised Dianne to go back home by taxi soon when the party after
   conference’
   b. Jennifer-ga [CP gakkai-go-no paati-ga owatta-ra suguni takusii-de uti-ni
   kaeru to] Dianne-ni yakuokusuta.

In (64), the CP contains more prosodic words than the CP in (63), but both orders in (64a) and (64b) are still equally acceptable (according to native speakers I consulted). This is predicted because in Japanese, it is the Prosodic Economy Condition, not the Prosodic Weight Condition, that determines linearization.15

4.3.1.2. Locality of XP-Scrambling

15 In 3.3.1.1, we saw that English also has word order alternations in the nominal domain called extraposition from NP. In Japanese, modifiers to a head N are all strictly pre-nominal and they are not displaced or separated from the head N (but see Yatabe 1993 for potential extraposition from NP examples in Japanese).
So far, our observations have been restricted to “short”-scrambling within the VP domain. Japanese also exhibits medial-scrambling, where a dependent XP appears in front of a subject, and long-distance-scrambling, where a dependent XP is fronted across a clause boundary. (In the following examples, the relevant XP is in brackets, and the position where the XP would appear in Order-I is indicated by underscore.)

(65) Short-scrambling
Randy-Nom Jennifer-Acc Dianne-Dat introduced
‘Randy introduced Jennifer to Dianne’

(66) Medial scrambling
‘Randy introduced Jennifer to Dianne’

(67) Long-distance scrambling
Jennifer-Acc Rachel-Nom Randy-Nom Dianne-Dat introduced Comp said
‘Rachel said that Randy introduced Jennifer to Dianne’

Following Miyagawa (1997), I assume that short-scrambling on the one hand and medial and long-distance scrambling on the other are syntactically different phenomena: the former is base-generated, whereas the latter involves leftward syntactic movement of XP. Long-distance scrambling can cross a clause boundary, because it is presumed to apply successive cyclically as A’-movement. The syntactic status of medial scrambling is not as clear as long-distance scrambling and whether medial scrambling involves movement or not is controversial. For example, Miyagawa (2001) analyzes medial scrambling as A-movement driven by EPP feature on T (see also Saito 1992 for the analysis of medial scrambling as A-movement.). On the other hand, a base-generation analysis of medial scrambling is argued for by Sugisaki and Isobe (2001) based on acquisition data, and by Nakayama (1995) based on processing data. A closer investigation of medial scrambling is outside the scope of this thesis.

In contrast to medial and long-distance scrambling, short-scrambling is local and restricted to the VP domain by the definition of “short”-scrambling. As long as it takes place within the VP domain, the Prosodic Economy Condition for Japanese (repeated below as (68)) determines the position of dependent phrases.

(68) The Prosodic Economy Condition for Japanese [= (30)]
Japanese employs default sentence stress rather than focal stress whenever possible, in order to prosodically encode disanaphoricity.

For example, when the VP consists of a head V and two dependents XP and YP, a dependent phrase XP shows a particular positional pattern in broad-focus sentences. When the intended focus interpretation is VP1 (= YP+XP+V), VP2 (= XP+V), or XP, the XP dependent appears immediately before the head V as in (69a) below. In order to achieve XP-focus or VP2-focus interpretations, the YP – XP – V order in (69a) is not only possible but preferred by the Prosodic Economy Condition, since the XP is in the default sentence stress position and these focus interpretations are achieved by focus projection from [Focus], without invoking focal stress. On the other hand, XP can appear in the VP-initial position as in (69b), when the intended focus is VP1 (= XP+YP+V), VP3 (= YP+V), or YP. The XP – YP – V order in (69b) is not only possible but preferred by the Prosodic Economy Condition in YP-focus or VP3-focus sentences, since YP-focus or VP3-focus interpretations are not achieved in any other order without invoking focal stress.

(69) Position of [XP] in the verbal domain that consists of V, [XP], [YP] in broad-focus sentences
a. When the intended focus interpretation is VP1, VP2, or XP,

\[
\text{default sentence stress}
\]

b. When the intended focus interpretation is VP1, VP3, or YP

\[
\text{default sentence stress}
\]

Therefore, the generalization about the locality effect of short XP-scrambling is that XP can be non-adjacent to the head V, to the extent that it does not violate the Prosodic Economy Condition for Japanese. Moreover, XP must be non-adjacent to a head V, if XP is not a part of focus and another dependent YP is a part of focus (e.g. YP-focus and VP3-focus in (69b)). This differs from XP-shift in English, where XP may appear in sentence-final position only if it satisfies the Prosodic Weight Condition. Therefore, in Japanese short-scrambling, the locality condition between a head V and its dependent XP is an instance of the placement of a focused XP into the default sentence stress position, which is left-adjacent to the V in Japanese. In other words, it is a positional pattern of dependents that matters in Japanese short-scrambling, rather than the locality requirement between the V and its dependent(s).

4.3.2. Arguments for Incremental Structure-Building in the Syntax

Assuming a strictly derivational model of the computational component, incremental structure-building in the syntax allows the most natural implementation of the Prosodic Phase Hypothesis: syntactic objects (SOs) are built incrementally from left to right, mapped to prosodic objects (POs), whereby SO-PO pairs are derivationally established and spelled out into the phonological component. In 4.3.2.1, I show that the results of syntactic constituency tests applied to scrambled sentences support incremental structure-building in the syntax. In 4.3.2.2, I show that right-branching structure resulting from incremental structure-building is consistent with structure-sensitive LF effects (e.g. binding) exhibited by scrambled sentences.

4.3.2.1. Syntactic Constituency Tests

Phillips (1996, 2003) convincingly shows that the conflicting results of different constituency tests are explained by assuming incremental structure-building: different constituents are visible to different syntactic processes (e.g. coordination, ellipsis, movement), according to which point of the derivation each process applies at. In 3.3.2.1, I showed that shifted sentences in English bear out the prediction in (70):
Constituency tests such as (i) coordination and (ii) ellipsis, which involve the coordination of the same type of strings $\alpha_1$ and $\alpha_2$, may refer to the string $\alpha_1$ that is a syntactic constituent at the point in the incremental derivation when the conjunction ‘and’ is merged. \[\text{[\textit{86}}\text{ in } 3.3.2.1\]

In this section, I show that scrambled sentences in Japanese also bear out this prediction.

4.3.2.1.1. Coordination Test

First, let us look at the coordination test. As is the case with shifted sentences in English, more than one $V$ can be semantically associated with the theme NP argument in the scrambled NP – AdvP – $V$ order in Japanese. For example, the string of AdvP – $V$ can be coordinated as is shown in (71).

\[(71)\]

\begin{enumerate}
\item \[\text{NP – [AdvP} - V_j] \text{‘and’ [AdvP}_j - V_j]\]
\begin{enumerate}
\item \[\text{Watashi-wa [NP kisei-kuusyo-nituite-no ronbun-o [AdvP}_j juubun]}\]
I-Top parasitic-gap-about-of paper-Acc thoroughly
\[\text{[v}_1 \text{ giron} \text{sureru} \text{ maen, [AdvP}_2 \text{ gakkai-de] [v}_2 \text{ happyosita].}\]
\text{discuss before conference-at presented}
\text{‘I presented _ at a conference before discussing _ thoroughly the paper about parasitic gaps’}
\item \[\text{Watashi-wa [NP kisei-kuusyo-nituite-no ronbun-o [AdvP}_j kinou]}\]
I-Top parasitic-gap-about-of paper-Acc yesterday
\[\text{[v}_1 \text{ happyositi-te, [AdvP}_2 \text{ juubun] [v}_2 \text{ giron-si-takatta].}\]
\text{present-and thoroughly discuss-wanted}
\text{‘I wanted to present _ yesterday and discuss _ thoroughly the paper about parasitic gaps’}
\end{enumerate}
\end{enumerate}

In (71a), the NP \textit{kisei-kuusyo-nituite-no ronbun-o ‘paper about parasitic gaps’} is the theme argument of the Vs, \textit{giron} \textit{sureru ‘discuss’ and happyosa} \textit{‘presented’. Likewise, in (71b), the NP is the theme argument of the Vs, happyosa ‘present’ and giron-si-takatta ‘wanted to discuss’. The conjunction ‘before’ in (71a) or ‘and’ in (71b) coordinates the adjacent AdvP – $V$ sequences, which is predicted to be possible by (70). The derivation of a coordinated scrambled sentence is schematized in (72):

\[(72)\]

\begin{enumerate}
\item step 1 \[\Rightarrow\] step2 \[\Rightarrow\] step3 \[\Rightarrow\] step4
\begin{enumerate}
\item \[\text{NP – [AdvP}_j - V_j]\]
\item \[\text{I-Top parasitic-gap-about-of paper-Acc thoroughly}\]
\item \[\text{[v}_1 \text{ giron} \text{sureru} \text{ maen, [AdvP}_2 \text{ gakkai-de] [v}_2 \text{ happyosita].}\]
\item \[\text{discuss before conference-at presented}\]
\item \[\text{‘I presented _ at a conference before discussing _ thoroughly the paper about parasitic gaps’}\]
\end{enumerate}
\end{enumerate}

The conjunct $\alpha_1 (= \text{AdvP}_j - V_j)$ is a constituent at the step 3 when the conjunction ‘and’ is merged. The conjunct $\alpha_1$ and conjunct $\alpha_2 (= \text{AdvP}_j - V_j)$ are strings of the same type.

Likewise, two NP – AdvP strings can be adjacent to each other only with a conjunction element in between, and hence can be coordinated:
In (73), the theme NP, 'the paper about parasitic gaps' and the AdvP, 'yesterday' are the dependents of the V 'wanted to present', and so are the theme NP, 'the paper about ATB extraction' and the AdvP, 'tomorrow'. The first NP – AdvP string forms a constituent (= α1) at the point when the conjunction 'and' is merged, and the following NP – AdvP string also forms a constituent (= α2) when it is merged (and before the V is merged).

(74) Derivation of the coordinated scrambled sentence in (73)

Therefore the coordinated sentence in (73) is correctly predicted to be possible by (70).

4.3.2.1.2. Ellipsis Test

Next, let us look at the ellipsis test. As the example in (75) shows, the elided VP (or VP anaphora) *soo-su* 'do so' in the second conjunct can take the NP – AdvP – V string in the first conjunct as an antecedent. In contrast, the ellipsis targeting the partial string of NP – AdvP as an antecedent is impossible, as is shown in (76).

(75) \([\alpha_1 \text{ NP} – \text{AdvP} – \text{V}] \text{‘and’ } [\alpha_2 \phi]\)

Ken-wa [\alpha_1 [\text{NP} kisei-kuusyo-nituite-no ronbun-o] [\text{AdvP} gengo-gakkai-de] [\text{V} happyosi-te, Megu-mo [\alpha_2 soo-si]-ta. ‘Ken presented a paper about parasitic gaps at the conference of linguistics, and Meg did so too’

(76) \([\alpha_1 \text{ NP} – \text{AdvP} – \text{V}] \text{‘and’ } [\alpha_2 \phi] – \text{V}

*Megu-wa [\alpha_1 [\text{NP} sushi nijukko-o] [\text{AdvP go-fun-de}][\text{V} tukuri], Megu-top sushi twenty-roll-Acc Ken-Top tabeta koto-ga aru. Ken-top ate fact-Nom have
Meg has cooked/has an experience of cooking twenty sushi rolls in five minutes, and Ken has eaten/has an experience of eating (twenty sushi rolls in five minutes)'

The contrast between (75) and (76) is correctly predicted by (70): in (75), the antecedent of the VP anaphora, i.e. the NP - AdvP - V string, forms a constituent at the point when the conjunction -te 'and' is merged, whereas in (76), the antecedent of the elided element, i.e. the NP - AdvP string, does not form a constituent at the point when the conjunction is merged because of the existence of the intervening V.

(77) The point when the conjunction 'and' is merged

\[ \alpha 1 \text{ in (75)} \]

non-

AdvP

AdvP - V

constituent

In this section, we saw that syntactic constituencies in scrambled sentences involving coordination and ellipsis are well captured by incremental structure-building in the syntax. In the next section, I will show that the LF effects exhibited by scrambled sentences are consistent with their right-branching structure resulting from incremental structure-building.

4.3.2.2. Base-Generation and LF Effects

4.3.2.2.1. Interpretation of the Scrambled XP

The Prosodic Phase Hypothesis argues that scrambled sentences as well as non-scrambled sentences are base-generated incrementally via Merge and do not involve syntactic movement. The legitimate derivations of non-scrambled and scrambled sentences in the verbal domain are schematized as follows:

(78) Derivation of non-scrambled and scrambled sentences in the verbal domain

a. Non-scrambled sentence: Order-I
   e.g. \[ \text{[Inf PP]} [\text{[Inf NP - V]}] \]
   Numeration
   Incremental structure-building

b. Scrambled sentence: Order-II
   e.g. \[ \text{[Inf NP]} [\text{[Inf PP - V]}] \]
   Numeration
   Incremental structure-building

Derivations of the non-scrambled sentence in (78a) and the scrambled sentence in (78b) start with the same Numeration, and the choice between them is optional in the syntax. Both non-scrambled and scrambled sentences are derived in the syntax freely, and illegitimate orderings are ruled out by correspondence conditions on the syntax-prosody mapping.

The observation that the scrambled NP - PP - V order (Order-II) exhibits the same prosodic
pattern as the non-scrambled PP – NP – V order (Order-I) provides evidence for a base-generation analysis of Order-II: if Order-II were derived by syntactic leftward movement of the NP from the corresponding Order-I, there would be a VP boundary between PP and V, \([_{VP}NP - \{_{VP}PP - \{_{VP}t_{NP} - V\}\}]\), and it is wrongly predicted that a new domain of downstep starts after the PP and before the verb.

\[(79)\] Base-generation vs. movement approaches

a. Base-generation of non-scrambled and scrambled sentences

\[\text{Numeration}\]
\[\text{non-scrambled sentence} \quad \text{scrambled sentence}\]
\[(\text{e.g. PP} – \text{NP} – \text{V}) \quad \text{(e.g. NP} – \text{PP} – \text{V})\]

b. Movement approach to scrambled sentences

\[\text{Numeration}\]
\[\text{non-scrambled sentence} \quad \text{scrambled sentence}\]
\[(\text{e.g. PP} – \text{NP} – \text{V}) \quad \text{(e.g. NP} – \text{PP} – \text{t}_{\text{NP}} – \text{V})\]

As is the case with English heavy NP shift, I assume, by stipulation, that verbal dependents can be semantically associated with the non-adjacent head V as long as they are in the same VP domain.\(^{16}\)

A base-generation approach to Japanese short-scrambling is compatible with the traditional view of Japanese as a non-configurational language, in the sense that both Order-I and Order-II are base-generated (Farmer 1980, Hale 1980, Miyagawa 2001). The present analysis, however, regards Japanese as a configurational language with right-branching structure: syntactic structures are built incrementally from left to right, and the resulting structure has a right-branching configuration. In the rest of this section, I will show that LF effects associated with the short-scrambling alternation are consistent with right-branching structure assigned to scrambled sentences.

4.3.2.2.2. LF Effects

Given that both Order-I and Order-II are base-generated by incremental structure-building and the resulting sentences all have right-branching structure, it is predicted that a preceding phrase c-commands a following phrase, and not vice-versa, regardless of whether the sentence is Order-I or Order-II. In this section, I will show that (i) reciprocal binding and (ii) Weak Crossover phenomena provide supporting evidence for the incremental approach to structure-building in the syntax, and (iii) scope reconstruction and (iv) bound variable pronominals are consistent with the right-branching structure resulting from incremental structure-building.

First, let us look at reciprocal binding in non-scrambled and scrambled sentences. A base-generation approach to the short-scrambling alternation predicts that the reciprocal \textit{otagai} 'each other' can be bound by a preceding NP, regardless of whether the sentence is Order-I or Order-II. This prediction is borne out.

\[16\] Yoshida (2000) also proposes a base-generation approach to short-scrambling, and relaxes locality of theta-relation. See 4.4.2 for the review of base-generation approaches to Japanese short-scrambling.
(80) Order-I (PP – NP – V)
   John-Nom Hanako-and Mary-Acc each.other-to introduced
   ‘John introduced each other to Hanako and Mary’
      ‘John introduced Hanako and Mary to each other’

Order-II (NP – PP – V)
   ‘John introduced Hanako and Mary to each other’ (Miyagawa 1997:5)
      ‘John introduced each other to Hanako and Mary’

Look at the scrambled sentence in (81a). Here, the theme NP Hanako-to Mari-o ‘Hanako-and Mary-Acc’ binds the reciprocal otagai ‘each other’ in the following PP. If we were to assume that short-scrambling involves (A’)-movement of NP and binding applies at LF, the theme NP may reconstruct to its original position at LF, from where it cannot bind the goal PP, contrary to the fact. The reciprocal binding effects shown in (80)-(81) support the right-branching structure resulting from incremental structure-building in the syntax, and are consistent with a base-generation approach to short-scrambling.

The Weak Crossover phenomena also provide supporting evidence for right-branching structure in Japanese. Weak Crossover is a typical diagnostic for configurationality. Based on the existence of a Weak Crossover phenomenon shown in (82), Hoji (1985, 1987) claims that Japanese VPs, independently of short-scrambling, have hierarchical structure:

(82) Order-I (PP – NP – V)
   John-Nom pro F-Acc gave teacher-to Bill-Acc had.meet
   ‘John had Bill; meet the teacher that gave him; an F’
   b. *Kimi-wa [pp proi F-o tuketa sensei-ni] [np darei-o] awaseta no?
      you-Top pro F-Acc gave teacher-to who-Acc had.meet Q
      ‘Who, did you have t, meet the teacher that gave him, an F?’ (Hoji 1987:177)

The name Bill can be coindexed with the preceding zero pronoun as in (82a), whereas the wh-phrase dare ‘who’ cannot be coindexed with the preceding zero pronoun as is shown in (82b). Hoji explains this contrast by LF raising of wh-phrases, which results in a Weak Crossover configuration:

(83) Weak Crossover configuration

Hoji’s explanation is based on the assumption that syntactic structure in Japanese is right-branching like English. Therefore, the Weak Crossover phenomenon independently motivates right-branching structure.

17 In Japanese, a zero pronoun, not the overt pronoun kare ‘he’, is used for the Weak Crossover test, because the pronoun kare cannot have a bound variable interpretation. See Hoji (1987:169-172) for discussion on this.
in Japanese, and is consistent with an incremental approach to structure-building in the syntax.

Regarding the relation between the short-scrambling alternation and the Weak Crossover phenomenon, Hoji (1985, 1987) argues that short-scrambling remedies Weak Crossover violations, based on examples like (84):

(84) a. Order-I (PP – NP – V)
   *Kimi-wa [pp pro, okuttekita hito-ni] [np nani-o] okurikaesita no?
you-Top pro sent-over person-to what-Acc sent.back Q
   ‘What did you send back to the person that had sent it to you?’
b. Order-II (NP – PP – V)
   Kimi-wa [np nani-o][pp pro, okuttekita hito-ni] okurikaesita no?
   ‘What did you send back to the person that had sent it to you?’ (Hoji 1987:178)

When the theme NP precedes the goal PP, as in the scrambled sentence in (84b), the zero pronoun can have a bound variable interpretation. This is predicted by incremental structure-building in the syntax: the scrambled sentence in (84b) as well as the non-scrambled sentence in (84a) have right-branching structure, and hence in (84b) the wh-phrase is always in a higher position than the zero pronoun. Therefore, LF-raising of the wh-phrase does not yield a Weak Crossover configuration. Thus, the contrast in (84) is explained independently of the issue of whether short-scrambling involves movement or not, because the contrast is predictable as long as both (84a) and (84b) have right-branching structure.

An argument in favor of a base-generation approach to short-scrambling comes from cases where the surface order exhibits a Weak Crossover configuration. When a wh-word follows a zero pronoun in the scrambled order, for example, a bound variable interpretation of the zero pronoun is unavailable.

(85) Order-II (NP – PP – V)
   *Kimi-wa [np pro, F-o tuketa sensei-o] [pp dare-ni] awaseta no?
you-Top pro F-Acc gave teacher-Acc who-Dat had.meet Q
   ‘Who did you have the teacher that gave him an F meet to?’

At LF, the wh-phrase dare-ni ‘who-Dat’ raises leftward to the highest sentence-initial position, crossing the zero pronoun. This results in the Weak Crossover configuration in (83). If we assume that the sentence is derived by (A’)-movement of the theme NP from the position after the wh-phrase (PP), it is wrongly predicted that the theme NP may reconstruct at LF, and therefore no Weak Crossover configuration arises.

An apparent counterexample to a base-generation approach to short-scrambling comes from scope reconstruction effects. Assuming that the scope relation between two elements is determined by the structural relation between them, a base-generation approach predicts that the preceding quantifier phrase always takes a wider scope over the following quantifier phrase. This is because the preceding element is always in a higher position than the following element in a right-branching configuration. The ambiguity of the non-scrambled sentence in (86b) and the scrambled sentence in (87b) appears to argue against this prediction, because the existential quantifier dareka ‘someone’ may take scope over the universal quantifier daremo ‘everyone’ to its left.

(86) Order-I (PP – NP – V)
      John-Nom someone-to everyone-Acc introduced
      ‘John introduced everyone to someone’
      (∃>∀, ∀>∃)
      (from Yatsushiro 1996, as cited in Williams 2003:158)
      ‘John introduced someone to everyone’
      (∀>∃, ∃>∀)

(87) Order-II (NP – PP – V)
'John introduced someone to everyone'

   'John introduced everyone to someone'

(3>V, ??V>3)

(4>V, 3>V)

(∀>3, 3>∀)

Ambiguous scrambled sentences such as (87b) have been taken as evidence for a movement analysis of short-scrambling (cf. Yatsushiro 1996). If the theme NP daremo-o 'everyone-Acc' is scrambled in front of the goal PP dareka-ni 'someone-to', leaving a trace, scope reconstruction should be possible by moving the phrase back into the trace position, and the ambiguity of the sentence is correctly predicted.

However, the ambiguity of scrambled sentences is not conclusive evidence for a movement approach to short-scrambling, because a scrambled sentence is not always ambiguous, as is seen in (87a), and a non-scrambled sentence can also be ambiguous, as in (86b). Therefore, the generalization that should be drawn from the paradigm in (86)-(87) is that in Japanese, a phrase with the existential quantifier (e.g. dareka 'someone') may take scope over another quantifier phrase, regardless of the surface word order (cf. Hoji 1986).

(88) a. ... ∀ ... 3 ...
   ∀ > 3, 3 > ∀ (ambiguous)

b. ... 3 ... ∀ ...
   3 > ∀

The generalization in (88) is not specific to Japanese. Kempson et al. (2001) point out that there is evidence that "quantifying expressions in natural language fall into two distinct classes according to their freedom with respect to scope construal" (p.225). The first class consists of non-indefinite quantifying expressions such as universal quantifiers ('every'), and 'most, few, almost, no', whose scope relations generally follow the sequences of words. The second class consists of indefinites such as existential quantifiers, which may all be construed as taking arbitrarily wide scope with respect to their order relative to other expressions in the string.18

If the scope relation is determined not only by the hierarchical relation between quantifiers but also by lexical properties of quantifiers themselves, scope interactions cannot be a reliable diagnostic for hierarchical relations. Therefore, the ambiguity of the scrambled sentence in (87b) does not constitute a counterexample to a base-generation approach to short-scrambling.19

Another apparent counterexample to a base-generation approach is found in bound variable pronouns, as discussed in Takano (1998). Takano argues for the specific version of Thematic Hierarchy, non-theme > theme, and claims that the scrambled order is derived via movement of a theme NP in front of a non-theme PP. His argument is based on paradigms such as (89)-(91) which involve bound variable pronouns. In the non-scrambled sentences in (a)-(b) in (89)-(91), a non-theme PP

18 The generalization in (88) holds not only for short-scrambling, but also for medial-scrambling:

   someone-to John-Nom everyone-Acc introduced
   'John introduced everyone to someone' (3>V, *∀>3)

   'John introduced everyone to someone'

('∀>3, *∀>3')

   'John introduced everyone to someone' (3>V, *∀>3)

(Yatsushiro 1996, as cited in Williams 2003:158)

19 Another factor that seems to influence scope interaction is focus: in the scrambled sentence in (87b), the existential quantifier takes a scope over the universal quantifier, only if the existential quantifier is focused and interpreted as a specific indefinite (Fodor and Sag 1982). (Thanks to Rose-Marie Dechaine for drawing my attention to this.) When a phrase with a universal quantifier is focused in an appropriate context, it is hard to obtain a wide scope reading of the existential quantifier. This is shown in (i):

(i) [context: John organized a party. As a host, he wanted to make sure that everybody at the party had someone to talk to and was having a good time. Some people came in late, and...]

Jon-wa [VP subeteno okurete kita hitsu-o] [VP dareka-ni] syookaisita.
John-Top every late came-in one-Acc introduced
   'John introduced everyone that came in late to someone'

(∀>3, #3>∀)
precedes a theme NP. In this case, a universal quantifier contained in the PP can bind a pronominal element contained in the NP, and the pronominal can have a bound variable interpretation as in (a), whereas a universal quantifier contained in the NP cannot bind a pronominal element contained in the PP, as is shown in (b). On the other hand, in the scrambled sentences in (c)-(d) in (89)-(91), a theme NP precedes a non-theme PP, and a universal quantifier can bind a pronominal element regardless of linear order between them.

   a. Mary-ga [PP subete-no gakusei-ni] [NP soitu-no sensei-o] syookaisita.
      Mary-Nom all-Gen student-to he-Gen teacher-Acc introduced
      ‘Mary introduced his teacher to every student’
   b. *Mary-ga [PP soitu-no sensei-ni] [NP subete-no gakusei-o] syookaisita.
      ‘Mary introduced every student to his teacher.’

Order-II (NP[theme] – PP[non-theme] – V)
   c. Mary-ga [NP subete-no gakusei-o] [PP soitu-no sensei-ni] syookaisita.
      ‘Mary introduced every student to his teacher’
   d. ?Mary-ga [NP soitu-no sensei-o] [PP subete-no gakusei-ni] syookaisita.
      ‘Mary introduced his teacher to every student’

   a. Mary-ga [PP subete-no kodomo-ni] [NP soitu-no fuku-o] kiseta.
      Mary-Nom all-Gen child-on he-Gen clothes-Acc put
      ‘Mary put his clothes on every child’
      Mary-Nom it-Gen owner-on all-Gen dress-Acc put
      ‘Mary put every dress on its owner’

Order-II (NP[theme] – PP[non-theme] – V)
   c. Mary-ga [NP subete-no doresu-o] [PP sore-no motinusi-ni] kiseta.
      ‘Mary put every dress on its owner’
   d. ?Mary-ga [NP soitu-no huku-o] [PP subete-no kodomo-ni] kiseta.
      ‘Mary put his clothes on every child’

   a. Mary-ga [PP subete-no gakusei-kara] [NP soitu-no syasin-o] karita.
      Mary-Nom all-Gen student-from he-Gen photo-Acc borrowed
      ‘Mary borrowed his picture from every student’
   b. *Mary-ga [PP sore-no tyosya-kara] [NP subete-no hon-o] karita.
      Mary-Nom it-Gen author-from all-Gen book-Acc borrowed
      ‘Mary borrowed every book from its author’

Order-II (NP[theme] – PP[non-theme] – V)
   c. Mary-ga [NP subete-no hon-o] [PP sore-no tyosya-kara] karita.
      ‘Mary borrowed every book from its author’
   d. ?Mary-ga [NP soitu-no syasin-o] [PP subete-no gakusei-kara] karita.
      ‘Mary borrowed his picture from every student’

(Takano 1998:828-829)

Based on this paradigm, Takano (1998) argues that the non-scrambled sentences in (a)-(b) are base-generated, whereas the scrambled sentences in (c)-(d) are derived via movement of a theme NP. Syntactic movement leaves a trace and hence a theme NP can reconstruct at LF to its original position lower than the non-theme PP, whereby the quantifier contained in the PP can bind the pronominal contained in the NP in the (d) sentences.

The crucial empirical contrast for Takano’s claim is the unacceptability of the (b) examples versus the acceptability of the (d) examples. However, there is inconsistency in his data. Notice that in (90)-(91),
the sentences do not form minimal pairs and the (b) and (d) sentences contain different lexical items. The
biggest problem with this is that the sentences contain different pronominal elements: the neutral pronoun
sore ‘it’ is used in (b) whereas the (masculine) epithet soitu ‘he/the guy’ is used in (d). This is problematic,
because the choice of pronominal elements influences the availability of bound variable interpretations.
For example, it seems that the epithet soitu is easier to be interpreted as a bound variable than the pronoun
sore. Look at (92):

   a. Mary-ga [PP subete-no sato-o-yo-a-ni] [PP soitu-no] kodomo-o miseta.
      Mary-Nom all-Gen foster-parent-to he-Gen child-Acc showed
      ‘Mary showed his child to every foster parent’
   b. ??Mary-ga [PP soitu-no sato-o-yo-a-ni] [NP subete-no] kodomo-o miseta.
      Mary-Nom he-Gen foster-parent-to all-Gen child-Acc showed
      ‘Mary showed every child to his foster parent’

Order-II (NP[theme] – PP[non-theme] – V)
   c. Mary-ga [NP subete-no kodomo-o] [PP soitu-no sato-o-yo-a-ni] miseta.
      ‘Mary showed every child to his foster parent’
   d. ?Mary-ga [NP soitu-no kodomo-o] [PP subete-no sato-o-yo-a-ni] miseta.
      ‘Mary showed his child to every foster parent’

If the epithet soitu is used instead of the pronoun sore ‘it’, the problematic (b) sentence improves.
Therefore, I argue that the Takano’s (1998) observation is not a matter of whether short-scrambling
involves movement or not. Rather, the contrast between (b) and (d) in (89)-(91) has more to do with the
lexical properties of pronominal elements. For the examples in (89), the contrast between (89b) and (89d)
is not as robust as Takano claims, and the exact degree of their acceptability varies between speakers.
Therefore, Takano’s paradigm does not necessarily constitute a counterexample to a base-generation
approach to short-scrambling.

To sum up the point of this section, the LF effects such as (i) reciprocal binding (e.g. (80)-(81))
and (ii) WCO effects (e.g. (84)-(85)) argue for the right-branching structure resulting from incremental
structure-building in the syntax, and argue against reconstruction of theme NPs, hence against a
movement approach to short-scrambling. Right-branching structure assigned to not only non-scrambled
but also scrambled sentences in Japanese is also consistent with the observations on (iii) scope interaction
(e.g. (86)-(87)) and (iv) bound variable pronominals (e.g. (89)-(91)), in that these data suggest that the
preceding phrase always c-commands the following phrase. Furthermore, we saw that the reverse scope
effect (e.g. (86b) and (87b)) and the bound variable interpretation of pronominal elements (e.g. (d)
sentences in (89)-(91)) do not necessarily constitute counterexamples to a base-generation approach to
short-scrambling. This is because they are influenced by lexical properties of quantifiers and pronominal
elements, respectively, and not determined solely by the syntactic position of relevant elements.

4.4. Situating the Prosodic Phase Hypothesis in Previous Analyses of XP-Scrambling

There are two types of analyses involved in the derivation of scrambled sentences in Japanese: one is to
derive the scrambled order (Order-I) from the corresponding non-scrambled order (Order-II) via
movement, either in the phonology or in the syntax; the other is to base-generate Order-II in the syntax.
The Prosodic Phase Hypothesis is a base-generation analysis: Order-II as well as Order-I are base-
generated in the syntax, and illegitimate orderings are filtered out by correspondence conditions on the
syntax-prosody mapping. This section reviews some representative analyses of XP-scrambling in
Japanese to highlight their differences from the Prosodic Phase Hypothesis.

4.4.1. Movement Approaches
4.4.1.1. Phonological Movement: Ross (1967/86)

Before various LF effects induced by short-scrambling in Japanese (such as the ones we have just seen in 4.3.2.2.2) were noticed, or when they were ignored as marked phenomena outside of the “core” computational component (Chomsky 1995:325), scrambling was treated as a stylistic rule in the phonological component. For example, Ross (1967/86) formulates scrambling (in Latin in particular) as a stylistic rule in (93).

(93) Scrambling (Ross 1986:51)

\[
\begin{array}{c}
\text{X} = \begin{cases}
\{\text{NP} \}
\begin{cases}
\{\text{VP} \}
\begin{cases}
\{\text{N} \}, \{\text{V} \}, \{\text{Adj} \}, \{\text{Adv} \}
\end{cases}
\end{cases}
\end{cases}
\end{array}
\]

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & 3 & 2 & 4
\end{array}
\]

Condition: Sj dominates 2 if and only if Sj dominates 3.

The scrambling rule in (93) optionally applies to the constituents 2 and 3, and permutes them under the condition that they are in the same clause (Sj). The rule in (93) seems consistent with the absence of prosodic markedness in the scrambled order (Order-II) in Japanese, in that the rule does not specify which order is base-generated or which order is derived. However, Ross assumes that the alternation is purely optional, as is indicated by “OPT”, and hence the rule in (93) does not explain the prosodic motivation for a particular order in narrow-focus sentences. LF effects associated with the short-scrambling alternation are not explained either, because the rule in (93) applies in the phonological component.

4.4.1.2. Syntactic Movement

4.4.1.2.1. Leftward Movement: Minimalist Approaches

In standard movement approaches of (short-)scrambling, the scrambled NP – PP – V order (Order-II) is derived from the corresponding PP – NP – V order (Order-I) via the optional application of leftward syntactic movement of the NP (see e.g. Hoji 1985, Saito 1985, Fukui 1993, Takano 1998 among others). Movement approaches assume that “UG allows optional movement (i.e. scrambling) in principle” (Takano 1998:856).

Optionality of scrambling as syntactic movement poses a problem within the Minimalist framework. The Minimalist Program proposes general principles of economy, which require that movement operations must be morphologically driven by the need to check some feature (Chomsky 1995:130). Suppose, for the sake of illustration, that the choice between Order-I and Order-II is optional in Japanese, as is in fact the case in broad-focus sentences. Under the Minimalist view, optionality in word order raises a problem for a movement approach, since scrambled sentences involve optional overt movement that has no driving force. Chomsky (1995) speculates that such sentences “may not really belong to the system we are discussing here […], the one that is concerned with Last Resort movement driven by feature checking” within the syntactic component (p.325). Therefore, in the Minimalist Program, scrambled sentences are considered to fall outside of the “core” computational component and not given an extensive discussion. More recently, Chomsky (1998) stipulates that stylistic operations that derive scrambled sentences might fall within the phonological component “because they are not feature-driven” (p.21).

Instead of dodging the issue of word order optionality by simply sweeping scrambling away from
the "core" computational component into the phonology alone, let us consider how a syntactic movement approach to scrambling could be modeled in the Minimalist framework. When two orders start from the same Numeration as in (94a) below, they are taken to belong to the same "reference set" and hence compete with each other (Chomsky 1995:227). In such a case, one order should be chosen over the other as more economical. When two orders start from different Numerations, one with α and one without α as in (94b) for example, they do not belong to the same reference set and hence both orders can be derived. In this case, however, the task to identify what is "α" in the Numeration that drives movement remains.

(94) Movement approaches to word order optionality in the Minimalist framework


Numeration: \{Y, X, V, \ldots\}
Base-generated: YP - XP - V (Order-I)
Derived: YP - XP - V (Order-I), and XP - YP - t_{NP} - V (Order-II), when movement applies

b. Obligatory syntactic movement, with optionality in the introduction of α into the Numeration that drives movement: Ishihara (2000) for medial-scrambling

Numeration: \{Y, X, V, \ldots\} “or” \{Y, X, V, α, \ldots\}
Base-generated: YP - XP - V YP - XP - V (Order-I) (Order-I)
Derived: YP - XP - V YP - XP - t_{NP} - V (Order-I) (Order-II)

Fukui’s (1993) treatment of scrambling is of type (94a): scrambled sentences are derived by optional leftward movement. In order to justify optionality in the Minimalist framework, Fukui proposes one specific measure of the cost of formal operations in the grammar, called the "Parameter Value Preservation" (PVP) measure:

(95) The Parameter Value Preservation (PVP) measure (Fukui 1993:400)
A grammatical operation (Move α, in particular) that creates a structure that is inconsistent with the value of a given parameter in a language is costly in the language, whereas one that produces a structure consistent with the parameter value is costless.

One specific interpretation of the PVP measure is that a costless application of Move α does not need a driving force (such as Case Filter and Spec-Head agreement) and can in principle be truly optional. The case in point for the word order alternation is the head-parameter. The PVP measure predicts that in
Japanese, a head-final language, leftward XP movement such as scrambling can be optional, whereas rightward XP movement must have a driving force. In fact, Saito (1985) argues that Japanese does not have rightward phrasal movement, and as a consequence, the language is strictly head-final. (But see Takami 1995 and Rosen 1996 for discussion on potential examples of rightward movement in Japanese.)

On the other hand, the PVP measure predicts that in English, a head-initial language, rightward XP movement can be optional, whereas leftward XP movement must have a driving force. This conforms to the standard assumption that leftward displacement in English (e.g. wh-movement and DP-movement) is grammatically motivated and hence obligatory, whereas rightward displacement (e.g. heavy NP shift and extraposition from NP) has a stylistic character and be optional in some sense (cf. 3.4).

The PVP measure is based on the Minimalist assumption that the computational component is influenced by economy considerations. In particular, the application of Move α is constrained by the PVP measure. However, Fukui (1993) simply assumes that scrambling in Japanese is purely optional, which is true for broad-focus, but not for narrow-focus cases. As a result, the PVP measure is merely an accommodation of optionality to a particular theoretical framework, and does not address the issue as to why the PVP measure is formulated as it is in (95).

I do not know of any (94b) type analysis for short-scrambling in Japanese. For medial-scrambling where an object appears to the left of a subject, Ishihara (2000) provides an obligatory syntactic movement analysis in the line of (94b). Ishihara identifies “α” in the Numeration in (94b) as the EPP feature, which is assigned to v under certain conditions. When the EPP feature is assigned to v, the object moves to [Spec, vP] in order to check this feature, yielding the medial-scrambling order (O-S-V). Importantly, the subject at the lower [Spec, vP] cannot move to this position because it is outside of the domain where the probe (= v in this case) looks for its goal. This is illustrated in (96):

(irrelevant details such as the structure above vP are omitted)

In line with Chomsky’s (1999) analysis of Scandinavian object shift, Ishihara claims that the EPP feature of v is assigned only when it has a semantic effect on the outcome: [Spec, vP] is the position where a certain interpretation (INT) is assigned (cf. Holmberg 1999, Chomsky 1999, Fox 2000). The relevant semantic effect is “definiteness” in Scandinavian object shift, and “defining the focus domain” in Japanese scrambling (Ishihara 2000:172-173).

Note here that the semantic effect that leads to the EPP feature assignment is not a motivation for medial-scrambling, but a necessary condition for it to take place. Behind this idea is Fox’s (2000) “Output Economy” which states that “optional operations must affect the output” (p.75). That is to say, all optional operations have either a phonological (e.g. word order) or a semantic (e.g. scope) effect. The INT can be regarded as a general term that covers semantic output effects. Although the Prosodic Phase Hypothesis acknowledges that the short-scrambling alternation in Japanese is partially determined by focus, the Prosodic Phase Hypothesis differs from Ishihara’s analysis in an important respect that the Prosodic Phase Hypothesis claims that linearization of verbal dependents is determined by the locus of prosodic prominence, which is affected by the locus of focus.

4.4.1.2.2. \textit{LF-Lowering: Bošković and Takahashi (1998)}

In contrast with a standard leftward movement analysis of XP-scrambling, Bošković and Takahashi
(1998) argue that scrambled phrases are base-generated in their surface non-θ positions and undergo obligatory LF-lowering to the position where they receive θ-roles. In the example in (97), the theme NP *sono-hon-o* ‘that book-Acc’ moves rightward at LF to the position where it receives the θ-role from the V.²⁰

(97) a. Order-II (scrambled order)
Ken-ga [vp sono hon-o [vp Naomi-ni ageta]] (koto)
   Ken-Nom that book-Acc Naomi-Dat gave (the.fact.that)
   ‘Ken gave Naomi that book’
⇒ LF

b. Order-I (non-scrambled order)
Ken-ga [vp Naomi-ni [vp sono-hon-o ageta]] (koto)
   ‘Ken gave Naomi that book’

The assumptions behind Bošković and Takahashi’s proposal are: (i) θ-roles are formal features that drive syntactic movement, (ii) LF is the only level where θ-criterion and the Projection Principle can be checked on Minimalist grounds, and (iii) movement does not have to leave a trace when no principle requires it (Lasnik and Saito 1992) and hence LF lowering as in (97b) does not interfere with the “Proper Binding Condition” that applies to traces (Fiengo 1974, Lasnik and Saito 1992).

The apparent conceptual advantage of Bošković and Takahashi’s proposal is that scrambling falls into the Minimalist framework, in that scrambled sentences are not the product of optional movement, but base-generated and involve obligatory LF-lowering, conforming to Last Resort. However, the flip side of this is that we are forced to allow base-generation of phrases in non-θ positions. Bošković and Takahashi (1998) speculate the possibility that whether a language allows scrambling or not depends on whether θ-roles (i.e. θ-features) are weak (Japanese) or not (English): “θ-roles need not be assigned or, more precisely, checked before LF in Japanese, whereas they must be assigned or checked in overt syntax in English” (p.352).²¹ In other words, Japanese has an option of base-generating the scrambled order (Order-II) as well as the non-scrambled order (Order-I). This is schematized in (98):

(98) Parametric differences in the nature of θ-roles and its relation to scrambling
a. Japanese type: optional scrambling (due to weak θ-roles)
   Base: [θ-position XP ] ‘or’ [non-θ-position XP ]
   Spelled out as: Order-I Order-II
   LF: [θ-position XP ] ⇒ obligatory LF-movement [non-θ-position _ ] [θ-position XP ]

b. English type: no scrambling (due to strong θ-roles)
   Base: [θ-position XP ]
   Spelled out as: canonical order
   LF: [θ-position XP ]

²⁰ Although the examples discussed in Bošković and Takahashi (1998) are all medial scrambling or long-distance scrambling, their analysis is extended to short-scrambling if we assume that the VP-adjoined position is a non-θ position.
²¹ See Fanselow (2000) for a similar analysis of German scrambling.
Thus, in Japanese type languages, (short-)scrambled as well as non-scrambled sentences are equally base-generated in their analysis, and the prosodic motivation for the short-scrambling alternation for narrow-focus sentences is not explained. It is also unclear how Bošković and Takahashi’s analysis accounts for the alternations in English such as heavy NP shift.

Notice further that the parametric differences given in (98) do not exhaust logical possibilities. It is predicted that there can be a language with weak θ-roles and does not allow the option of base-generating XPs in θ-positions, hence forces obligatory scrambling:

(99) **Obligatory Scrambling (due to weak θ-roles)**

Base: \[ \text{[non-θ-position XP]} \]

Spelled out as: scrambled order

LF: \[ \Rightarrow \text{obligatory LF-movement} \]

\[ \text{[non-θ-position \_]} \text{[θ-position XP]} \]

Polysynthetic languages where argument XPs are realized as syntactic adjuncts may fall into this category (Jelinek 1984, Baker 1996). Bošković and Takahashi’s analysis of Japanese scrambling is incomplete in that it does not tell us anything about the difference between Japanese type of scrambling languages in (98a) and obligatory scrambling languages in (99).

### 4.4.2. Base-Generation Approaches

Under the Prosodic Phase Hypothesis pursued in this thesis, both Order-I and Order-II are base-generated in the syntax incrementally via Merge. This analysis is based on the observation that Order-I and Order-II sentences exhibit the same intonation pattern, and the choice between them depends on the locus of a narrow-focus.

Miyagawa (1997) provides a base-generation approach to short-scrambling. However, Miyagawa’s argument for a base-generation approach is syntactic, not prosodic. He uses Rizzi’s (1986) Chain Condition on A-movement as a basis of the argument for base-generation of the scrambled XP – YP (–V) order. Let us see how his argument works.

First, Miyagawa argues that the unacceptable example in (100) shows that medial scrambling observes the Chain Condition:

(100) \[ \text{[Ken-to Mary-o]} \text{otagai-ga t} \text{mita.} \]

Ken-and Mary-Acc each.other-Nom t saw

‘Ken and Mary, each other saw’ (Miyagawa 1997:4)

The theme NP *Ken-to Mary-o* (‘Ken-and Mary-Acc’), which is the intended antecedent of the reciprocal *otagai* “each other”, appears in front of the reciprocal subject. If we assume that medial-scrambling involves movement, the reciprocal locally e-commands the trace of its antecedent thereby violating the Chain Condition and the unacceptability of (100) is predicted.

If short-scrambling also involves movement, the same violation is expected to arise. However, it does not: in (101), the theme NP *Yoko-to Naomi-o* ‘Yoko-and Naomi-Acc’ appears before the goal PP *otagai-ni* ‘each other-to’, and the sentence is acceptable.

22 Thanks to Rose-Marie Déchaine for discussion on the θ-role typology illustrated in (98)-(99).

23 The Chain Condition is defined as follows: \( C=\alpha_{1},...\alpha_{n} \) is a ‘chain’ iff, for \( 1 \leq i < n, \alpha_{i} \) is the local binder of \( \alpha_{i+1} ; \alpha \) is a “binder” of \( \beta \) iff, for \( \alpha, \beta = \text{any category}, \alpha \) and \( \beta \) are co-indexed, and \( \alpha \) e-commands \( \beta ; \alpha \) is the “local binder” of \( \beta \) iff \( \alpha \) is a binder of \( \beta \) and there is no \( \gamma \) such that \( \gamma \) is a binder of \( \beta \), and \( \gamma \) is not a binder of \( \alpha \) (Rizzi 1986:66).
Ken introduced Yoko and Naomi to each other (at the party) (Miyagawa 1997:5)

The contrast between (100) and (101) leads Miyagawa to argue that there is no trace of the theme NP in (101), and therefore short-scrambling involves no movement. Although he explicitly argues against “optional” scrambling as syntactic movement, optionality still exists in short-scrambling regarding the base-generation of Order-I or Order-II.

Yoshida (2000) furthers a base-generation approach to short-scrambling, and accounts for the optionality of scrambling by relaxing the locality of the theta-domain (cf. 4.3.2.2.1):

(102) Internal arguments of a lexical head (e.g. V or A) can be freely merged in their projection irrespective of their argument structure as long as the head finality holds. (Yoshida 2000:104)

Yoshida accounts for the intuitive distinction between Order-I and Order-II in terms of the semantic argument structure of Vs in the lexicon, following the framework of Lexical Decomposition Semantics (Wunderlich 1997). For example, a give-type ditransitive verb projects the agent-goal-theme order as the canonical order in Japanese. The semantic argument structure is mapped to syntactic word order transparently in an unmarked case, but can be mapped differently as long as (102) is satisfied in a marked case. Under this view, the scrambled order is marked yet still permitted. Although Yoshida’s analysis is compatible with the Prosodic Phase Hypothesis in attributing markedness of Order-II in broad-focus sentences solely to the lexical property of head verbs, it does not explain the cases where the Prosodic Economy Condition chooses one order over the other, overriding the lexical selectional property of verbs, as in narrow-focus cases.

4.4.3. Summary

The Prosodic Phase Hypothesis is a base-generation approach to Japanese XP-scrambling: the scrambled order as well as the non-scrambled order are based-generated in the syntax, and illegitimate orderings are filtered out by correspondence conditions on the syntax-prosody mapping. The Prosodic Phase Hypothesis is based on the observation that there is no prosodic markedness associated with the short-scrambling alternation, and the positional pattern of XP is determined by the Prosodic Economy Condition for Japanese. It differs from any movement or base-generation approaches reviewed in this section, which do not discuss the prosodic motivation for the short-scrambling alternation.

Appendix. Details of the Elicitation

A1. Purpose

A systematic elicitation was designed as a test of how natural or unnatural sentences with the PP – NP – V order (Order-I) and sentences with NP – PP – V order (Order-II) sound to native speakers of Japanese (i) with certain focus interpretations and (ii) with certain prosodic patterns (listening test). As a follow-up test, native speakers were asked to judge the written version of the stimuli sentences (written test).

A2. Procedure and Methodology

A2.1. Focus patterns

A question-answer format was used to construct sentences with certain focus interpretations. The
alternation involves the V moratta ‘received’, which takes a theme NP and a source PP. For the word order alternation between PP – NP – V and NP – PP – V, five types of focus interpretations were examined: a broad-focus (i.e. IP-focus) interpretation, two single narrow-focus (NP-focus and PP-focus) interpretations, and two double narrow-focus (NP+V-focus and PP+V-focus) interpretations.

Stimuli sentences

Both (A1) and (A2) sentences are prosodically controlled: the (A1) sentences do not have any focal stress, and the (A2) sentences contain focal stress on a narrow-focus.

The following notational conventions are used:

default
lexical item in bolding carries default sentence stress
HEAVY
lexical item in upper case letters carries extra prosodic prominence
focused
intended narrow-focus is underlined


(1) IP-focus
Order-I
Q: Kinou nani-ga okita tte?
yesterday what-Nom happened Q
‘What happened yesterday?’
A1: Mayumi-ga Mamoru-ni nuigurumi-o moratta yo.
Mayumi-Nom Mamoru-from . doll-Acc received yo
‘Mayumi received a doll from Mamoru’

Order-II
Q: Kinou nani-ga okita tte?
A1: Mayumi-ga nuigurumi-o Mamoru-ni moratta yo.
(2) NP-focus
Order-I
Q: Mamoru-to Mayumi-ga nanika-o motteiru keredo,
Mamoru-and Mayumi-Nom something-Acc having and
karera-wa nani-o siteiru no?
they-Top what-Acc doing Q
‘Mamom and Mayumi are holding something. What are they doing?’
A1: Mayumi-ga Mamoru-ni nuigurumi-o moratta mitai yo.
Mayumi-Nom Mamoru-from doll-Acc received it seems yo
‘It seems that Mayumi received a doll from Mamoru’

Q: Mamoru-to Mayumi-ga nanika-o motteiru keredo, karera-wa nani-o siteiru no?
A2: Mayumi-ga Mamoru-ni NUIGURUMI-o moratta mitai yo.
Order-II
Q: Mamoru-to Mayumi-ga nanika-o motteiru keredo, karera-wa nani-o siteiru no?
A1: Mayumi-ga nuigurumi-o Mamoru-ni moratta mitai yo.
Q: Mamoru-to Mayumi-ga nanika-o motteiru keredo, karera-wa nani-o siteiru no?
A2: Mayumi-ga NUIGURUMI-o Mamoru-ni moratta mitai yo.
(3) PP-focus
Order-I
Q: Mayumi-ga daijisouni nuigurumi-o kakaeteiru keredo,
Mayumi-Nom carefully doll-Acc holding.is and
dareka-kara moratta no?
someone-from received Q
‘Mayumi is holding a doll carefully. Did she receive it from someone?’

A1: Kanojo-wa Mamoru-ni are-o moratta rassii yo.
she-Top Mamoru-from that-Acc received they.say yo
‘They say that she received it from Mamoru’

Q: Mayumi-ga daijisouni nuigurumi-o kakaeteiru keredo, dareka-kara moratta no?
A2: Kanojo-wa MAMORU-ni are-o moratta rassii yo.

Order-II
Q: Mayumi-ga daijisouni nuigurumi-o kakaeteiru keredo, dareka-kara moratta no?
A1: Kanojo-wa are-o Mamoru-ni moratta rassii yo.
Q: Mayumi-ga daijisouni nuigurumi-o kakaeteiru keredo, dareka-kara moratta no?
A2: Kanojo-wa are-o MAMORU-ni moratta rassii yo.

(4) NP+V-focus
Order-I
Q: Mamoru-to Mayumi-ga nakayosasouni mieru keredo, karera-ni Mamoru-and Mayumi-Nom intimate look and they-to nani-ga atta no?
what-Nom happened Q
‘Mamoru and Mayumi look intimate. What happened to them?’
A1: Mayumi-ga Mamoru-ni nuigurumi-o moratta mitai yo.
Mayumi-Nom Mamoru-from doll-Acc received it.seems yo
‘It seems that Mayumi received a doll from Mamoru’
Q: Mamoru-to Mayumi-ga nakayosasouni mieru keredo, karera-ni nani-ga atta no?
A2: Mayumi-ga Mamoru-ni NUIGURUMI-o MORATTA mitai yo.

Order-II
Q: Mamoru-to Mayumi-ga nakayosasouni mieru keredo, karera-ni nani-ga atta no?
A1: Mayumi-ga nuigurumi-o Mamoru-ni moratta mitai yo.
Q: Mamoru-to Mayumi-ga nakayosasouni mieru keredo, karera-ni nani-ga atta no?
A2: Mayumi-ga NUIGURUMI-o Mamoru-ni MORATTA mitai yo.

(5) PP+V-focus
Order-I
Q: Mayumi-ga daijisouni nuigurumi-o kakaeteita keredo, Mayumi-Nom carefully doll-Acc holding.was and nazeka sitteru?
why know
‘Mayumi was holding a doll carefully. Do you know why?’
A1: Kanojo-wa Mamoru-ni are-o moratta rassii yo.
she-Top Mamoru-from that-Acc received they.say yo
‘They say that she received it from Mamoru’
Q: Mayumi-ga daijisouni nuigurumi-o kakaeteita keredo, nazeka sitteru?
A2: Kanojo-wa MAMORU-ni are-o MORATTA rassii yo.

Order-II
Q: Mayumi-ga daijisouni nuigurumi-o kakaeteita keredo, nazeka sitteru?
A1: Kanojo-wa are-o Mamoru-ni moratta rassii yo.
Q: Mayumi-ga daijisouni nuigurumi-o kakaeteita keredo, nazeka sitteru?
A2: Kanojo-wa are-o MAMORU-ni MORATTA rassii yo.

A2.2. Prosodic Patterns

I had two native speakers of Japanese read the stimuli sentences: the speaker AF (male) pronounced the
(Q) sentences and the speaker KS (female) pronounced the (A) sentences. Both were graduate linguistics students at UBC. The recording was done at separate times for (Q) sentences and (A) sentences, and the (Q)-(A) sentences were combined later by KS using the Praat program. For each order in each context, two prosodically controlled answers were constructed: the (A1) sentences without any focal stress but with default sentence stress on the phrase in the default sentence stress position, and the (A2) sentences with focal stress on a narrow-focus. Look at the NP-focus sentences in (2), for example:

(2) NP-focus
   Order-I
   A1: Mayumi-ga Mamoru-ni nuigurumi-o moratta mitai yo.
      Mayumi-Nom Mamoru-from doll-Acc received it.seems yo
   ‘It seems that Mayumi received a doll from Mamoru’
   A2: Mayumi-ga Mamoru-ni NUIGURUMI-o moratta mitai yo.
   Order-II
   A1: Mayumi-ga nuigurumi-o Mamoru-ni moratta mitai yo.
   A2: Mayumi-ga NUIGURUMI-o Mamoru-ni moratta mitai yo.

The (A1) sentences do not contain focal stress and exhibit the default intonation pattern. On the other hand, focal stress with a higher pitch peak was intentionally assigned to the NP in the (A2) sentences. Double narrow-focus sentences were constructed in the same way.

(4) NP+V-focus
   Order-I
   A1: Mayumi-ga Mamoru-ni nuigurumi-o moratta mitai yo.
      Mayumi-Nom Mamoru-from doll-Acc received it.seems yo
   ‘It seems that Mayumi received a doll from Mamoru’
   A2: Mayumi-ga Mamoru-ni NUIGURUMI-o MORATTA mitai yo.
   Order-II
   A1: Mayumi-ga nuigurumi-o Mamoru-ni moratta mitai yo.
   A2: Mayumi-ga NUIGURUMI-o Mamoru-ni MORATTA mitai yo.

Notice that the two occurrences of focal stress are adjacent to each other in Order-I, whereas they are non-adjacent in Order-II.

Subjects

Three native speakers of Japanese (AF, SU, MM) participated in the listening test, and six native speakers of Japanese (AF, SU, TT, MK, AT, NH) participated in the written test. The number of (A) sentences judged by subjects was eighteen in the listening test, and ten in the written test (because two sentences in each four narrow-focus sentences are the same when written). Two of the subjects (= AF, SU) participated in both listening and written tests, and the written test was conducted about ten days after the listening test. In the listening test, two out of three (= AF, SU) were linguists, and the other (= MM) was not. In the written test, all the six subjects were linguists. None of them were told the purpose of the test. (Note that the subject AF was the one who pronounced the (Q) sentences in constructing stimuli sentences, but the person did not know or hear the (A) sentences until participating in the tests, and was not told the purpose of the tests either.)

Procedure

In the listening test, the subjects were asked to evaluate the naturalness of (A) sentences as answers to the corresponding (Q) sentences in six scales: 6 = completely natural; 1 = completely unnatural. The sound
files were presented as speaker icons on Microsoft Powerpoint program, with each Q-A pair corresponding to each icon. A subject put on a headset connected to the computer and heard a Q-A pair by clicking on an icon. Subjects were allowed to listen to any pair of sentences anytime and as many times as they wanted.

In the written task, the stimuli sentences were presented in a Microsoft Word format on a computer screen.

**Equipment**

The stimuli sentences were recorded on a portable Marantz cassette recorder, sent to a PC computer, and saved as sound files. Recorded files were digitized and analyzed using the Praat program.

**A3. Specific Predictions**

The table in (6) summarizes the specific predictions for naturalness judgments in the listening test.

(6) Predictions for naturalness of Order-I and Order-II in the listening test

<table>
<thead>
<tr>
<th>(i) focus pattern and (ii) prosodic pattern</th>
<th>naturalness judgment</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) IP-focus</td>
<td></td>
</tr>
<tr>
<td>Order-I A1: PP - NP - V</td>
<td>6</td>
</tr>
<tr>
<td>Order-II A1: PP - NP - V</td>
<td>6</td>
</tr>
<tr>
<td>(2) NP-focus</td>
<td></td>
</tr>
<tr>
<td>I A1: PP - NP - V</td>
<td>6</td>
</tr>
<tr>
<td>A2: PP - NP - V (focal stress on NP)</td>
<td>1-5 (*1)</td>
</tr>
<tr>
<td>II A1: NP - PP - V</td>
<td>1 (*2)</td>
</tr>
<tr>
<td>A2: NP - PP - V (focal stress on NP)</td>
<td>6</td>
</tr>
<tr>
<td>(3) PP-focus</td>
<td></td>
</tr>
<tr>
<td>I A1: PP - NP - V</td>
<td>1 (*2)</td>
</tr>
<tr>
<td>A2: PP - NP - V (focal stress on PP)</td>
<td>6</td>
</tr>
<tr>
<td>II A1: NP - PP - V</td>
<td>6</td>
</tr>
<tr>
<td>A2: NP - PP - V (focal stress on PP)</td>
<td>1-5 (*1)</td>
</tr>
<tr>
<td>(4) NP+V-focus</td>
<td></td>
</tr>
<tr>
<td>I A1: PP - NP - V</td>
<td>6</td>
</tr>
<tr>
<td>A2: PP - NP - V (adjacent focal stress on NP+V)</td>
<td>1-5 (*1)</td>
</tr>
<tr>
<td>II A1: NP - PP - V</td>
<td>1 (*2)</td>
</tr>
<tr>
<td>A2: NP - PP - V (non-adjacent focal stress on NP...V)</td>
<td>1</td>
</tr>
<tr>
<td>(5) PP+V-focus</td>
<td></td>
</tr>
<tr>
<td>I A1: PP - NP - V</td>
<td>1 (*2)</td>
</tr>
<tr>
<td>A2: PP - NP - V (non-adjacent focal stress on PP...V)</td>
<td>1</td>
</tr>
<tr>
<td>II A1: NP - PP - V</td>
<td>6</td>
</tr>
<tr>
<td>A2: NP - PP - V (adjacent focal stress on PP+V)</td>
<td>1-5 (*1)</td>
</tr>
</tbody>
</table>

The prediction is that the order where (a part of) a narrow-focus occupies the default sentence stress position should be natural (= 6 in the scale) and the other order should not (= 1 in the scale). Therefore, in broad-focus (i.e. IP-focus) sentences, where there is no narrow-focus, it is predicted that the two orders should be equally natural.
When (a part of) a narrow-focus occupies the default sentence stress position, the narrow-focus should carry only default sentence stress, not focal stress, conforming to the Prosodic Economy Condition for Japanese (see *1 in the table (6)).

When (a part of) a narrow-focus is not in the default sentence stress position, the sentence with the default intonation pattern should be unnatural (see *2 in the table (6)). This is because in this case, the narrow-focus does not carry default sentence stress or focal stress, and the intended focus interpretation cannot be obtained.

In the written test, where the prosodic pattern of stimuli sentences is not controlled, the order involving non-adjacent focus may be judged to be natural, to the extent that the subject is not conscious of the prosodic pattern.

A4. Results

A4.1. Listening Test

The results of naturalness judgment by three native speakers of Japanese (AF, SU, MM) in the listening test are summarized in the table (7):

<table>
<thead>
<tr>
<th>Focus pattern</th>
<th>Word order</th>
<th>Order-I (PP – NP – V)</th>
<th>Order-II (NP – PP – V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad-focus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single Narrow-focus</td>
<td>NP</td>
<td>A1: default stress on NP 6-5 (ave. 5.7)</td>
<td>A1: default stress on NP 6-5 (ave. 5.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt; A2: focal stress on NP 5-4 (ave. 4.7)</td>
<td>= A2: focal stress on NP 6-3 (ave. 4.7)</td>
</tr>
<tr>
<td></td>
<td>PP</td>
<td>A1: default stress on PP 5-4 (ave. 4.7)</td>
<td>A2: focal stress on PP 6-2 (ave. 4.3)</td>
</tr>
<tr>
<td>Double Narrow-focus</td>
<td>NP+V</td>
<td>A1: default stress on NP 6-5 (ave. 5.7)</td>
<td>A1: default stress on NP 6-2 (ave. 4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt; A2: adjacent focal stresses on NP and V 5-4 (ave. 4.7)</td>
<td>= A2: adjacent focal stresses on NP and V 5-2 (ave. 3.3)</td>
</tr>
<tr>
<td></td>
<td>PP+V</td>
<td>A1: default stress on PP 5-4 (ave. 4.3)</td>
<td>A2: adjacent focal stresses on PP and V 5-4 (ave. 4.3)</td>
</tr>
</tbody>
</table>

Firstly, Order-I is judged to be better than Order-II with broad-focus in (1).
Secondly, for narrow-focus sentences, the order where (a part of) a narrow-focus occupies the default sentence stress position is judged to be better than the other order. Among them, the narrow-focus with default sentence stress in (A1) is more natural than that with focal stress in (A2). For example, for the
PP – NP – V order (Order-I) with NP-focus in (2), the (A1) pattern is preferred over the (A2) pattern; for the NP – PP – V order (Order-II) with PP-focus in (3), the (A1) pattern is preferred over the (A2) pattern.

Thirdly, for double narrow-focus sentences, non-adjacent occurrences of focal stress (e.g. the (A2) pattern with the NP – PP – V order with NP+V-focus in (3), and the (A2) pattern with PP – NP – V order with PP+V-focus in (4)), are judged to be less natural, compared with the corresponding (A2) pattern with the order, where occurrences of focal stress are adjacent to each other; also compared with the corresponding (A1) pattern with the same order, which does not employ focal stress.

When a narrow-focus does not occupy the default sentence stress position in a single narrow-focus sentence (e.g. the NP – PP – V order with NP-focus in (2), and the PP – NP – V order with PP-focus in (3)), no preference between the prosodic pattern (A1) and (A2) is found, despite the prediction that the (A1) pattern should be impossible in such cases (see *2 in the table (6)).

When a part of narrow-focus does not occupy the default sentence stress position in a double narrow-focus sentence (e.g. NP – PP – V order with NP+V-focus in (4), and PP – NP – V order with PP+V-focus in (5)), the prosodic pattern (A1) is more natural than the pattern (A2), despite the prediction that the (A1) pattern should be impossible in such cases (see *2 in the table (6)).

More details

(8) Results of the listening test

<table>
<thead>
<tr>
<th>Focus-pattern</th>
<th>Subject</th>
<th>AF</th>
<th>SU</th>
<th>MM</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) IP-focus</td>
<td>Order-I A1: PP - NP - V</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Order-II A1: NP - PP - V</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>(2) NP-focus</td>
<td>I     A1: PP - NP - V</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II    A1: NP - PP - V</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II    A2: NP - PP - V</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>(3) PP-focus</td>
<td>I     A1: PP - NP - V</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II    A2: NP - PP - V</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II    A2: PP - NP - V</td>
<td>6</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>(4) NP+V-focus</td>
<td>I     A1: PP - NP - V</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II    A2: NP - PP - V</td>
<td>4</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>II    A2: PP - NP - V</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>(5) PP+V-focus</td>
<td>I     A1: PP - NP - V</td>
<td>4</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II    A2: PP - NP - V</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>II    A1: NP - PP - V</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II    A2: NP - PP - V</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
The results of naturalness judgment by five native speakers of Japanese (AF, SU, TT, KM, AT, NH) in the written test are summarized in the table (9):

(9) Results of the written test

<table>
<thead>
<tr>
<th>Focus-pattern</th>
<th>Subject</th>
<th>AF</th>
<th>SU</th>
<th>TT</th>
<th>MK</th>
<th>AT</th>
<th>NH</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) IP-focus</td>
<td>Order-I A1: PP – NP – V</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Order-II A1: NP – PP – V</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>(2) NP-focus</td>
<td>I A1: PP – NP – V</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>II A1: NP – PP – V</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>(3) PP-focus</td>
<td>I A1: PP – NP – V</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>II A1: NP – PP – V</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(4) NP+V-focus</td>
<td>I A1: PP – NP – V</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>II A1: NP – PP – V</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>(5) PP+V-focus</td>
<td>I A1: PP – NP – V</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>II A1: NP – PP – V</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

Firstly, for broad-focus sentences in (1), Order-I is judged to be completely natural (= 6) by all the six subjects. The preference for Order-I over Order-II is present but not as robust as in the listening test.

Secondly, for single narrow-focus sentences in (2)-(3), the order where a narrow-focus occupies the default sentence stress position is judged to be better than the other: PP – NP – V order (Order-I) is preferred over Order-II for NP-focus in (2), and NP – PP – V order (Order-II) is preferred over Order-I for PP-focus in (3) except for one speaker (MK).

Thirdly, for double narrow-focus sentences, the order where the two foci are adjacent to each other is consistently preferred over the other order where they are non-adjacent: Order-I is preferred over Order-II for NP+V-focus in (4), and Order-II is preferred over Order-I for PP+V-focus in (5).
CHAPTER V
Implications and Conclusions

In the preceding chapters, I have presented evidence and arguments in favor of the following generalization: linear ordering between dependents in the verbal domain in English and Japanese is determined by correspondence conditions on the syntax-prosody mapping. This claim is the "Prosodic Phase Hypothesis". Assuming that the computational component is strictly derivational rather than representational, I argued that syntactic objects are built incrementally, paired with prosodic objects, and then spelled out into the phonological component. This conclusion has been empirically motivated by a close analysis of prosodic properties associated with English heavy NP shift and Japanese short-scrambling. The Prosodic Phase Hypothesis makes prosody a key part of an explanatory model of the grammar and goes beyond a purely syntax-oriented view of linearization. The methodology that I have employed in reaching this conclusion has consisted of acceptability judgments of orally presented sentences. This way, I have succeeded in taking the prosody of stimuli sentences into consideration.

The data dealt with so far has been limited to English and Japanese. This chapter goes beyond those two languages and explores some further cross-linguistic predictions of the Prosodic Phase Hypothesis. Section 5.1 provides a summary of the thesis. In 5.1.1, I summarize the general principles proposed in the preceding chapters, and consider their implications from a broader typological perspective. In section 5.1.2, I recapitulate the theoretical significance of the Prosodic Phase Hypothesis. Section 5.2 considers further issues regarding the prosodically based approach to linearization. In particular, I attempt to reduce cross-linguistic differences in linearization possibilities to a "lexical pitch parameter", i.e. whether lexically specified pitch features in a particular language is distinctive or not.

5.1. Summary of the Thesis

5.1.1. Summary from a Typological Perspective

5.1.1.1. Syntax-Prosody Mapping Algorithm

Let us start by reviewing the grammatical principles that I claim hold across languages. First and foremost, the main claim of this thesis is the Prosodic Phase Hypothesis in (1), which is represented in the model of the language faculty illustrated in (2).

(1) The Prosodic Phase Hypothesis
A syntactic object is spelled out as a prosodic object.

(2) The model of the language faculty in the Prosodic Phase Hypothesis
Adopting a strictly derivational approach to the computational component, syntactic structure must be built incrementally from left to right in the order in which syntactic words are pronounced (Phillips 1996, 2003), so syntactic objects (SOs) can be derivationally paired with prosodic objects (POs). The conditions on syntactic and prosodic structure-building are summarized in (3) below. Syntactic structure is built from left to right by merging a syntactic word at the right edge of the structure. Since POs are mapped from SOs, the way prosodic structure is built follows the way syntactic structure is built. However, due to prosody-specific conditions such as non-recursivity of Intonational Phrases (IntPs), a prosodic word is introduced to the root node of the structure, and accordingly prosodic structure allows multiple branching.

(3) Conditions on syntactic and prosodic structure-building

<table>
<thead>
<tr>
<th>Operation</th>
<th>Syntactic structure</th>
<th>Prosodic structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary unit</td>
<td>syntactic word</td>
<td>prosodic word</td>
</tr>
<tr>
<td>Branching</td>
<td>binary</td>
<td>multiple</td>
</tr>
<tr>
<td>Merge</td>
<td>Direction</td>
<td></td>
</tr>
<tr>
<td></td>
<td>left-to-right</td>
<td>left-to-right</td>
</tr>
<tr>
<td></td>
<td>Locus</td>
<td>any accessible target</td>
</tr>
</tbody>
</table>

The syntax constructs a series of SOs (SO₁, SO₂…), each of which can potentially be realized by a corresponding PO (PO₁, PO₂…). However, only certain SO-PO pairs (e.g. {SO₂, PO₂} and {SO₄, PO₄} in (2)) are directly spelled out and realized by the phonology. Originally, Chomsky (1998, 1999, 2001) proposed a “Phase” as the unit of spell-out and defined it in syntactic terms as either CP or vP. Under the Prosodic Phase Hypothesis, Phase is a derivational SO-PO pairing operation, and the unit of spell-out is a pair consisting of an SO and a PO, where the PO is an Intonational Phrase (IntP). The size of an IntP varies cross-linguistically depending on the general prosodic properties of a particular language.

The SO-PO mapping is implemented differently depending on the size of the PO. When the relevant PO is IntP, the SO-PO mapping is implemented by the alignment constraints given in (4) below. In broad-focus sentences where no focal stress is involved, a clause (CP/IP) corresponds to an IntP in English. The size of an IntP is smaller than a clause in languages with distinctive lexical pitch. For example, in Japanese, lexical pitch accent is distinctive, and the left edge of a syntactic phrase (XP) starts a

(4) **SO-PO mapping for maximal prosodic objects (= Intonational Phrases)**

<table>
<thead>
<tr>
<th>Edge</th>
<th>SO$^\text{max}$ (i.e. CP/IP)</th>
<th>SO$^\text{max-min}$ (i.e. XP)</th>
<th>SO$^\text{min}$ (i.e. X)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Align (SO, R; PO, R) and Align (SO, L; PO, L)</td>
<td>English, German, Dutch, Portuguese etc.</td>
<td></td>
<td>Tone languages?</td>
</tr>
<tr>
<td>ii. Align (SO, L; PO, L)</td>
<td>Japanese, Korean etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iii. Align (SO, R; PO, R)</td>
<td>O'odham</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In terms of size, syntactic clauses are maximal SOs (= SO$^\text{max}$) and syntactic phrases are neither maximal or minimal (= SO$^\text{max-min}$). Since there are languages where a maximal SO (i.e. CP/IP) corresponds to an IntP, and there are languages where an intermediate SO (i.e. XP) corresponds to an IntP, we might expect to find languages where a minimal SO (= SO$^\text{min}$), i.e. a syntactic head X, can correspond to an IntP. Tone languages, which can have a tonal melody within a word or a syllable, might belong to this category. However, this is not trivial because tone languages can also have intonation which may modify lexical tones (Shen 1990:31).

The edge of the SO-PO alignment is also subject to parametric variation. For example, the intonation contour of Tohono O'odham (formerly known as Papago) sentences consists of a sequence of instances of the tonal pattern identified as (L)HL, and it is the right edge, not the left edge, of a syntactic phrase that corresponds to the right edge of the (L)HL unit (Hale and Selkirk 1987).

The mapping from SO to PO may take place anytime after syntactic structure-building begins. The elementary unit of PO-building is a prosodic word, which is prosodically identified as the domain of lexical stress. When the relevant PO is a prosodic word, the SO-PO mapping is implemented by the alignment constraints given in (5) below, which specify that the edge of a lexical word (Lex) is aligned with the edge of a prosodic word. In English, alignment occurs at the right edge, and in Japanese, alignment occurs at the left edge.

(5) **SO-PO mapping for minimal prosodic objects (= prosodic words)**

<table>
<thead>
<tr>
<th>Edge</th>
<th>Lexical word (Lex)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Align (SO, R; PO, R)</td>
<td>English, German, French, Shona etc.</td>
</tr>
<tr>
<td>ii. Align (SO, L; PO, L)</td>
<td>Japanese, Korean, Shanghai etc.</td>
</tr>
</tbody>
</table>

German, French and Shona belong to the same category as English, and Korean and Shanghai (Chinese) belong to the same category as Japanese (Hale and Selkirk 1987:176-178).

The conditions on syntactic and prosodic structure-building summarized in (3) as well as the SO-PO mapping constraints in (4)-(5) concern sentences with the default intonation pattern (e.g. broad-focus sentences). In such sentences, the rightmost lexical stress in the sentence is defined as default sentence

---

1 The other logical possibilities are: Align (SO, R; PO, L) and Align (SO, L; PO, R), where there is a mis-alignment of the edges between SOs and POs. Whether these patterns exist or not remains to be investigated.
stress. In particular, the rightmost prosodic word in the rightmost IntP carries default sentence stress in e.g. English, Italian (Ladd 1996:191-192, Donati and Nespor 2003), Spanish, French (Zubizarreta 1998), and Bengali (Hayes and Lahiri 1991). In Japanese and Korean (Ladd 1996:196), it is the leftmost prosodic word in the rightmost IntP that carries default sentence stress, specifically because the sentence-final verb is not prominent prosodically.

(6) Default Sentence Stress Assignment
Default sentence stress is assigned to a prosodic word in the rightmost Intonational Phrase (= POₙ).

<table>
<thead>
<tr>
<th>Edge</th>
<th>Domain</th>
<th>POₙ</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Right (i.e. rightmost prosodic word)</td>
<td>English, Italian, Spanish, French, Bengali etc.</td>
<td></td>
</tr>
<tr>
<td>ii. Left (i.e. leftmost prosodic word)</td>
<td>Japanese, Korean etc.</td>
<td></td>
</tr>
</tbody>
</table>

The locus of default sentence stress may be more variable in some languages than in others (Ladd 1996:193-195). For example, in e.g. German, Dutch, and Turkish, default sentence stress may fall either on the rightmost prosodic word, e.g. verb in head-final contexts, or on the preceding noun (cf. Cinque 1993, Ladd 1996, Reinhart 1996, Neeleman and Reinhart 1998, Zubizarreta 1998). This contrasts with Japanese, where default sentence stress never falls on the rightmost prosodic word (i.e. verb).

Remember that default sentence stress is not necessarily more prominent than normal lexical stress. In English, default sentence stress is sometimes only relatively more prominent than anaphoric (and hence deaccented) elements, and therefore is not necessarily the most prominent prosodic peak in a sentence. But in Japanese, default sentence stress boosts the pitch peak of lexical High tone, and hence is prosodically distinct. (See (iii) vs. (iv) in the Disanaphoricity Scale in (10) below for comparison.) However, even in languages like English, default sentence stress is still theoretically significant, in that it explains the availability of broad-focus interpretations as an instance of focus projection.

Next, let us review generalizations about non-default cases. In narrow-focus sentences, for example, the occurrence of extra prosodic prominence (e.g. focal stress) systematically affects the intonation pattern of sentences. In English, the right edge of a prosodic word with extra prosodic prominence often induces an intonational break and hence gets aligned with the right edge of an IntP. The same observation is reported for Spanish, Italian (Nespor and Vogel 1986, Zubizarreta 1998:84), and Russian (Pereltsvaig 2004). In Japanese, the left edge of a prosodic word with extra prosodic prominence starts a new IntP, and all the following elements are deaccented and incorporated into the same IntP. Ladd (1996), citing Jun (1993), reports the same effect in Korean (Ladd 1996:196). The alignment constraint in (7i) formulates the effect of focal stress in English, and the alignment constraint in (7ii) together with the Prosodic Condition in (8) specify the effect of focal stress in Japanese:

(7) SO-PO mapping for maximal prosodic objects (= Intonational Phrases) with [Focus]

<table>
<thead>
<tr>
<th>Edge</th>
<th>The type of SO</th>
<th>SO with [Focus]SO</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Align (SO, R; PO, R)</td>
<td>English, Spanish, Italian, Russian etc.</td>
<td></td>
</tr>
<tr>
<td>ii. Align (SO, L; PO, L)</td>
<td>Japanese, Korean etc.</td>
<td></td>
</tr>
</tbody>
</table>
(8) **The Prosodic Condition on Japanese [Focus]₀**

In Japanese, [Focus]₀ must be in the rightmost Intonational Phrase (= PO₀).

Concerning the deaccenting phenomenon in Japanese, I argued that deaccenting after focal stress can be regarded as a general phonological phenomenon across languages when we limit the domain of deaccenting to IntPs.

(9) **Deaccenting**

Any elements that follow a prosodic word with extra prosodic prominence are deaccented within an Intonational Phrase.

In the (7i) type languages like English, the deaccenting phenomenon is not visible, because a prosodic word with extra prosodic prominence is always at the right edge of an IntP and there is no element to be deaccented. In the (7ii) type languages like Japanese, the deaccenting phenomenon has the effect of extending an IntP and eliminating following intonational break(s). The Prosodic Condition in (8) formalizes this effect.

Remember that the SO-PO mapping constraints for IntPs with [Focus] in (7) are vacuously satisfied in sentences with the default intonation pattern. In English, the prosodic word that carries default sentence stress is at the right edge of the whole sentence (IP), and hence at the right edge of an IntP (cf. (6i)). In Japanese, a prosodic word that carries default sentence stress is at the left edge of the rightmost IntP (cf. (6ii)). Therefore, the SO-PO mapping constraints for IntPs with [Focus] in (7) apply as a specific case, and the SO-PO mapping constraints for IntPs in general in (4) apply elsewhere.

The SO-PO mapping constraints in (7) are relevant to the cases where not only focused elements but also anaphoric elements are overtly expressed in a sentence. However, remember that the most natural way of expressing anaphoric elements in Japanese is to delete them, which is prosodically the weakest process of expressing anaphoric elements. Focused elements are semantically more prominent than anaphoric elements, and semantic prominence is correlated with prosodic prominence. The Disanaphoricity Scale in (10) describes the various prosodic devices employed by natural languages to encode disanaphoricity:

(10) **Disanaphoricity Scale**

<table>
<thead>
<tr>
<th>anaphoric</th>
<th>focused/disanaphoric</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) deletion</td>
<td>(i) null NP</td>
</tr>
<tr>
<td>(ii) deaccenting</td>
<td>(ii) ---</td>
</tr>
<tr>
<td>(iii) normal prominence (i.e. lexical stress)</td>
<td>(iii) pronominal</td>
</tr>
<tr>
<td>(iv) extra prominence</td>
<td>(iv) default sentence stress</td>
</tr>
<tr>
<td><em>(v)</em> extra-extra prominence</td>
<td><em>(v)</em> focal stress</td>
</tr>
</tbody>
</table>

e.g. In English,  
(i) ---  
(ii) pronominal  
(iii) default sentence stress  
(iv) focal stress  
(v) ---

e.g. In Japanese,  
(i) null NP  
(ii) ---  
(iii) pronominal  
(iv) default sentence stress  
(v) focal stress

The scale in (10) specifies that anaphoric elements can be: (i) deleted, (ii) deaccented, or (iii) realized with normal (lexical) prosodic prominence; and focused/disanaphoric elements can carry: (iii) normal (lexical) prosodic prominence, or (iv) extra prosodic prominence. The star "*" sign in front of an extra higher peak in (v) indicates that an extra boosting of a pitch peak is avoided cross-linguistically (cf. 4.2.1).

The way English and Japanese encode disanaphoricity in narrow-focus sentences is summarized...
in the table in (11) below. In English, deletion of anaphoric arguments (= (11i)) is not productive. Instead, English uses deaccenting (= (11ii)) to express anaphoric elements, and anaphoric elements are often expressed by deaccented pronominals. In contrast, focused elements carry focal stress (= (11iv)).

In contrast to English, Japanese uses deletion productively for anaphoric arguments (= (11i)) when their informational content is recoverable from discourse context. When anaphoric elements are deleted, the sentence ends up consisting only of focused elements. When anaphoric elements cannot be deleted for pragmatic reasons, they can be retained as lexical words or pronominals. The difference between Japanese and English, however, is that in Japanese, pronominals are not deaccented even when anaphoric, and carry lexical stress (= (11iii)). When anaphoric elements are retained as pronominals, focused elements must carry extra prosodic prominence. In that case, Japanese places a narrow-focus into the default sentence stress position, because the default sentence stress in Japanese is prosodically distinct and realized with extra prosodic prominence (= (11iv)) independently of focus. If focal stress with an extra higher pitch peak is employed instead (= (11v)), the elements following the focal stress are all deaccented, as formulated by the Prosodic Condition on Japanese [Focus]0 we saw in (8) above.

(11) Prosodic encoding of a narrow-focus in English and Japanese

<table>
<thead>
<tr>
<th>Prosodic processes in (10)</th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
<th>*(v)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Narrow-focus</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td>Narrow-focus w/deletion</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Narrow-focus w/ alternation</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td></td>
<td>#/*Narrow-focus w/o alternation</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

Remember that employment of focal stress is the least preferred option (if possible at all) in Japanese. I have formulated this effect as the Prosodic Economy Condition which triggers word order alternations in Japanese:

(12) The Prosodic Economy Condition for Japanese
Japanese employs default sentence stress rather than focal stress whenever possible, in order to prosodically encode disanaphoricity.

This suggests a general cross-linguistic prohibition of extra-extra prominence (= (11v)).

To sum up, English uses different degrees of prosodic prominence, e.g. deaccenting (= (11ii)), normal prominence (= (11iii)), and extra prominence (= (11iv)), in order to set up a contrast between anaphoric and focused elements. On the other hand, Japanese uses deletion (= (11i)) or word order alternations in order to use default sentence stress, which is prosodically distinct independently of focus (= (11iv)).

5.1.1.2. Generalizations about Linearization of Verbal Dependents

Differences in the prosodic encoding of disanaphoricity lead to differences in linearization properties between English and Japanese. The generalizations about linearization of verbal dependents in English and Japanese are shown in (13) and (14), respectively:

2 Although deletion of anaphoric arguments is restricted to specific registers (e.g. cooking instruction) in English, English deletes anaphoric VPs and IPs (e.g. do-φ) productively (see e.g. Merchant 2001). I do not have explanation for what restricts deletion of anaphoric arguments in particular in English.
Linearization in the verbal domain in English

Let the C-order = V – DP – PP with the prosodic pattern [Intp ... V – DP – PP], and the M-order = V – PP – DP with the prosodic pattern [Intp ... V – PP] [Intp DP],

a. select M-order or C-order when M-order satisfies the prosodic weight condition, and
b. select C-order elsewhere.

Linearization in the verbal domain in Japanese

Let the Order-I = PP – NP – V with the prosodic pattern [Intp PP] [Intp NP – V], and the Order-II = NP – PP – V with the prosodic pattern [Intp NP] [Intp PP – V],

a. select Order-I or Order-II if there is no narrow-focus.

b. select Order-I is NP is (a part of) a narrow-focus, and
c. select Order-II if PP is (a part of) a narrow-focus.

The heavy NP shift alternation in English is independent of the prosodic encoding of disanaphoricity. Instead, it is motivated by the Prosodic Weight Condition, which requires a prosodically heavy phrase to be in sentence-final position.

The Prosodic Weight Condition
POn (= the rightmost Intonational Phrase) must be prosodically heavy by containing

a. a larger number of prosodic words than POni (= the second rightmost
   Intonational Phrase), or
b. a prosodic word with extra prosodic prominence.

In English, the Prosodic Weight Condition is empirically motivated by the observation that the two orders, the canonical V – DP – PP order (C-order) and the marked V – PP – DP order (M-order), exhibit different intonation patterns. C-order exhibits the default intonation pattern where the whole sentence (IP) corresponds to an IntP (cf. (4i)), whereas in M-order, the rightmost DP must form its own Intonational Phrase (IntP). M-order is available only if it satisfies the Prosodic Weight Condition in (15).

Thus, M-order is prosodically more marked than C-order. The SO-PO Mapping Condition on English PPs in (16) formulates prosodic markedness of the marked V – PP – DP order in English.

The SO-PO Mapping Condition on English PPs
Align (SO, R; PO, R), where SO = a PP dependent, and PO = Intonational Phrase.

Although the Prosodic Weight Condition in (15) is irrelevant to word order alternations in Japanese, we can assume that the condition holds cross-linguistically because it is vacuously satisfied in Japanese sentences. In Japanese, the left edge of a syntactic phrase is aligned with the left edge of an IntP (cf. (4ii)), regardless of the linear ordering between dependents. Therefore, the rightmost IntP is heavier than the preceding IntP automatically by virtue of containing a verb. Even if a preceding syntactic phrase (i.e. PP in Order-I and NP in Order-II) appears to be “heavy” by virtue of containing a modifier, the modifier forms an IntP on its own and hence does not lead to a violation of the Prosodic Weight Condition. In Japanese, the linear ordering between dependents is determined by the locus of a narrow-focus, as was formulated by the Prosodic Economy Condition in (12), and a narrow-focus is placed into the default sentence stress position.

To summarize, in the default case, the whole sentence (IP) corresponds to an IntP in English by the SO-PO mapping constraint in (4i). When a DP dependent appears in W-final position as in M-order, an intonational break is inserted after the other PP dependent and before the DP, as is formulated by the SO-PO Mapping Condition on English PPs in (16). In contrast, word order does not affect the intonation pattern in Japanese, and the left edge of a syntactic phrase (XP) is aligned with the left edge of an IntP by the SO-PO mapping constraint in (4ii) in both Order-I and Order-II.

Thus, English word order alternations are associated with a prosodic markedness effect, whereas Japanese word order alternations are not. In English, the prosodic property of dependents directly
determines word order (see (13)), whereas in Japanese, the locus of a narrow-focus determines word order (see (14)). The driving force for Japanese word order alternations is the maximal use of the default sentence stress position. In this sense, Japanese word order alternations are also prosodically conditioned.

The generalizations in (13)-(14) have gone unnoticed under purely syntax-oriented approaches to linearization, which regard English C-order as analogous to Japanese Order-I, and English M-order as analogous to Japanese Order-II. The prosodically based approach to linearization correctly predicts the presence of prosodic markedness in English, and the absence of prosodic markedness in Japanese. The next section recapitulates the theoretical significance of the Prosodic Phase Hypothesis.

5.1.2. Theoretical Significance of the Prosodic Phase Hypothesis

5.1.2.1. Determining Linearization

The central idea behind the Prosodic Phase Hypothesis (repeated below) is that a certain type of linearization in natural languages is prosodically constrained.

(1) The Prosodic Phase Hypothesis
A syntactic object is spelled out as a prosodic object.

My case study of English and Japanese word order alternations in the verbal domain provides empirical support for the adequacy of the Prosodic Phase Hypothesis. In particular, the prosodically based approach to these alternations derives the existence of prosodic markedness in English and its absence in Japanese.

In order to prosodically encode disanaphoricity, English uses different degrees of prosodic prominence, dividing a sentence into smaller Intonational Phrases (IntPs). In contrast, Japanese deletes anaphoric elements when possible, or uses word order alternations to place a focused element into the default sentence stress position. In English, the availability of mobile prosodic prominence limits the extent of word order alternations. In Japanese, the availability of deletion makes the word order alternation vacuous. When the deletion strategy is not available for pragmatic reasons, word order alternation is employed more freely in Japanese than in English, due to the unavailability of mobile prosodic prominence in Japanese.

The Prosodic Phase Hypothesis predicts not only markedness and relative rigidity of word order, but also a cluster of prosodic and syntactic properties associated with linearization of verbal dependents (cf. 3.3 and 4.3).

5.1.2.2. A Division of Labor between Syntax and Prosody

One of the crucial hypotheses the Prosodic Phase Hypothesis uses is incremental structure-building in the syntax (Phillips 1996, 2003). The Prosodic Phase Hypothesis argues that certain types of linearization, in particular, English heavy NP shift and Japanese short-scrambling, are constrained by the syntax-prosody mapping. The Prosodic Phase Hypothesis, by defining a Phase as a pairing operation of syntactic objects (SOs) and prosodic objects (POs), explains the characteristic properties of heavy NP shift and short-scrambling, notably sensitivity to prosodic prominence, which have been problematic for purely syntactic approaches to these alternations.

Defining a Phase as SO-PO pairing imposes a division of labor between syntax and prosody, by shifting an explanatory basis of prosodically conditioned linearization from the syntax proper to the syntax-prosody interface. However, this does not mean that the prosodically based definition of a Phase completely replaces the syntax-based definition of a Phase with respect to the empirical domain of explanation: whether or not the empirical phenomena that have been explained in terms of syntactic Phases (e.g. the Strict Cyclicity) can be explained in the context of left-to-right structure-building is an independent issue yet to be investigated.

One of the empirical strengths of a derivational approach to the syntax-prosody mapping is that it
can derive the difference in the size of an intonational domain between English and Japanese from the difference in the size of a syntactic category, without assigning different prosodic categories to them. It follows from the SO-PO mapping constraints for IntPs in (4) that a clause is the maximal SO that can be mapped to an IntP in English, whereas a phrase is the maximal SO that can be mapped to an IntP in Japanese. This way, we do not have to stipulate that the IntP in English is a higher level category than the IntP in Japanese in the prosodic hierarchy.

5.1.2.3. The Role of Intonational Phrases in Defining Phonological Processes

An Intonational Phrase (IntP) plays an essential role in implementing the Prosodic Phase Hypothesis. This is reflected in phonological processes such as the Default Sentence Stress Assignment and deaccenting. For example, cross-linguistically, default sentence stress is always found in the rightmost IntP, whose size is determined by general prosodic properties of a particular language, and hence does not have to be stipulated.

(6) Default Sentence Stress Assignment
Default sentence stress is assigned to a prosodic word in the rightmost Intonational Phrase (= PO

| Edge          | Domain                      | PO

| i. Right (i.e. rightmost prosodic word) | English, Italian, Spanish, French, Bengali etc. |
| i. Left (i.e. leftmost prosodic word)  | Japanese, Korean etc. |

Although the exact locus of default sentence stress within the rightmost IntP (= PO

Independently of the edge parameter, the Default Sentence Stress Assignment in (6) refers to the rightmost IntP. I take this “rightmost-ness” of default sentence stress as a universal PF property imposed by the sensori-motor system, but how it is so remains to be investigated.

5.2. Implications and Further Issues: The Lexical Pitch Parameter

The gist of the Prosodic Phase Hypothesis is that the size of an Intonational Phrase (IntP) in a particular language constrains the linearization of verbal dependents in that language. For example, in English a clause (CP/IP) corresponds to an IntP, and in Japanese a syntactic phrase (XP) corresponds to an IntP. Therefore English IntPs are bigger than Japanese IntPs. The SO-PO mapping constraints for maximal prosodic objects in (4) specify the difference in the size of IntP between the two languages, and hence different prosodic categories do not have to be postulated for English and Japanese IntPs. However, it remains to be explained why it is CP/IP in English and XP in Japanese that is paired with an IntP. In this section, I attempt to explain the differences between English and Japanese in terms of one prosodic parameter: whether lexically specified pitch features in a language are distinctive (e.g. Japanese) or not (e.g. English). I call this the “Lexical Pitch Parameter”. Let me quote Pierrehumbert and Beckman (1988), who compare intonation systems in English and Japanese concisely:

169
In Japanese pitch accents are specified in the lexicon. Phrasal tones are assigned
in a determinate way on the basis of prosodic structure. In English, by contrast,
pitch accents are not part of the lexical specification of words. Instead, they are
elements in an inventory of intonational morphemes, which also includes the H or L
phrase accent and the H% or L% boundary tone. (Pierrehumbert and Beckman 1988:237)

Therefore, English and Japanese are distinguished according to the level at
which pitch is relevant to meaning: English is an “intonation language” and
Japanese, a “word-pitch language” (Pike 1948).

Correlated with this difference in size is a difference in the degree of
variability associated with intonational phrasing, which is much greater in
English than in Japanese. I propose that this difference also derives from
the Lexical Pitch Parameter: Japanese is a word-pitch language, and the lexical
origin of the pitch features makes “the range of intonational variation considerably
smaller than in the English intonation system” (Beckman and Pierrehumbert
1986:306). In other words, English prosody enjoys more mobility than
Japanese prosody. The table in (17) summarizes prosodic properties related to
the distinctiveness of lexically specified pitch features.

(17) The Lexical Pitch Parameter and related prosodic properties

<table>
<thead>
<tr>
<th>Prosodic properties</th>
<th>Language-type</th>
<th>Intonation language (English)</th>
<th>Word-pitch language (Japanese)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical Pitch Parameter</td>
<td></td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>(lexical pitch as a distinctive feature)</td>
<td></td>
<td>CP/IP</td>
<td>XP</td>
</tr>
<tr>
<td>i. Size of Intonational Phrase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ii. Mobility of prominence/ variability of intonational phrasing</td>
<td></td>
<td>+</td>
<td>--</td>
</tr>
</tbody>
</table>

The idea behind the Lexical Pitch Parameter is that the way English uses
prosodic prominence is relational and hence holistic, whereas the way
Japanese uses prosodic prominence is absolute and hence local (cf. Abe 1955).
Recall that default sentence stress (i.e. the rightmost lexical stress in a sentence)
is not necessarily prosodically distinct in English, whereas it is prosodically
distinct in Japanese. In English, default sentence stress can be perceived as
relatively prominent if there is a contrast between anaphoric and focused elements
within a sentence, as in partial broad-focus sentences. In contrast, default sentence
stress exhibits a higher peak than the preceding stress in Japanese, not only in
partial broad-focus sentences, but also in broad-focus sentences. This leads to
a systematic contrast between the two languages in the prosodic processes used to encode
disanaphoricity (as indicated by the bi-directional arrows in (10) below):

(10) Disanaphoricity Scale

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>anaphoric</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) deletion</td>
<td>(ii) deaccenting</td>
<td>(iii) normal prominence (i.e. lexical stress)</td>
<td>(iv) extra prominence</td>
<td>*(v) extra-extra prominence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>focused/disanaphoric</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The differences in the value of the Lexical Pitch Parameter in English and Japanese seem to be correlated with another set of differences in the prosodic encoding of disanaphoricity, which in turn may determine how linearization in a language is prosodically constrained. The non-distinctiveness of lexical pitch in English allows mobile prosodic prominence, and hence English uses a contrast between deaccenting (= (10ii)) and normal prominence (= (10iii)). In contrast, distinctive lexical pitch in Japanese does not allow mobile prominence, particularly (partial) deaccenting, and Japanese uses a higher level contrast between normal prominence (= (10iii)) and extra prominence (= (10iv)). In other words, English deaccents anaphoric elements, and Japanese assigns extra prominence to a focused element by placing it into the default sentence stress position, which is prosodically distinct (= (10iv)). Therefore, the prosodic encoding of disanaphoricity does not determine linearization in English, whereas it does in Japanese.

Thus, the Lexical Pitch Parameter divides languages into at least two groups, each of which is identified with a particular set of prosodic properties. According to Ladd (1996:196), German and Dutch belong to the English type (which Ladd calls “deaccenting” languages), and Korean belongs to the Japanese type (which Ladd calls “dephrasing” languages). Regarding word order alternations, the prediction is that intonation languages should show a prosodic markedness effect, where one order is prosodically more marked than the other and has to satisfy the Prosodic Weight Condition in (15), whereas such markedness should be absent in word-pitch languages.

German and Dutch seem to bear out the first half of this prediction in the following sense. First of all, they share the value of the Lexical Pitch Parameter with English in that their lexical pitch is non-distinctive and they use variable degrees of prosodic prominence (Bolinger 1989:42-43, Ladd 1996, Gibbon 1998:90-91). Secondly, German and Dutch exhibit alternations between dependents not only in the pre-verbal domain but also in the post-verbal domain, due to their less rigid verb-finality, as opposed to the rigid verb-finality of Japanese (Greenberg 1966). This makes it possible to look into whether or not German and Dutch exhibit the prosodic weight effect at the right edge of a sentence as English does. Previous studies of word order alternations in German and Dutch reveal that: (i) in the post-verbal domain, prosodic markedness exists in rightward XP-shift alternations, and the marked order exhibits the right-edge heaviness effect (see e.g. Hawkins 1986, Truckenbrodt 1994, Cohan et al. 2001, Shiobara to appear, among others); and (ii) in the pre-verbal domain, mobile prosodic prominence (e.g. deaccenting and accenting) is possible and hence the Prosodic Economy Condition in (12) does not trigger word order alternations (see e.g. Cinque 1993, Reinhart 1996, Krifka 1998, Neeleman and Reinhart 1998, De Hoop 2003 among others). Therefore, as far as English-type languages are concerned, the attempt to reduce linearization properties to the Lexical Pitch Parameter seems to be a promising approach.3

5.3. Final Remarks

This thesis provided a prosodically based explanation for certain word order alternations in English and Japanese which have been problematic for purely syntax-oriented approaches to linearization (in particular, heavy NP shift in English and short-scrambling in Japanese). The Prosodic Phase Hypothesis

---

3 Tone languages should share the value of the Lexical Pitch Parameter with word-pitch languages, because their lexically specified pitch features are distinctive. Other things being equal, tone languages should behave the same way as word-pitch languages with regards to linearization. However, this is not obvious, since lexical tones may be modified by intonation (Shen 1990). In addition, it seems often to be the case that tone languages have a syntactically designated focus construction as in Yoruba (Ajiboye p.c. 2004) or a focus position as in Chinese (Xu 2004). I leave further typological investigation for future research.
claims that the different behavior of English and Japanese word order alternations is due to differences in the general prosodic properties of the two languages. English makes use of mobile prosodic prominence (e.g. accenting and deaccenting) in order to encode disanaphoricity in a sentence, whereas Japanese does not. In the previous section, I have attempted to reduce this difference to the Lexical Pitch Parameter, which differentiates languages according to whether their lexically specified pitch features are distinctive (Japanese) or not (English).

There has been a tendency within purely syntax-oriented approaches to linearization to relegate prosodically conditioned word order alternations such as heavy NP shift in English and short-scrambling in Japanese to the “marked periphery”. However, I have shown that systematic linearization possibilities are found when we examine prosody, which is in part determined by semantic/pragmatic factors such as focus.

Generative syntax has matured sufficiently in the study of “core” grammar, and enough research has been done on a broad range of languages that the limits of syntactic explanation are beginning to be evident. In the introductory chapter of Syntax and Semantics volume 29 in 1998 with the subtitle “the limits of syntax”, Culicover and McNally provide a concise description of the situation:

> Over the past 35 years, generative grammarians have made a great deal of progress in understanding the nature of human language by characterizing a very wide range of linguistic phenomena in syntactic terms. This inquiry has focused on developing theories that illuminate speaker competence and has led to many interesting and successful predictions about cross-linguistic similarities and differences. However, as this research has grown to incorporate a greater variety of within-language and cross-linguistic data, it has become increasingly clear that certain phenomena that have been given widely accepted syntactic explanations [...] quite possibly cannot be analyzed in exclusively syntactic terms without an undesirable impact on the explanatory value of the syntactic theories themselves. (Culicover and McNally 1998:1)

The time is right for exploring the question of how the incorporation of nonsyntactic explanations into syntactic research may deepen our understanding of problematic linguistic phenomena. Among all nonsyntactic areas of research, Culicover and McNally (1998) note that “in the last 10 years there has been considerable research on semantic and formal pragmatic approaches to classic linguistic problems such as anaphora and focus” (p.2). This research trend also has consequence for our understanding of prosody, because semantic/pragmatic properties of focus, for example, are closely tied to their prosodic realization, as we have seen throughout the thesis.

In this thesis, I have attempted to pursue a parallel research strategy on the prosodic side of the grammar. My major hypothesis has been that systematic investigation of the syntax-prosody interface can account for certain properties which are problematic for purely syntactic explanation. In particular, sensitivity to prosodic factors such as weight and sentence level stress is diagnostic of a class of word order alternations which are best dealt with by correspondence conditions on the mapping from syntax to prosody. By removing these alternations from “core” syntax, I have simultaneously simplified the syntax itself, and provided an explanatory account of phenomena previously relegated to the “periphery”.

The development of research in the areas of semantics, formal pragmatics, and prosody is also accelerated by the guiding idea of the Minimalist Program, which seeks to derive all conditions imposed on the grammar from the interfaces (Chomsky 1995, 1998, 1999, 2001). The Prosodic Phase Hypothesis, which imposes a particular division of labor between the syntax and the prosody, shifts the explanatory basis of certain types of linearization from pure syntax toward the PF interface. In this sense, it clearly follows the spirit of the Minimalist Program.

A particular implementation of the Prosodic Phase Hypothesis employed in this thesis is based on two crucial hypotheses. Firstly, syntactic structure is incrementally built in the order in which syntactic words are merged, which is independently motivated for explaining a wide range of grammatical and performance phenomena (Phillips 1996, 2003). Secondly, the computational component includes a
derivational mapping from the syntax to the prosody, where a Phase is defined as derivational pairing of syntactic objects and prosodic objects. Therefore, the Prosodic Phase Hypothesis, which is based on a close analysis of prosodically conditioned linearization, provides supporting evidence for incremental structure-building in the syntax and a derivational approach to the syntax-prosody mapping. It is hoped that the Prosodic Phase Hypothesis will contribute to our further understanding of the architecture of the grammar.
REFERENCES


Massachusetts, Amherst. [Published in 1985 by New York: Garland.]
Rochemont, Michael and Peter W. Culicover. 1990. English Focus Constructions and the 
British Columbia.
Doctoral thesis, MIT.
1: 69-118.
Cambridge, MA: MIT Press.
Interdisciplinary Approaches to Language Essays in Honor of S.-Y. Kuroda, ed. by C. 
Research: University of Tokyo Working Papers in English Linguistics 18: 61-96.
Shiobara, Kayono. 2002a. “On the interaction between end-weight and end-focus,” Paper 
presented at the annual meeting of Canadian Linguistic Association, University of 
Toronto. [Published in 2004. Université du Québec à Montréal Working Papers, 273-284]
Poster presented at NELS 33, MIT.
The Proceedings of the Fourth Tokyo Conference on Psycholinguistics, ed. by Y. Otsu, 
UBC Working Papers in Linguistics 11.
Sugisaki, Koji and Miwa Isobe. 2001. “What can child Japanese tell us about the syntax of 
scrambling?”, The Proceedings of West Coast Conferences of Formal Linguistics 20: