Linearizing Structure with Silence:
A Minimalist Theory of Syntax-Phonology Interface

A Dissertation
Submitted to the University of Tsukuba
In Partial Fulfillment of the Requirements for
the Degree of
Doctor of Philosophy in Linguistics

Hisao TOKIZAKI

2006

Linearizing Structure with Silence: A Minimalist Theory of Syntax-Phonology Interface

by

Hisao TOKIZAKI

Submitted to the University of Tsukuba

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Linguistics

December 2006

ABSTRACT

This thesis investigates how phrase structure of sentences is mapped onto phonological representations. The bare mapping theory is proposed which interprets syntactic boundaries as phonological boundaries. Prosodic phrases are formed by deleting a number of boundaries according to the level of phrase and the rate of speech. This theory supports the idea of bare phrase structure rather than X-bar theoretic phrase structure. The theory of cyclic Spell-Out enables us to do away with the readjustment rule. The effect of edge parameter is derived by syntactic head parameter. Optionality of phrasing is also explained by the deletion of a number of boundaries. Further consequences of the theory are discussed which include the effects of constituent length, i.e. secondary phrasal stress and Heavy NP Shift in English and optional phrasing in Korean and Japanese. The theory offers an alternative analysis to the Early Immediate Constituent analysis (Hawkins 1994) and help us to explore the relation between phrase structure and sentence processing.

Prosody and punctuation in English and Japanese, topic/focus and phrasing, semantics and phrasing, and derivation and parsing are also discussed.

Acknowledgments

Until quite recently I did not know if the day really comes when I can finish my dissertation and write the acknowledgment. Now I am sure that a lot of people around me made it possible for me to do it. First I would like to thank my thesis committee: Nobuhiro Kaga, who gave me a chance to submit my dissertation to University of Tsukuba and has been encouraging me to finish it, Masao Okazaki, Hiromi Onozuka, Masaharu Shimada, and Norio Yamada, who gave me valuable comments and suggestions. I am grateful to Yukio Hirose, whose thesis interested me when I was a graduate student.

I would like to thank Seizo Kasai, who has been my advisor at Hokkaido University for more than twenty years. His love of languages as well as his deep insights into them awoke my interest in linguistics. The people of Sapporo Linguistics Circle gave me wonderful opportunity of presenting and discussing my ideas. I am grateful to Masanobu Ueda, Katsuhide Sonoda, Kimihiro Ohno, Haruhiko Ono, Yoshihiro Yamada, Seiji Ueno, Hiroyuki Kageyama, and Satoshi Oku. I would like to thank the people in the linguistic department at Hokkaido University: Jiro Ikegami, Toshiro Tsumagari, Fubito Endo, Osamu Kawagoe, and Yuko Kawauchi. The syntax reading group was the starting point of my career as a linguist.

The year I spent at Amherst, Massachusetts widened my research area. There I met a number of wonderful linguists. I would like to thank Barbara Partee, Elisabeth Selkirk, Lyn Frazier, Norvin Richards, Junko Shimoyama, Kiyomi Kusumoto, Mariko Sugahara, Masako Hirotani, Bozena Cetnarowska, Anthi Revithiadou, Kim Hak-Yoon, Hainer Drenhouse, and Takahiro Tsutsui. During my stay in Amherst, I often visited Boston, where I met Inada Toshiaki, Tsuyoshi Oishi, Minoru Amanuma, and Akihiko Uechi. I would like to thank Susumu Kuno for giving me a chance to talk at Harvard University.

I also met wonderful linguists at academic meetings and summer institutes both in Japan and in the states. I would like to thank Naoko Hayase, Masatoshi Koizumi, Shin-ich Tanaka, Takeru Homma, Kayono Shiobara, and Yoshihito Dobashi.

I am also grateful to Shosuke Haraguchi, Takeshi Kohno, Koichi Tateishi, and Haruo Kubozono for their valuable comments and suggestions. I would also like to thank Hideyuki Hirano who encouraged me with his kind words and warm smile.

William Jones and William Green kindly corrected my stylistic errors. I would like to thank Heiko Narrog for his help with German.

People at Sapporo University kindly helped me with discussion and gave me a lot of support. I am grateful to Hideto Hamada, Yoshihisa Goto, and Yasutomo Kuwana.

Finally, I would like to thank my family, especially my parents, Hiroshi and Kinuko Tokizaki for their love and support. This thesis is dedicated to them.

Abbreviations

The following is a list of abbreviations used in the glosses in this thesis.

Acc	accusative
Cl	classifier
Def	definite
E	Xiamen *e* (cf. Chen 1987: 145, n. 11)
Emph	emphatic
Fut	future
Neg	negative
Nml	nominalizer
Nom	nominative
Obj	objective
Pl	plural
PP	past participle
Prog	progressive
Q	question marker
Prt	particle
Rel	relative
Sg	singular
Top	topic marker
1	first person
3	third person

Contents

Chapter 1

Introduction

The structure of sentences has been one of the general issues in linguistics. In the theory of generative grammar, it has been assumed that sentences have hierarchical structure that is schematized with tree diagrams. It is poorly understood, however, how phrase structure is interpreted in the phonological component of the grammar. In other words, the exact nature of Spell-Out has not been well discussed. The goal of this thesis is to take a closer view of the relation between syntax and phonology.

In this thesis I will propose a theory of syntax-phonology mapping in the minimalist framework (Chomsky 1995). Chapter 1 is an overview of the mapping theories proposed so far. I will illustrate the idea of phrase structure in the minimalist framework and will argue that those interface theories are not tenable in the current theory of grammar. In Chapter 2, I will propose a bare theory of syntax-phonology interface which is based on the idea of bare phrase structure. I will argue that the bare interface theory supports the bare phrase structure theory rather than the standard X-bar theory. I will show various prosodic phenomena as supporting evidence. In Chapter 3, I will argue that the bare mapping theory derives the Edge Parameter (Selkirk and Tateishi 1988, 1991) from the syntactic head parameter. Chapter 4 is a discussion of optional phrasing phenomena in several languages. In Chapter 5, I will discuss the effect of the length of constituents on phonology and syntax. In Chapter 6, I will extend the analysis to discourse. In Chapter 7, I will investigate the relation between topic/focus and phrasing and explain when topic and focus sometimes make a separate prosodic phrase. Chapter 8 is a discussion of semantic effects on phrasing. In Chapter 9, I will argue that lexical items and syntactic brackets are Spelled Out and parsed one by one. Chapter 10 is devoted to Conclusion.

1.1 Architecture of Grammar: Components and Their Interface

First I would like to show the framework in which the thesis is written. A recent theory of generative grammar, called minimalist program, assumes syntactic derivation and two interface levels with sound and meaning (Chomsky 1995, 1998, 2000). Lexical items are introduced into derivation by the operation Merge, which combines two syntactic objects. For example, a noun *cats* is merged with a verb *loves* and makes a verb phrase *[$_{VP}$ loves cats]*. Then the VP is merged with another lexical item or phrase, and so on. Constituents made by Merge are interpreted as sound by Spell-Out at some points of derivation (phases), and are sent to the phonological component (Phonetic Form: PF).[1] At the same time, meanings are also interpreted and are sent to the semantic component (Logical Form: LF). The following diagram sketches the overall architecture of grammar proposed by Chomsky (1999):

(1)

$$\begin{tikzcd} & \text{LF1} & \text{LF2} & \\ \text{Numeration} \arrow[r] & \text{Phase1} \arrow[u] \arrow[d] \arrow[r] & \text{Phase2} \arrow[u] \arrow[d] \arrow[r] & \ldots \\ & \text{PF1} & \text{PF2} & \end{tikzcd}$$

Note that the interpretive processes are cyclic and are called cyclic Spell-Out. This is a characteristic point of the recent development in minimalist theory of grammar. Given this model of grammar, the goals of the research in syntax-phonology interface is to determine how syntactic structure is interpreted as phonological representation and what information in syntax is mapped onto phonology in what way.

---

[1] I will not go into detail of phase here. Following Chomsky (1999), I assume that CP and *v*P are (strong) phases.

In Chapter 9, I will consider an alternative model of derivation and parsing, in which each syntactic object is Spelled-Out and parsed incrementally.

1.2 Previous Proposals: Overview of Their Differences

There are a number of proposals for the analysis of syntax-phonology interface. I would like to briefly review some of them which are relevant to the discussion below: (i) relation-based mapping (e.g., Nespor and Vogel 1982, 1986, Hayes 1989); (ii) end-based mapping (e.g., Chen 1987, Selkirk 1986, Selkirk and Shen 1990); (iii) arboreal mapping (Zec and Inkelas 1990).[2]

1.2.1 Relation-Based Mapping

Nespor and Vogel (1986:168) propose the principles for the definition of phonological phrases (ϕ) as in (2).

(2) The domain of ϕ consists of a C which contains a lexical head (X) and all Cs on its nonrecursive side up to the C that contains another head outside of the maximal projection on X.

To illustrate how (2) works, let us look at an example from Italian.

[2] See Inkelas and Zec (1995) for another review of these three proposals. Note that some research has been done which is based on the optimality theory (e.g., Truckenbrodt 1995, 1999, Selkirk 2000). Other than these, various proposals are published occasionally (Jackendoff 1987, Steedman 2000).

(3) $[[\text{Aveva}]_C\ [\text{giá}]_C\ [_V\ \text{visto}]_C\ ]_\phi\ [[\text{molti}]_C\ [_N\ \text{canguri}]_C\ ]_\phi$

'He had already seen many kangaroos.'

As Italian is syntactically right branching, the recursive side with respect to the head is the right side. The verb and the noun, which are lexical heads, incorporate the words to their left into their phonological phrases as in (3).

1.2.2 End-Based Mapping

Selkirk (1986) argues that phonological phrasing can be predicted by the end-based theory, which can be summarized as in the following algorithm:[3]

(4) a. $_{Xmax}$ [...

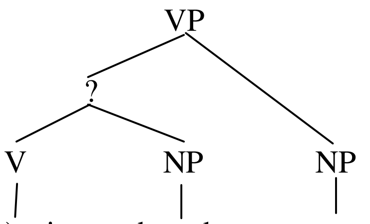
b. ...] $_{Xmax}$

The phrasing position is parameterized so that a language chooses the left (4a) or right (4b) end of a maximal projection as a phrasing boundary. Selkirk (1986:382) gives an example from Chi Mwi:ni, which chooses the right end setting (4b).

---

[3] Cho (1990) also discusses the relation-based theory (Nespor and Vogel 1986, among others) and the direct syntax approach (Kaisse 1985), the latter of which I will not discuss here.

(5)



a. pa(:)nzize cho:mbo mwa:mba

'he ran the vessel on to the rock'

b. ....$]_{Xmax}$ ....$]_{Xmax}$

c. ($)_{PPh}$ ($)_{PPh}$

If we apply (4b) to the sentence (5a), we get the correct phrasing (5c).

1.2.3 Arboreal Mapping

Zec and Inkelas (1990:370) show an algorithm of phonological phrasing shown in (6).

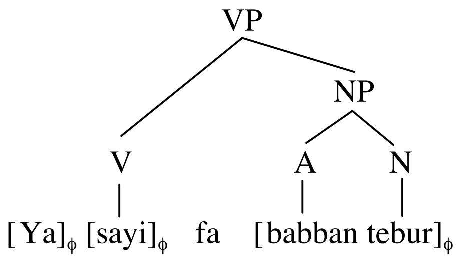
(6) a. Prominent elements are mapped into their own phonological phrases.

b. From the bottom up, branching nodes are mapped into phonological phrases.

c. No two phonological words on opposite sides of an XP boundary may be phrased together to the exclusion of any material in either XP.

Zec and Inkelas illustrate the result of applying this algorithm to some example sentences in Hausa.

(7)



he bought EMP big table

'He bought a big table.'

The adjective and the noun are syntactic sisters and are grouped into a phonological phrase. Since no nesting of phonological phrases is permitted, the verb is forced to phrase separately in the verb phrase.

1.2.4 Similarities and Differences among the Interface Theories

It is almost impossible to critically review all of the approaches to syntax-phonology interface proposed so far. Instead I would like to recapitulate their similarities and differences. The first point of difference is the input to the mapping. Most of the theories assume that the input to the mapping is the S-structure of sentences or the phrase structure at the Spell-Out in the minimalist framework. Steedman (2000) is an exception in that he assumes the derived constituent structure in Combinatory Categorial Grammar as the input to the mapping. Even in the theories that are based on generative grammar, the phrase structures they assume are somewhat different from each other. This is mainly due to the rapid development of the generative syntax. I will argue that bare phrase structure should be the input to the phonological component in Chapter 2.

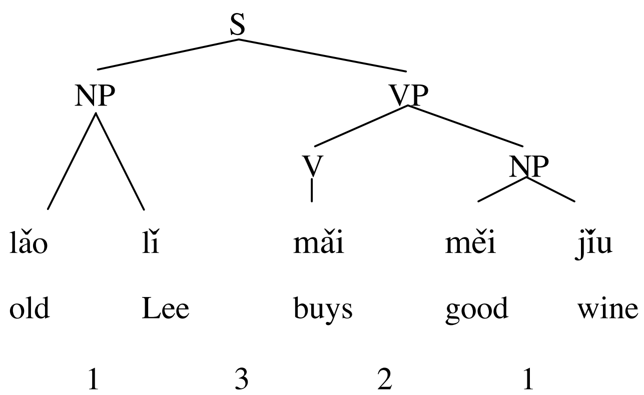
The second point of difference among theories is what counts as the crucial factor in prosodic phrasing. The head-complement relation is crucial in the relation-based theory. The end-based theory treats the left or right edge of X^n as the boundary in phrasing. The arboreal mapping makes a prosodic phrase by grouping sister constituents. The optimality approach claims that there are a number of factors involved in phrasing, such as Align XP and Wrap XP. In this thesis I would like to develop a theory in which the whole phrase structure of sentence is crucial in mapping from syntax into phonology.

1.3 Previous ideas of syntactic depth

1.3.1 Depth of Syntactic Boundaries

Before starting with the new mapping theory, let us review some previous ideas of syntax-phonology connection that are relevant to the discussion below. First, Cheng, R. (1966:150) refers to the idea of depth of syntactic boundaries proposed by Wang, W. S-Y. (TRIP Report, the Ohio State University. Mimeographed, 1965). Let us look at his example (8).

(8)



The numerals 1-3 approximately indicate closeness of syntactic relationships, which Wang calls depths of syntactic boundaries. The underlying tone of all the words in (8) is the third tone (ˇ). The tone sandhi rules in Mandarin Chinese are (9a) and (9b).

(9) a. ˇ --> ˊ / __ ˇ

b. ˊ --> ˉ / [ˉ] __ [ ] or [ˊ] __ [ ]

(9a) states that a third (dipping) tone changes into a rising tone when it is followed by another third tone. (9b) states that a rising tone changes into a light (level) tone when it is not final and preceded by either a level or rising tone. Interestingly, (9a) does not apply across boundaries of some depth according to the rate of speech. In slow speech, (9a) applies only across boundaries of depth 1 in (8) to produce the tone sequence ˊ ˇ ˇ ˊ ˇ. In faster speech, (9a) applies across boundaries of depth 1 and 2 to give ˊ ˇ ˊ ˉ ˇ. In rapid speech, (9a) applies across all the boundaries to give ˊ ˉ ˉ ˉ ˇ. This tone sandhi phenomenon shows that the notion of syntactic depth plays an important role in syntax-phonology interface.

Wang's idea of depth of syntactic boundaries is interesting and appealing, but it is not entirely clear how to assign numerals to each position between words in longer and more complex sentences than (8).

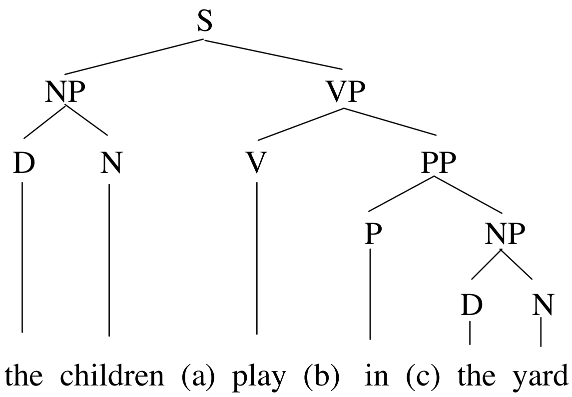
1.3.2 Branching Depth

Another attempt to define the notion 'relative strength of junctures' is Clements' (1978) branching depth.[4] Let us consider the following phrase structure:

---

[4] Clements (1978) in fact proposes five approaches to the syntax-phonology interface: (A) depth of embedding, (B) branching depth, (C) categorial domains, (D) categorial hierarchies, and (E) non categorial hierarchies. I discuss only (B) here which is relevant to the theory of interface to be developed in the later chapters.

(10)



The strength of a juncture is expressed as the total number of categorial nodes it dominates (other than itself) along the two paths connecting it with each of the flanking items. In (10), juncture (a) has a branching depth of 4. The lowest node dominating both *children* and *play* is S, which dominates two categorial nodes other than itself, NP and N, along the path connecting it to *children*, and also two categorial nodes other than itself, VP and V, along the path connecting it to *play*. Similarly, juncture (b) and (c) are assigned a branching depth of 3. This measure describes the intuition that the juncture between the subject and the predicate is the strongest.

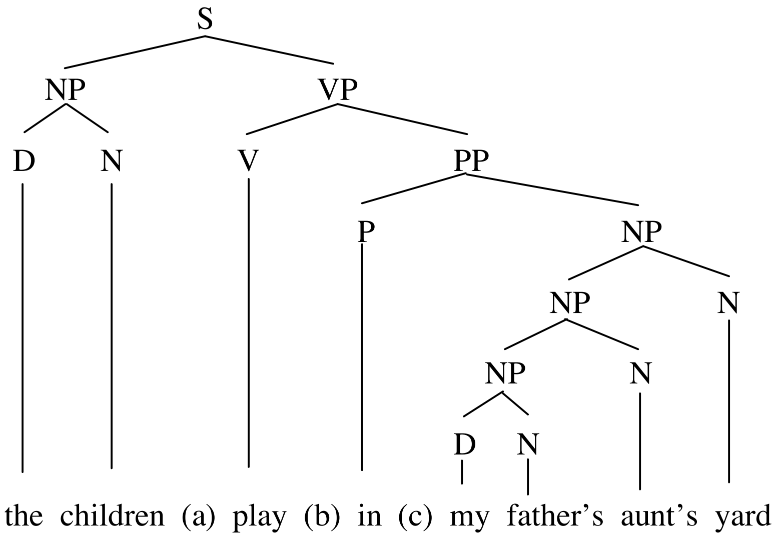
Clements notes that if we express the phrase structure with brackets, branching depth is directly encoded into representations.

(11) $[_{S}\, [_{NP}\, [_{D}\, \text{the}]\, [_{N}\, \text{children}]]\, [_{VP}\, [_{V}\, \text{play}]\, [_{PP}\, [_{P}\, \text{in}]\, [_{NP}\, [_{D}\, \text{the}]\, [_{N}\, \text{yard}]]]]]$

The branching depth of any juncture is identical to the number of brackets intervening between the lexical items that flank them.

Clements also points out a problem with this theory in the case of (12).

(12)



In (12) juncture (c) has the branching depth of 5, and is predicted to be stronger than the corresponding juncture (c) in (10). Clements argues that a given preposition generally shows the same phonological behavior with respect to the following item no matter how deeply embedded this item may be.

I will propose a theory of syntax-phonology mapping in Chapter 2, which contains the same kind of notion as Clements' branching depth.

1.3.3 Silent Demibeat Addition

Lastly, let us review Selkirk's (1984: 314) Silent Demibeat Addition, which articulates the syntactic timing of a sentence.

(13) Silent Demibeat Addition

Add a silent demibeat at the end of the metrical grid aligned with

a. a word

b. a word that is the head of a nonadjunct

c. a phrase

d. a daughter phrase of S.

This rule applies to the sentence (14) to assign the silent demibeats (x) in (15).[5]

(14) $[_S\ [_{NP}\ [_N$ Mary]] $[_{VP}\ [_V$ finished] $[_{NP}$ [her] $[[_{AP}\ [_A$ Russian]] $[_N$ novel]]]]]

```
(15)  x          x               x        x
      x          x           x   x        x
      x    xxx   x x     xx  x   x  x  x  x x  xxxxx
     Mary   ↑   finished  ↑  her Russian↑ novel   ↑
          (a,b,d)      (a,b)           (a)     (a,b,c,d)
```

In (15), *Mary* is followed by three silent positions, because *Mary* is a word (13a), an argument (of VP) (13b), and a daughter of S (13d).[6] Similarly, the other silent demibeats are assigned by (13).

Notice that Silent Demibeat Addition is different from the depth of boundaries and the branching depth in that it counts only the *end* of a constituent as shown in the first line

---

[5] In (15), the stress beats (x) are assigned by other rule than Silent Demibeat Addition (13).

[6] Somehow Selkirk (1984: 317) does not argue that (13c) applies to *Mary* in (15) in spite of the fact that it is assumed to be an NP as shown in (14).

of (13). Furthermore, Selkirk assumes the Principle of Categorial Invisibility of Function Words (PCI), whose effect is to make function words (e.g., determiners, auxiliary verbs, personal pronouns, conjunctions, prepositions, etc) invisible to rules of the grammar. PCI confines (13a) and (13b) of SDA to applying only to words of the categories N, V, A, Adv. These points might be good for describing the data, but at the same time they arouse questions in our mind. Why does SDA count only the end? Why are function words invisible to the rules of grammar? At least we need to know the reasons.

1.3.4 Summary

We have seen three approaches to dealing with syntactic depth or juncture. I have also mentioned some problems they have. Moreover, in the current framework of syntax, we cannot rely on these theories as they are, because these theories of syntax-phonology interface are proposed in 1960s, 70s, and 80s. First, the syntactic structures they assume are different from the ones we assume in the contemporary syntactic theory. Second, they lack functional categories and projections such as I(nfl) and I'. Third, they assume (pre-) X-bar theoretic phrase structure (cf. Chomsky 1986), which assumes three bar levels, XP, X' and X. Thus we cannot rely on these theories as they are. I will propose a new theory of syntax-phonology interface that is compatible with the current theory of syntax.

1.4 Outline of the Theory

1.4.1 The Essentials of the Theory

The theory to be developed below has the following rules and constraints as its essentials:

(16) Syntax-Phonology Mapping (*Linearization*) (Ch.2 (3)):

Interpret boundaries of syntactic constituents [...] as prosodic boundaries / ... /.

(17) Boundary Deletion (*Zoom-Out*) (Ch.2 (5)):

Delete *n* boundaries between words. (*n*: a natural number)

(18) A Constraint on Boundary Deletion (*Consistency*) (Ch.4 (46)):

In a sentence (or paragraph), the number of boundaries to be deleted (*n*) should be as constant as possible.

(19) Avoid Pause (*Continuity*) (Ch.5 p.143):

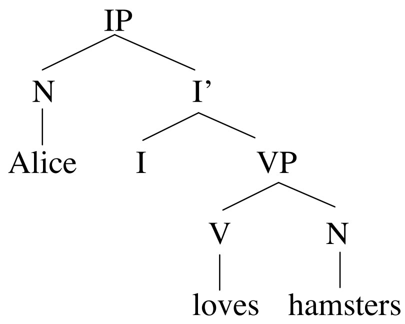
A long pause in a clause should be avoided.

I will show briefly how these rules and constraints apply to sentences.

1.4.2 Application of the Rules and Constrains: Overview

The basic idea of the theory is that hierarchical phrase structure is linearized and represented with brackets enclosing constituents. This is the first step to Spell-Out at the syntax-phonology interface. Consider the following sentence, for example:

(20) a.



b. [$_{IP}$ [$_{N}$ Alice] [$_{I'}$ [$_{VP}$ [$_{V}$ loves] [$_{N}$ hamsters]]]]

The hierarchical structure in (20a) is linearized into (20b) with pairs of brackets. I assume that phonologically null elements and the constituents made by merging them with other syntactic objects are invisible to phonological rules. Then I and I' in (20a) and (20b) are invisible at the syntax-phonology interface. (20b) can be represented as (21).

(21) [$_{IP}$ [$_{N}$ Alice] [$_{VP}$ [$_{V}$ loves] [$_{N}$ hamsters]]]

The brackets and labels such as IP and N are not objects interpretable at the PF interface. They do not have phonetic features of their own. Labels can be eliminated from syntax if we take Collins' (2001) approach (cf. Tokizaki 2005b). Then, (21) can be represented as (22).

(22) [[Alice] [[loves] [hamsters]]]

I also assume that brackets should be transformed into boundaries at the syntax-PF interface and be represented as pause of some length in the PF output. In Chapter 2, I will propose a syntax-phonology mapping rule (16) above, which can also be represented schematically as (23).

(23) $$\left\{ \begin{array}{c} [\\ ] \end{array} \right\} \rightarrow \ /$$

The mapping rule (23) applies to (22) and gives (24).

(24) // Alice /// loves // hamsters ///

The phonological representation in (24) shows the basic juncture of sentence. Words are separated by a number of boundaries between them. I assume that prosodic phrases are also separated by a number of boundaries. To make a larger prosodic phrase, we delete a certain number of boundaries. If we delete one boundary between words, we have (25a). If we delete two, we have (25b). If three boundaries are deleted, we have (25c).

(25) a. / Alice // loves / hamsters //

b. Alice / loves hamsters /

c. Alice loves hamsters

We could represent (25a), (25b), and (25c) as (26a), (26b), and (26c), respectively.

(26) a. (Alice) (loves) (hamsters)

b. (Alice) (loves hamsters)

c. (Alice loves hamsters)

We could argue that each prosodic phrase enclosed by a pair of parentheses in (26a), (26b), and (26c) corresponds to a phonological word, a phonological phrase, and an intonational phrase, respectively. However, since prosodic categories are not without problems, we will try to develop a theory without prosodic category labels in section 3.2.

Boundaries between words can be realized as a certain length of pause. I assume that a boundary is realized as a silent demibeats in the sense of Selkirk (1984) (cf. Liberman 1975:284). We can formulate the realization process as the rule (27).

(27) / → x

If we apply (27) to (25a), (25b) and (25c), we get (28a), (28b), and (28c) as their phonetic representation, respectively.

(28) a. x Alice xx loves x hamsters xx

b. Alice x loves hamsters x

c. Alice loves hamsters

The number of boundaries to be deleted between words can also be related to the speed of utterance. If the basic juncture of sentence in (24), repeated here as (29), is uttered as it is, it is in fact the slowest speech with prosodic boundaries between words.

(29) // Alice /// loves // hamsters ///

The faster the utterance becomes, the longer each prosodic phrase becomes. I assume that the number of brackets to be deleted corresponds to speech rate. When the sentence in (29) is uttered slowly, at most one boundary is deleted between words, as shown in (30a).

When the speech rate is normal, two boundaries between words are deleted, as shown in (30b). In the fastest speech, three boundaries are deleted as shown in (30c).

(30) a. / Alice // loves / hamsters //

b. Alice / loves hamsters /

c. Alice loves hamsters

These patterns may seem to be the same as those in (25), which show the hierarchy of prosodic categories. However, I assume that (30a), (30b), and (30c), utterance at a speech rate, are the input to further application of boundary deletion, which gives various prosodic categories as shown in (26).

Variable boundary deletion also explains variable prosodic phrasing. It is well known that as the speech rate becomes faster, the sentence is divided into fewer prosodic phrases such as phonological phrases (cf. Nespor and Vogel 1986). The paradigm in (30) shows variable prosodic phrasing demarcated by boundaries.

Note that the number of boundaries to be deleted between words should be as constant as possible throughout a sentence or a discourse, as shown above. If it is not constant, the phrased sentence becomes unacceptable, as shown in (31c).

(31) a. [[Two [of [our horses]]] [suddenly [got restive]]]

b. // Two ~~/~~ of ~~/~~ our horses ~~///~~ suddenly / got restive ///

<------------------------ *n*=4 --> <-- *n*=0 ----------->

c. *(Two of our horses suddenly) (got restive)

To make the phrasing in (31c), we have to delete four boundaries between *horses* and *suddenly*, and we cannot delete the boundary between *suddenly* and *got*, as shown in (31b). The value *n* is not consistent and the phrasing (31c) is unacceptable.

The theory has interesting consequences for syntax. First, heaviness of constituents can be represented as the number of boundaries at their right edge.

(32)
a. ϕ (zero)
b. [it] (stressed/independent pronoun)
c. [a [book]] (DP)
d. [a [new [book]]] (modified DP)
e. [a [book [on [French]]]] (modified DP)
f. [a [book [on [the [desk]]]]] (modified DP)

Generally, the longer a constituent becomes, the more boundaries it has at its right edge. A pronoun has only one boundary at its right edge as in (32b), while a modified DP can have five or more as in (32f). Thus, the number of boundaries at the edge of constituents represents the information status such as given/new. Generally, given constituents tend to be shorter and have less number of brackets at their right edge than new constituents, as (32b) shows. Then we have alternative to functional explanation using the notion given/new, which is not easy to define and formalize (cf. Newmeyer 1998).

Once heaviness is formulated as the number of brackets, we can explain why heavy constituents tend to be positioned at the right edge of a sentence.

(33)
a. [Ken [gave [[a [book [about [small hamsters]]]] [to Alice]]]]
b. [Ken [[gave [to Alice]] [a [book [about [small hamsters]]]]]]

There are five brackets between *hamsters* and *to* in (33a) while three between *Alice* and *a* in (33b). Assuming “Avoid Pause,” (33b) is preferred than (33a) because it does not have a long pause in the sentence. Thus this analysis provides an alternative to Hawkins’ (1994) theory of Early Immediate Constituents (EIC).

Linguistic structure goes well beyond a sentence. Merge combines sentences to make a paragraph, which are in turn merged with other paragraphs to make various units of discourse. The mapping theory also predicts longer pauses between sentences if they are separated by a large number of brackets.

(34) a. [[It’s [late]] [I’m [leaving]]] -> ... la[*ɾ*] I’m ...

b. [[It’s [very [late]]] [[Irene [and [I]]] [are [leaving]]]] -- … late Irene …

In (34a), *late* and *I’m* are separated by only three brackets, and Flapping applies to change *late* into *la[ɾ]*. Flapping does not apply in (34b), where *late* and *Irene* are separated by five brackets.

The mapping theory shed a new light on topic/focus and movement (Chapter 7). A focused constituent tends to make its own prosodic phrase because the other parts of the sentence are presupposed to lose its internal structure.

(35) [What do you think of a California rolls?]

a. ~~[~~[California rolls] ~~[~~I ~~[~~love ~~[~~to eat~~]]]]~~

b. [California rolls] I love to eat

c. / California rolls / I love to eat

d. California rolls I love to eat (*n*=1)

In (35a), the other constituent structures than the focused constituent *[California rolls]* are deleted to be (35b) because they are presupposed. The output of the mapping rule is (35c), which is easily changed into (35d) by deleting just one boundary between words.

Parsing of sentence structure is also affected by pause duration between words. Hearers interpret silent demibeats as syntactic brackets. If a silent demibeats is immediately followed by a word, it is interpreted as a left bracket. If it is immediately followed by another silent demibeat, it is interpreted as a right bracket.

(36) a. $\underline{x}\ \alpha \rightarrow [\alpha$

b. $\underline{xx} \rightarrow\ ]\underline{x}$

Then we can produce and percept a sentence from left to right. A speaker utters words incrementally with silent demibeats before Merge combines them. Hearers build phrase structure incrementally by changing silent demibeats into syntactic brackets. The derivation, Spell-Out, and parsing proceeds roughly in the following order:

(37)

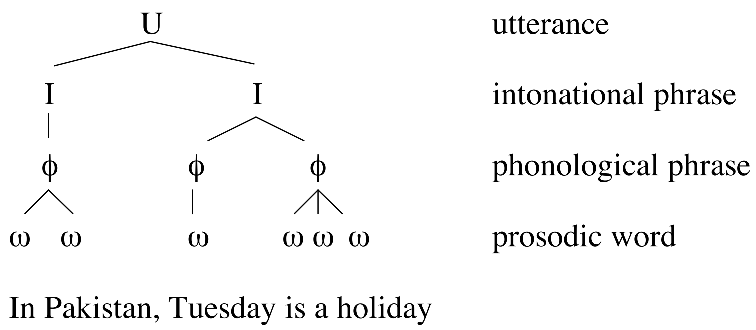
Speaker	PF	Hearer
[Meg	$\underline{x}$ Meg	[Meg
[Meg [loves [cats	$\underline{x}$ Meg $\underline{x}$ loves	[Meg [loves
[Meg [loves [cats	$\underline{x}$ Meg $\underline{x}$ loves $\underline{x}$ cats	[Meg [loves [cats
[Meg [loves [cats]	$\underline{x}$ Meg $\underline{x}$ loves $\underline{x}$ cats $\underline{x}$	[Meg [loves [cats
[Meg [loves [cats]]	$\underline{x}$ Meg $\underline{x}$ loves $\underline{x}$ cats $\underline{xx}$	[Meg [loves [cats]
[Meg [loves [cats]]]	$\underline{x}$ Meg $\underline{x}$ loves $\underline{x}$ cats $\underline{xxx}$	[Meg [loves [cats]]
[Meg [loves [cats]]] [She	$\underline{x}$ Meg $\underline{x}$ loves $\underline{x}$ cats $\underline{xxxx}$ She	[Meg [loves [cats]]] [She

As I have outlined above, this mapping theory has a lot of consequences for various aspects of grammar. I will elaborate each topic below.

1.4.3 Prosodic Phrases: Definition and Nature

Before turning to the discussion of each topic, let us define prosodic phrases and consider their nature. I will use the term "prosodic phrases" to refer to any level of prosodic categories, which include prosodic word, phonological phrase, intonational phrase, and utterance. For example, consider the following hierarchy of prosodic categories (cf. Selkirk 1984):

(38)



These categories represent grouping of phonological elements just as syntactic phrases such as DP and IP represent grouping of syntactic elements. The smallest unit in the hierarchy is a prosodic word, which basically corresponds to a syntactic word but may consist of a syntactic word and a clitic (e.g. *t'aime* in *Je t'aime* consisting of *te* and *aime*). A phonological phrase is a group of prosodic words and is considered to be a rhythmic unit such as foot. An intonational phrase is the domain where an intonation contour such as fall or rise appears. An utterance is the largest domain corresponding to a syntactic sentence.

A number of prosodic categories other than these have been proposed in the literature of prosodic phonology. Among them are clitic group, major and minor phrases, intermediate phrases, and focus phrases. I will argue that all these prosodic categories are just a variety of strings demarcated by prosodic boundaries in Section 3.2.

Chapter 2

Prosodic Phrasing and Bare Phrase Structure

In this chapter, I will propose a new theory of mapping from syntax to phonology in the minimalist framework. I will argue that the phrasing data from a number of languages, together with this mapping theory, give evidence for the bare phrase structure theory (Chomsky 1995).[1]

2.1. Bare Syntax-Phonology Mapping

Cinque (1993: 244) proposes a simplified version of Halle and Vergnaud's (1987) Nuclear Stress Rule. One of the rules is (1), which maps syntactic constituents onto metrical boundaries, as shown in (2):

(1) Interpret boundaries of syntactic constituents as metrical boundaries.

(2) ((*) (* (* (*))))
    [[Jesus] [preached [to the people [of Judea]]]]

In the first line of (2), Cinque shows metrical boundaries as parentheses that have directions, right [( ] and left [ )]. However, if the function of boundaries is to show the border or division between two strings, they do not need to have directions. The mapping rule I propose here is (3).

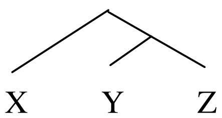
(3) Interpret boundaries of syntactic constituents [...] as prosodic boundaries / ... /.

---

[1] A part of this chapter is a revised version of Tokizaki (1999b).

This rule interprets boundaries of syntactic constituents as prosodic boundaries that have no direction, like bar lines in music. I assume here that the input to the rule (3) is the bare phrase structure, and not the X-bar theoretic phrase structure. I will argue about this point in section 2.3. For example, the rule (3) maps the right branching structure (4a) into the PF representation (4b), where X, Y, and Z schematically represent a word.

(4) a. [[X] [[Y][Z]]]



b. // X /// Y // Z ///

In (4b), we have two prosodic boundaries before X, three between X and Y, two between Y and Z, and three after Z.[2]

In this bare mapping theory, prosodic phrasing is to group words by deleting prosodic boundaries between them. The phrasing process can be formulated into the rule shown in (5), where *n* is a variable.

(5) Delete *n* boundaries between words. (*n*: a natural number)

---

[2] As we have seen in Chapter 1, the basic idea of the rule (3) is not unprecedented. There are similar ideas such as depth of syntactic boundaries (Cheng 1966:150), depth of embedding (Clements 1978: 29), Silent Demibeat Addition (Selkirk 1984:314, 1986:376, 388). However, as I argued in Section 1.4, these analyses cannot hold in the minimalist framework because they assume rather old versions of syntactic structure. See also Tokizaki (1988).

If we apply (5) to (4b) with *n*=1, 2, or 3, we get (6a), (6b), and (6c), respectively.

(6) a. / X // Y / Z // (*n*=1) --> (X) (Y) (Z)

b. X / Y Z / (*n*=2) --> (X) (Y Z)

c. X Y Z (*n*=3) --> (X Y Z)

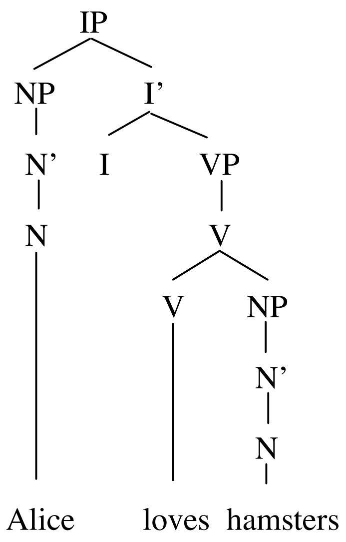
In (6a), one boundary is deleted in every sequence of boundaries, and there are two boundaries between X and Y, and one boundary between Y and Z. Thus we get three prosodic phrases (X), (Y), and (Z). In (6b), two boundaries are deleted in every sequence of boundaries, and there is one boundary between X and Y, but no boundary between Y and Z. Thus we get two prosodic phrases (X) and (Y Z). There is no boundary left in (6c) after three boundaries are deleted in every sequence of boundaries. The whole string is contained in a prosodic phrase (X Y Z).

To illustrate how the rules (3) and (5) work with the actual sentences, consider the following sentence:

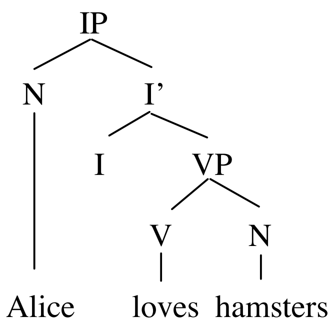
(7) Alice loves hamsters.

As I will argue later in section 2.3, I simply assume here that phrase structure is bare in the sense of Chomsky (1995). As Chomsky (1995:246) notes, “there is no such thing as a non-branching projection.” This is a consequence of the operation Merge, which combines two syntactic objects. Then the phrase structure of (7) is not the X-bar theoretic structure (8a) but the bare structure (8b).

(8) a. $[_{IP}\ [_{NP}\ [_{N'}\ [_{N}\ \text{Alice}]]]_i\ [_{I'}\ \text{I}\ [_{VP}\ [_{V'}\ [_{V}\ \text{loves}]\ [_{NP}\ [_{N'}\ [_{N}\ \text{hamsters}]]]]]]]$



b. $[_{IP}\ [_{N}\ \text{Alice}]\ [_{I'}\ \text{I}\ [_{VP}\ [_{V}\ \text{loves}]\ [_{N}\ \text{hamsters}]]]]$



I also assume the following convention for invisible syntactic objects:[3]

---

[3] Nespor and Scorretti (1984) also argue that empty categories have no effect on the various PF rules. For *wanna* contraction, see Tokizaki (1991), where I propose an analysis by Empty Category Principle instead of an intervening trace between *want* and *to*. See also Goodall (1991, 2006) for other analyses without intervening trace.

(9) Phonologically null elements and the constituents made by merging them with other syntactic objects are invisible to phonological rules.

By "phonologically null elements", I refer to trace, PRO, Infl, *v*, and so on. Given the convention (9), I and I' in (8b) are invisible to phonological rules. I is phonologically null and I' is made by merging I with VP. Thus phonological rules can "see" only some parts of the structure, which is shown in (10).[4]

(10) $[_{\text{IP}}$ $[_{\text{N}}$ Alice] $[_{\text{VP}}$ $[_{\text{V}}$ loves] $[_{\text{N}}$ hamsters]]]]

Following Chomsky (1995) and Collins (2002), I also assume that there are no labels in syntactic structure. With these assumptions, the mapping rule (3) applies to the "completely bare" structure (11).

(11) [[Alice] [[loves] [hamsters]]]

---

[4] If we assume VP-internal subject hypothesis as in (ia), the result is almost the same as (10) as shown in (ib) because the trace of subject and the VP are invisible in (ia).

(i) a. $[_{\text{IP}}$ $[_{\text{N}}$ Alice]$_{\text{i}}$ $[_{\text{I'}}$ I $[_{\text{VP}}$ t_{i} $[_{\text{V'}}$ $[_{\text{V}}$ loves] $[_{\text{N}}$ hamsters]]]]]]

b. $[_{\text{IP}}$ $[_{\text{N}}$ Alice] $[_{\text{V'}}$ $[_{\text{V}}$ loves] $[_{\text{N}}$ hamsters]]]]

The only difference between (10) and (ib) is that the sequence *love hamsters* is VP in (10) and V' in (ib), which disappears if we assume the label free structure as in (11).

The rule interprets the brackets in (11) and changes them into prosodic boundaries as in (12).

(12) // Alice /// loves // hamsters ///

Now the phrasing rule (5) deletes a number of boundaries between words to make longer prosodic phrases. If we apply this rule with *n*=1 to (8b), it deletes one boundary between words to give (13a). The three words are still separated by boundaries, and each word makes a prosodic phrase by itself.

(13)	a. / Alice // loves / hamsters //	(*n*=1)	-->	(Alice) (loves) (hamsters)
	b. Alice / loves hamsters /	(*n*=2)	-->	(Alice) (loves hamsters)
	c. Alice loves hamsters	(*n*=3)	-->	(Alice loves hamsters)

I assume here that the number of boundaries to be deleted (*n*) corresponds to the speed of utterance. The basic idea is that if the speaker utters the sentence faster, the more boundaries are deleted, and the longer phrases we get. If we suppose that *n*=2, that is, when the speaker talks faster, then we get (13b) as the result of applying the deletion rule (5). If *n*=3, the fastest in this case, the whole sentence is included in a prosodic phrase as in (13c), because there is no boundary left between words after deletion.[5] Thus we can capture the relation between the rate of speech and the length of prosodic phrases.

---

[5] I will discuss the level of prosodic phrases in Section 3.2. We will argue that *n* also relates to the levels of prosodic categories. If *n* is larger, then (5) makes larger prosodic domains (e.g. phonological phrases or

2.2 Branching and Prosodic Phrasing

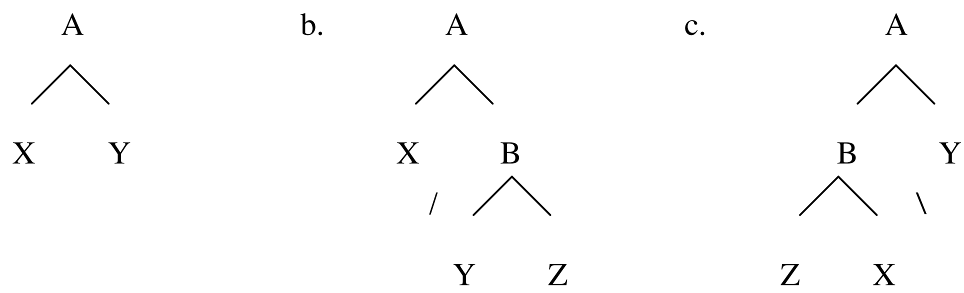
The bare mapping theory gives us a new insight into the relation between branching and prosodic phrasing. In a number of languages, there are some phonological rules that apply between X and Y in (14a), but not in (14b) or (14c).[6]

(14) a. $[_A$ X Y]

b. $[_A$ X $[_B$ Y Z]]

c. $[_A$ $[_B$ Z X] Y]

(15) a. b. c.



2.2.1 Right-Branching

Let us begin by looking at the data that show the difference between non-branching (14a) and right branching (14b). First, Cowper and Rice (1987:189f) show that Consonant Mutation in Mende applies in (16a) and (17a) but not in (16b) and (17b).[7]

---

intonational phrases). We will also argue that with this theory we could dispense with prosodic category hierarchy altogether.

[6] Left branching structure (14c), as well as right branching structure (14b), makes a prosodic boundary, as we will see in section 2.2.2. These cases pose an interesting problem on the view that the right/left branching structures are asymmetry as argued in Kubozono (1992:26, 1993:159). I will not go into detail here, however.

(16) a. [$_{S}$ [$_{NP}$ ndóláà] [$_{VP}$ w̱òtéà]] <- p̱òté ‘turn’

baby turn

‘the baby turned’

b. [$_{S}$ [$_{NP}$ tí] [$_{VP}$ [$_{V}$ ḵàkpángà] [$_{PP}$ ngì má]]] -> *tí g̱àkpángà ngì má

they surround him on

‘they surrounded him’

(17) a. mɛ̀hɛ̂ mɛ́ [$_{PP}$ [$_{P}$ à] [$_{NP}$ ḻòkó]] <- ṯòkó ‘hand, forearm’

food eat with hand

‘eat with fingers’

b. hɛ́1ɛ́ [$_{PP}$ [$_{P}$ a] [$_{DP}$ [$_{NP}$ ng̱úlí] [$_{D}$ í]]]

hang from tree Det

‘hang from the tree’

That is, the rule applies if the constituent in question does not branch, but it does not apply if the constituent branches. In (16a) the VP *wòtéà* does not branch and in (17a) the complement NP of P *lòkó* does not branch. On the other hand, in (16b) the VP *kàkpángà ngì má* branches, and in (17b) the complement DP of P *ngúlí í* branches.

---

[7] Cowper and Rice (1987) do not show the mutated form in the case of (17b). I suppose that it may be acceptable when uttered in a rapid speech rate. I show Cowper and Rice’s category labels instead of bare phrase structure here.

Second, Zec and Inkelas (1990:369) argue that the discourse particle *fa* in Hausa needs to be followed by a branching constituent as shown in (18).

(18) a. *Ya [$_{VP}$ [$_{V}$ sayi] <u>fa</u> [$_{NP}$ teburin]]
he bought table-DEF
'He bought the table.'

b. Ya [$_{VP}$ [$_{V}$ sayi] <u>fa</u> [$_{NP}$ [$_{A}$ babban] [$_{N}$ tebur]]]
he bought big table
'He bought a big table.'

In (18a), the object NP *teburin* does not branch, and *fa* cannot be inserted. In (18b), the object NP *babban tebur* branches, and *fa* is allowed to occur in the position preceding it.[8]

Third, Nespor and Vogel (1986:175) show that Italian Stress Retraction, which occurs to avoid stress crash, applies in (19a), but not in (19b).

---

[8] In fact, Hausa *fa* needs to be followed by a branching *constituent*, not just by more than one word, as shown in (i) (Zec and Inkelas 1990:370).

(i) * Ya [$_{S}$ [$_{VP}$ [$_{V}$ sayi] <u>fa</u> [$_{NP}$ teburin]] [$_{Adv}$ jiya]]
he bought table-DEF yesterday
'He bought the table yesterday.'

What is crucial to the insertion of *fa* is not just the length of strings following it but the length of the constituent following it. I will argue that we can explain the fact by the bare mapping theory in section 2.3.1. See note 14.

(19) a. Le [$_{NP}$ [$_{N}$ cítta] [$_{AP}$ nórdiche]] non mi piacciono. (<- cittá)

'I don't like Nordic cities.'

b. Le [$_{NP}$ [$_{N}$ cittá] [$_{AP}$ [$_{Adv}$ mólto] [$_{A}$ nordiche]]] non mi piacciono. (-> *cítta)

'I don't like very Nordic cities.'

The stress on the final syllable of *cittá* moves to the first syllable in (19a), but not in (19b). The AP in (19a) is non-branching and the AP in (19b) is branching.

Fourth, Rhythm Rule in English applies in (20a), but not in (20b) (Nespor and Vogel 1986:178, cf. Inkelas and Zec 1995:543).[9]

(20) a. John [$_{VP}$ [$_{V}$ pérseveres] [$_{Adv}$ gládly]] (<- persevéres)

b. John [$_{VP}$ [$_{V}$ persevéres] [$_{\&P}$ [$_{Adv}$ gládly] [$_{\&'}$ and diligently]] (-> *pérseveres)

In (20b), two adverbs are conjoined to make a branching &P. Inkelas and Zec (1995) also show a similar example as in (21).

(21) a. [$_{S}$ [$_{NP}$ Ánnemarìe] [$_{VP}$ héard]] (<- Ànnemaríe)

b. [$_{S}$ [$_{NP}$ Ànnemaríe] [$_{VP}$ [$_{V}$ héard] [$_{PP}$ about it already]]]

---

[9] For the analysis of coordinate structure as projection of a conjunction head, see Larson (1990) and Kayne (1994). The discussion holds even if we assume the traditional structure for coordinate structure because it is branching into two conjuncts and a conjunction.

(i) John [$_{VP}$ [$_{V}$ persevéres] [$_{AdvP}$ [$_{Adv}$ gládly] [$_{CONJ}$ and] [$_{Adv}$ diligently]]

Stress Retraction applies to *Annemarie* in (21a) where the VP is non-branching, but it does not apply to (21b) where the VP is branching.[10]

2.2.2 Left-Branching

Let us turn to left branching structure as shown in (14c). The phenomena that show the left branching effect on phrasing are not as many as the right branching effect we have just seen. However, they really exist.

First, According to Bickmore (1990:14), High Deletion in Kinyambo applies in (22a), but not in (22b).

(22) a. [$_S$ [$_{NP}$ abakozi] [$_{VP}$bákajúna]] <- abakózi 'workers'

workers they-helped

'the workers helped'

---

[10] Inkelas and Zec (1995:544) also show the following examples and argue that phonological phrasing (ϕ) is sensitive to complexity at the prosodic level.

(i) a. [Ánnemarìe áte it]$_\phi$

b. [Ánnemarìe áte]$_\phi$

c. [Ànnemaríe]$_\phi$ [áte]$_\phi$ [with her fingers]$_\phi$

d. [Ànnemaríe]$_\phi$ [áte and drank]$_\phi$

(ia) might be considered as a counterexample to the analysis here because *ate it* is a branching VP and induces Rhythm Rule as (1b). However, the pronoun *it* seems to be cliticized to the preceding verb *ate* to make a (phonological) word *ate-it*. The clitic nature of *it* can be seen in its contracted form *'t* as in *see't*. I will return to this point in Section 3.1.4.

b. [$_S$ [$_{NP}$ [$_N$ abakozi] [$_A$bak<u>ú</u>ru]] [$_{VP}$bákajúna]] -- bak<u>ú</u>ru 'mature'

workers mature they-helped

'The mature workers helped.'

High Deletion states that a High tone (´) in one word deletes the High tone in the word to its left. So the high tone in *abakózi* is deleted in (22a) where the subject NP does not branch, but the high tone in *bakúru* in (22b) is not deleted. In (22b) the subject NP is branching and the whole sentence has left branching structure.

Second, the first mora in a Japanese unaccented word loses its low tone and becomes high when the word is in the same prosodic phrase (minor phrase) with the preceding unaccented word which ends in high tone. Here I refer to this phenomenon as Low Deletion. Low Deletion applies to the second conjunct NP *nira* in (23a), but it does not apply in (23b). Initial low tones are shown with grave accents (`).[11]

(23) a. [$_{NP}$ [$_{NP}$ Mòmo-to] [$_{NP}$ n<u>i</u>ra-o]] yome-ni ageta. (<- nìra)

peach-and leek-Acc daughter-in-law-to gave

'I gave peaches and leeks to my daughter in law.'

b. [$_{NP}$ [$_{NP}$ [$_A$ Àmai] [$_N$ momo-to]] [$_{NP}$ n<u>ì</u>ra-o]] yome-ni ageta.

sweet peach-and leek-Acc daughter-in-law-to gave

'I gave sweet peaches and leeks to my daughter in law.'

---

[11] I owe to Azuma (1992), who argues that F_0 resetting disambiguates syntactically ambiguous sentences similar to (23). Low Deletion could be called lack of Initial Lowering (Selkirk and Tateishi 1988). I use the former term because it represents the sandhi nature of the phenomenon straightforwardly.

The first conjunct NP in (23a) is not branching, and the NP in (23b) is branching. The pitch contours are shown as Figure 1 and 2 below.

Figure 1

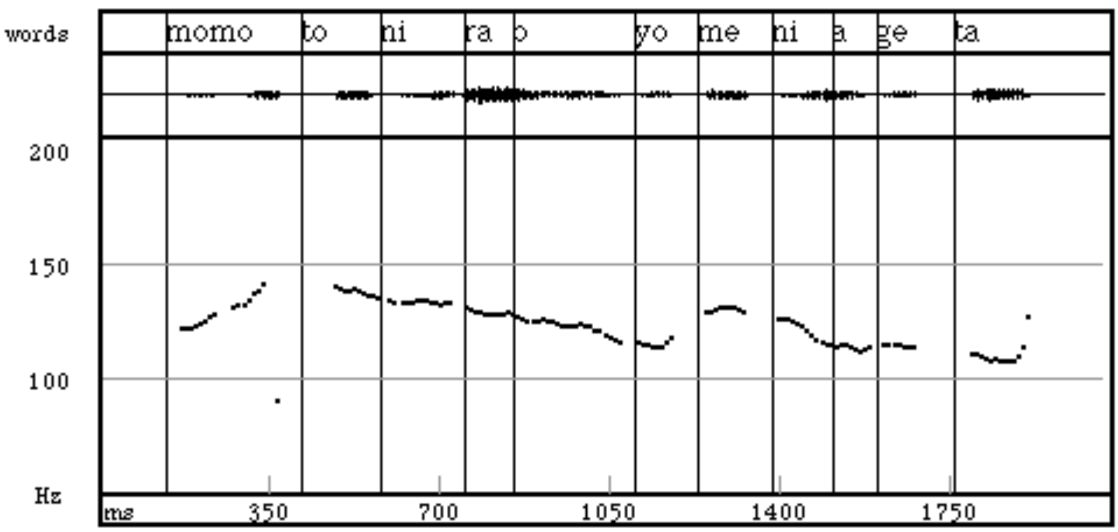
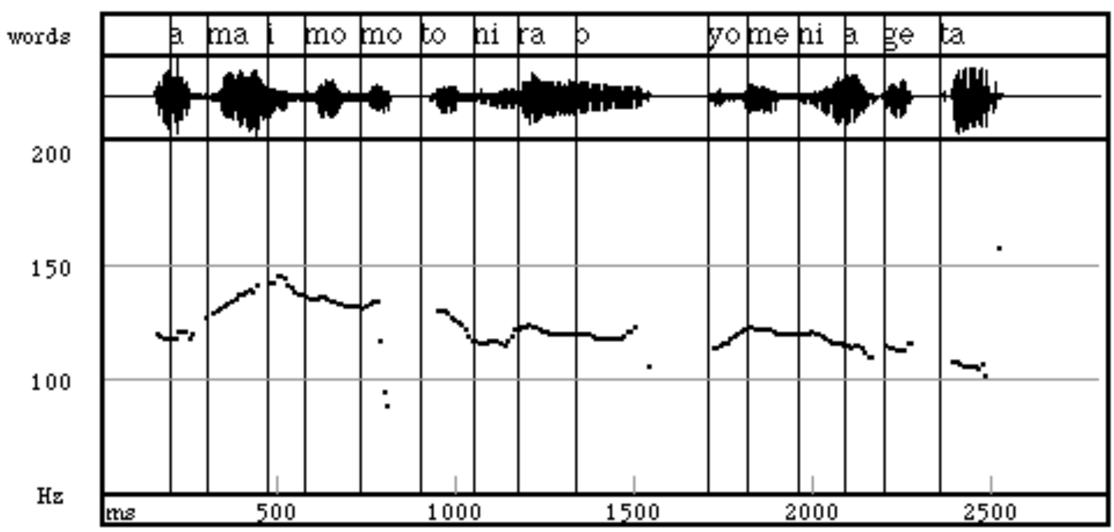


Figure 2



The low tone of the first mora in *nira* is deleted in Figure 1, and the low tone is not deleted in Figure 2.

In this section we have seen branching categories block the application of sandhi rules. I have argued that the blocking effect is caused by both right branching and left branching.

2.3 Bare Phrase Structure

2.3.1 Phrasal Phonology Supports the Bare Phrase Structure

All of the examples in section 2.2 show that the application of sandhi rules is blocked by branching categories. In this section, I will argue that this fact gives an empirical support to the bare phrase structure theory proposed by Chomsky (1995).

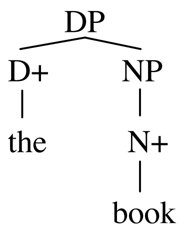
Chomsky (1986:4) posed the question about the existence of intermediate projection X'. He adopted the convention that single bar level structure as in (24a) need not be present when not required, as shown in (24b):

(24) a. $[_{NP}\ [_{N'}\ [_{N}$ pictures] [of John]]]

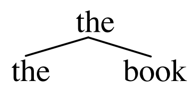
b. $[_{NP}\ [_{N}$ pictures] [of John]]

Chomsky (1995) further proposes a radical elimination of standard X-bar theory, a bare phrase structure theory, in which there are no such entities as XP, X^0 or a non-branching projection. For example, the string *the book* has (25b) instead of (25a):[12]

(25) a.



b.



(Chomsky 1995:246)

---

[12] See Collins (2001) for the idea of eliminating labels and its consequences.

If we assume this theory, rules specifying XP or X^0 in their formulation cannot be maintained as they are. They must be reformulated without using such entities. Among them are Phonological Phrase Formation (Nespor and Vogel 1986), the end based theory (Selkirk 1986), Phonological Phrase Algorithm (Zec and Inkelas 1990), and Wrap-XP (Truckenbrodt 1995). This problem is discussed in detail in Tokizaki (2005b). Here I will argue that the bare mapping and the prosodic facts we have just seen support the assumption that there is no non-branching projection.

First let us assume the standard X-bar theory and consider the case where the sister of XP is non-branching as in (26a) and the case where it is branching as in (26b). Notice that the X-bar theory is different from bare phrase structure in the numbers of syntactic boundaries. For example, let us consider the following X-bar theoretic structure:

(26) a. ... $[_{\alpha}$ A] $[_{XP}$ $[_{X'}$ $[_{X}$ B]]] ...

b. ... $[_{\alpha}$ A] $[_{XP}$ $[_{X'}$ $[_{X}$ B] $[_{\beta}$ ...]]] ...

In (26a) B is a head that has no complement while in (26b) B has a complement β. There are four brackets between A and B in both (26a) and (26b). However, the examples have the following structure in bare phrase theory:

(27) a. ... $[_{\alpha}$ A] $[_{X}$ B] ...

b. ... $[_{\alpha}$ A] $[_{X'}$ $[_{X}$ B] $[_{\beta}$ ...]] ...

In (27a) B has no complement and does not project while in (27b) B (X) projects into X' (or XP). The number of boundaries between A and B is two in (27a) and three in (27b).

Now let us consider the examples we have seen in 2.2. In the standard X-bar theory, the structures of (16a) and (16b) in Mende would be (28a) and (28b), respectively.

(28) a. [$_S$ [$_{NP}$ [$_{N'}$ [$_N$ ndóláà]]] [$_{VP}$ [$_{V'}$ [$_V$ wòtéà]]]]

b. [$_S$ [$_{NP}$ [$_{N'}$ [$_N$ tí]]] [$_{VP}$ [$_{V'}$ [$_V$ kàkpángà] [$_{PP}$ ngì má]]]]

Our mapping rule (3) would not make any difference in the number of prosodic boundaries between (28a) and (28b) if we assumed the standard X-bar theory with non-branching projections. If we applied (3) to (28a) and (28b), we would get the same number of boundaries, six boundaries, before *wotea* and *kakpanga* as shown in (29a) and (29b).

(29) a. //// ndóláà ////// wòtéà ////

b. //// tí ////// kàkpángà // ngì má ////

If we apply the deletion rule with *n*=6 to (29a) and (29b), there would be no boundary left in both of the cases as shown in (30).

(30) a. ndóláà wòtéà (*n*=6)

b. tí kàkpángà ngì má (*n*=6)

Alternatively, if we suppose that *n* is equal to or smaller than 5, there would be a prosodic boundary in both of the cases. For example, if *n*=5, the rule (3) deletes five boundaries between words to make (31).

(31) a. ndóláà / wòtéà (*n*=5)

b. tí / kàkpángà ngì má (*n*=5)

Neither (30) nor (31) is a welcomed result. We must explain the fact that Consonant mutation applies to (a) but not to (b) in (28)-(31).

On the other hand, if we assume bare phrase structure, the structures of (16a) and (16b) (with no labels) are the input to the rule (3), and their output is (32a) and (32b).

(32) a. // ndóláà // wòtéà //

b. // tí /// kàkpángà // ngì má ///

There are two boundaries before the verb *wòtéà* in (32a), and three boundaries before the verb *kàkpángà* in (32b). Then if we assume that *n* is 2 in the boundary deletion rule (5), there still remains a boundary before *kàkpángà* in (33b) and not before *wòtéà* in (33a).

(33) a. ndóláà wòtéà (*n*=2) <- pòté 'turn'

b. tí / kàkpángà ngì má / (*n*=2) -> *tí gàkpángà ngì má

Let us assume that Consonant Mutation applies to (33a) and (33b), the output of the boundary deletion rule, i.e. phonological phrasing. Then we can account for the difference in applicability of the mutation rule between (33a) and (33b). In (33b) the mutation process is blocked by the prosodic boundary between *tí* and *kàkpángà*. Thus we can explain the difference in phrasal phonology only when we assume the bare phrase structure instead of X-bar structure. Therefore the data in Mende support the bare phrase theory instead of X-bar theory.

Similarly, the data from (17) to (23) show that the bare mapping rule (3) and the phrasing rule (5), together with bare phrase structure, correctly predict the difference between the application cases and the non-application cases.[13] I will show the results of applying (3) to each of the structures (17) to (23) below.[14]

(34) a. mɛ̀hɛ̃́ mɛ // à // lòkó // (<- tòkó)

b. hɛ́lɛ́ // a /// ngúlí // í ///

(35) a. *Ya // sayi / fa / teburin //

b. Ya // sai / fa // babban // tebur ///

(36) a. Le // cítta // nórdiche // non mi piacciono. (<- cittá)

b. Le // cittá /// mólto // nordiche /// non mi piacciono. (-> *cítta)

(37) a. John // pérseveres // gládly // (<- persevéres)

b. John // persevéres /// gládly // and diligently // (-> *pérseveres)

(38) a. // Ánnemarìe // héard // (<- Ànnemaríe)

b. // Ànnemaríe /// héard // about it already ///

(39) a. // abakozi // bákajúna // (<- abakózi)

b. /// abakozi // bakúru /// bákajúna // (-- bakúru)

(40) a. // Mòmo-to // nira-o // yome-ni ageta. (<- nìra)

b. /// Àmai // momo-to /// nìra-o // yome-ni ageta.

[13] Uechi (1998) independently argues that non-branching XPs are invisible to phonology in Japanese.

[14] If we apply (3) to the example (i) in note 8, we have the following representation:

(i) * Ya /// sayi / fa / teburin /// jiya //

There are only two boundaries between *sayi* and *teburin*. Thus *fa* cannot be inserted there in (i) as well as in (35a).

In (34) to (40), there are two boundaries between the words in question in the (a) sentences and three boundaries there in the (b) sentences. We can represent the facts schematically as (41a) and (41b).

(41) a. ... A // B ...

b. ... A /// B ...

The boundary deletion rule with *n*=2 applies to (41a) and (41b) to give (42a) and (42b), respectively.

(42) a. ... A B ... (*n*=2)

b. ... A / B ... (*n*=2)

Assuming that (42) is the representation to which phonological rules apply, we can explain why the rules can apply to (42a) but not to (42b). This explanation is possible only when we assume the bare phrase structure without non-branching projection. Thus the phonological facts in (17) to (23) as well as (16) supports the bare phrase theory of Chomsky (1995).

2.3.2 Korean Voicing

Data from Korean seem to be a problem for this analysis. The rule of Obstruent Voicing voices plain consonants but not aspirated or tense consonants. Cho (1990) observes that Obstruent Voicing occurs between possessive noun and noun (43a), between verb of relative clause and noun (43b), and between object and verb (44a), but not between subject and verb (44b).[15]

(43) a. [$_{NP}$ [$_{NP}$ Suni–ɨy] [$_{N}$ c̲ip]] —> Suniɨy j̲ip

Suni's house

'Suni's house'

b. [$_{NP}$ [$_{S}$ [$_{NP}$ kɨ–ka] [$_{VP}$ mək–nɨn]] [$_{N}$ p̲ap]] —> kɨga məŋnɨn b̲ap

he-Sub eat-Mod rice

'the rice he is eating'

(44) a. [$_{VP}$ [$_{NP}$ kɨlim–ɨl] [$_{V}$ p̲ota]] —> kɨrimɨl b̲oda

picture-Acc see (— kɨrimɨl p̲oda)

'look at the picture'

b. [$_{S}$ [$_{NP}$ kæ–ka] [$_{VP}$ c̲anta]] — kæga c̲anda

dog-Nom sleep

'The dog is sleeping.'

Note that Cho observes that Obstruent Voicing occurs in object-verb case as (44a). All of my three informants, however, pronounce the voiceless labial sound as shown in the parenthesis in (44a). Here I assume that there is no voicing between verb and its object.

---

[15] For Korean Obstruent Voicing, see Sohn (1999) and Hirano (2001).

If we assume the bare phrase structure without bar-levels and category labels, the structure of (43) to (44) are as follows:

(45) a. [[Suni–ɨy] [c̲ip]]

b. [[[kɨ–ka] [mək–nɨn]] [p̲ap]]

(46) a. [[kɨlim–ɨl] [p̲ota]]

b. [[kæ–ka] [c̲anta]]

If we apply the mapping rule (3) to these syntactic brackets, we get (47) and (48).

(47) a. // Suni–ɨy // c̲ip //

b. /// kɨ–ka // mək–nɨn /// p̲ap

(48) a. // kɨlim–ɨl // p̲ota //

b. // kæ–ka // c̲anta //

There is no difference in the number of boundaries between voicing cases (47a-b) and non-voicing cases (48a-b). It seems that the domain of this voicing rule is restricted within the topmost NP and that speech rate does not affect the process. Note that this case is a problem for the end-based theory and other mapping theories as well. The right edge of NP, as well as the left edge of NP, seems to make a boundary in object-verb cases as shown in (44a). More research in detail must be done, but I will leave this matter open here.

2.4 Readjustment with Multiple Spell-Out

It might be argued that this mapping theory cannot handle the so-called syntax-prosody mismatches, like the following example from Chomsky and Halle (1968:372):

(49) a. [[This] [[is] [[the] [[cat] [[that] [[caught] [[the] [[rat] [[that] [[stole] [[the] [cheese]]]]]]]]]]]]

b. // This /// is /// the /// cat /// that /// caught /// the /// rat /// that /// stole /// the // cheese ////////////

c. (This is the cat) (that caught the rat) (that ate the cheese)

Our rule (3) maps (49a) to (49b). Given that there are no more boundaries after <u>cat</u> and <u>rat</u> than in any other place in (49b), how do we get the actual phrasing (49c)?

Chomsky (1998:20) argues that a phase of derivation is CP or *v*P, and that derivation proceeds phase by phase. (50), for example, has the four phases bracketed:

(50) [John [<u>t</u> thinks [Tom will [*t* win the prize]]]]

Chomsky (1998:48) further proposes that Spell-Out is contingent on feature-checking operations and that Spell-Out applies cyclically, possibly at the phase level, in the course of the (narrow syntactic) derivation. Let us assume that this approach is correct and consider the derivation of (49a). Then (49a) has the six phases bracketed below:

(51) [This [is the cat [that [caught the rat [that [stole the cheese]]]]]]

The following structures are sent to PF in turn in the course of derivation:

(52) a. [[stole] [[the] [cheese]]]

b. [that]

c. [[caught] [[the] [rat]]]

d. [that]

e. [[is] [[the] [cat]]]

f. [this]

If we assume that the mapping rule (3) applies every time a structure is sent to PF, the outputs are (53).

(53) a. // stole /// the // cheese ///

b. / that /

c. // caught /// the // rat ///

d. / that /

e. // is /// the // cat ///

f. / this /

After the whole sentence (49a) is sent to PF, its PF representation is (54).

(54) / this / // is /// the // cat /// / that / // caught /// the // rat /// / that / // stole /// the // cheese ///

In (54), there are four boundaries before the two occurrences of *that*. Thus we predict the phrasing (49c) straightforwardly. If we apply the phrasing rule (5) with n=3, we get the right result (55):

(55) this is the cat / that caught the rat / that stole the cheese

Thus we can explain this case without the readjustment rule assumed in Chomsky and Halle (1968:372), which converts sentences with (multiple) embedded clauses into sentences dominating sister-adjoined clauses.

2.5 Summary

In this chapter, I have proposed a bare theory of syntax-phonology mapping. I argued that prosodic data involving the branching/non-branching effect support the bare phrase structure theory rather than the X-bar theory. I also showed that Multiple Spell-Out solves the problem of syntax-phonology mismatch in the case of multiply embedded relative clauses.

Chapter 3

An Alternative to End-Based Prosodic Theory

The bare mapping theory proposed in Chapter 2 has a number of consequences. In this chapter, I would like to present an alternative analysis to the end-based prosodic theory proposed by Selkirk (1986) among others. I will argue that the edge parameter of phonological phrasing is derived from the syntactic head parameter. I will also argue that her strict layer hypothesis is explained in terms of prosodic boundaries. The goal of this chapter is to show that if we assume the bare mapping theory, the edge parameter and the strict layer hypothesis are no longer necessary.

3.1 Deriving the Edge Parameter from the Head Parameter

3.1.1 Bare Phrase Structure and Bare Mapping

First, let us reconsider the edge parameter proposed by Selkirk (1986) and Chen (1987). As I reviewed it in section 1.2.2, the end-based theory assumes that languages have the edge parameter in prosodic phrasing whose values are right or left. For example, Chi Mwi:ni shows that the right edge of a lexically headed XP is a phonological phrase boundary.

(1) a. $[_{VP}$ $[_{V'}$ $[_{V}$ pa(:)nzize] $[_{NP}$ cho:mbo]] $[_{NP}$ mwa:mba]]

'He ran the vessel on to the rock'

b. ....$]_{Xmax}$ ....$]_{Xmax}$

c. $_{PPh}($______________________$)$ $_{PPh}($________$)$

Selkirk argues that in (1) the left edge of the NP *cho:mbo* does not make a prosodic boundary but the right edge of it does.

On the other hand, Selkirk and Tateishi (1988, 1991) argue that Japanese has left as the value of the phrasing parameter. The following example shows that verbs take their complements to their left (Selkirk and Tateishi 1991:524):

(2) a. $[_{S}$ $[_{NP}$ $[_{NP}$ Ao'yama-no] $[_{N}$Yama'guchi-ga]] $[_{VP}$ $[_{NP}$ ani'yome-o] $[_{V}$yonda]]]

Aoyama-from Yamaguchi-Nom sister-in-law-Acc called

'Mr. Yamaguchi from Aoyama called his sister-in-law.'

b. $_{MaP}$(Ao'yama-no Yama'guchi-ga) $_{MaP}$(ani'yome-o yonda)

They argue that the right edge of the NP *Ao'yama-no* does not make a prosodic boundary but the right side of it does.

In this way, according to the end-based theory, languages can be grouped in terms of the edge parameter of prosodic phrasing. The following is a list of languages that have right and left as the edge parameter value:

(3) Right edges of lexically headed XPs:

Chi Mwi:ni (Kisseberth and Abasheikh 1974, Selkirk 1986)

Kimatuumbi (Odden 1987)

Xiamen (Chen 1987)

Papago (Hale and Selkirk 1987)

(4) Left edges of lexically headed XPs:

Ewe (Clements 1978)

Japanese (Selkirk and Tateishi 1991)

Korean (Cho 1990)

Northern Kyungsang Korean (Kenstowicz and Sohn 1997)

Shanghai Chinese (Selkirk and Shen 1990)

Notice that there seems to be a parallelism between the syntactic head parameter and the prosodic edge parameter. Head-initial (*i.e.* complement-right) languages such as Chi Mwi:ni (cf. (1)) and Xiamen have right edge as the parameter value, and head-final (*i.e.* complement-left) languages such as Japanese (cf. (2)) and Korean have left as the value. If we are on the right track, we can dispense with the edge parameter by deriving its effect from the head parameter with the bare mapping theory.[1]

First, let us reconsider the example (1) from Chi Mwi:ni in terms of bare phrase structure.

(5) $[_{VP}$ $[_{V'}$ $[_{V}$ pa(:)nzize] $[_{N}$ cho:mbo]] $[_{N}$ mwa:mba]]

---

[1] We cannot explain optional tone sandhi in Shanghai straightforwardly if we suppose that the phrase structure of Shanghai is the same as that of Xiamen as Hale and Selkirk (1987:179) argue. One possible explanation is to suppose that the value of *n* in (8) in Shanghai is smaller than that in Xiamen. See also Selkirk and Shen (1990:335). I will discuss this issue in 3.1.3.

Chi Mwi:ni is head-initial (*i.e.* complement-right) and has right as the edge parameter value. We can explain why this is the case with our bare mapping theory. In Chapter 2, I proposed the following mapping rule as shown in (6).

(6) Interpret boundaries of syntactic constituents […] as prosodic boundaries / … /.

Now let us consider the reported data in turn. First consider Chi Mwi:ni (5). As we have seen, the mapping rule places the minimum number of prosodic boundaries, that is 2, between heads and non-branching complements because they are sisters in phrase structure. It also places three boundaries between the first object and the second object if they are non-branching as shown in (7).

(7) /// pa(:)nzize // cho:mbo /// mwa:mba //

If we apply the boundary deletion rule (8) with *n*=2, which I proposed in Chapter 2, to (7), we have the demarcated string as in (9).

(8) Delete *n* boundaries between words. (*n*: a natural number)

(9) / pa(:)nzize cho:mbo / mwa:mba (*n*=2)

This is the right prosodic phrasing for the sentence. The left edge of the first object does not make a prosodic boundary because the object is the sister of the preceding verb. The right edge of the object makes a prosodic boundary because the second object is not the sister of the first object but the sister of the category branching into the verb and the first object. Thus we do not have to specify the edge parameter of the language as right. The phrasing pattern is predicted from phrase structure.

This also holds with head-final languages like Japanese. As the examples in (2) show, verbs take their complements to their left. I will show bare phrase structure and the result of applying the mapping rule (6) together below. Consider (10) for example.

(10) a. $[_S$ $[_{NP}$ $[_{NP}$ Ao'yama-no] $[_N$ Yama'guchi-ga]] $[_{VP}$ $[_{NP}$ ani'yome-o] $[_V$ yonda]]]
Aoyama-from Yamaguchi-Nom sister-in-law-Acc called
'Mr. Yamaguchi from Aoyama called his sister-in-law.'

b. $_{MaP}$(Ao'yama-no Yama'guchi-ga) $_{MaP}$(ani'yome-o yonda)

In (10a), the subject NP branches. So there are four boundaries between the head of the subject NP *Yamaguchi-ga* and the object NP *ani'yome-o*, and only two boundaries between the verb *yonda* and its object *ani'yome-o*, as shown in (11a):

(11) a. /// Ao'yama-no // Yama'guchi-ga //// ani'yome-o // yonda ///

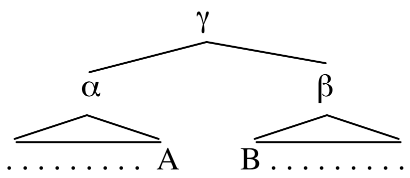
b. / Ao'yama-no Yama'guchi-ga // ani'yome-o yonda / (*n*=2)

We can explain the phrasing (10b) straightforwardly as shown in (11b) without assuming that Japanese has left as the edge parameter value.

3.1.2 Syntactic Constituents and Prosodic Boundaries

Let us consider the relation between syntactic constituents and prosodic boundaries in general. Suppose that α and β are sisters of γ, and that A and B are as follows. A is a word dominated by and is the right edge of α. B is a word dominated by and is the left edge of β. Or α equals A and β equals B. This is shown with a tree diagram in (12).

(12)



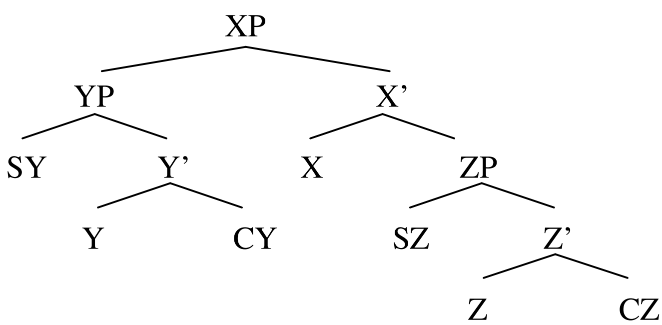
This is shown with brackets in (13) where possible brackets are italicized.

(13) $[_{\gamma}\, [_{\alpha} \ldots \mathit{[} \ldots \mathit{[} \ldots \mathrm{A}\mathit{]]}\,]\, [_{\beta}\, \mathit{[[}\, \mathrm{B} \ldots \mathit{]} \ldots \mathit{]} \ldots]]$

We can make the following generalization. The number of the boundaries between words is minimum when both α and β are non-branching. The deeper A or B is embedded in α or β, the larger the number of brackets between A and B becomes.

Let us consider what phonological representations the mapping rule makes in different syntactic structures. First consider the syntactic structure of head initial languages. For example, look at the following right-branching structure where X-Z is a head word, S specifier, and C complement. SY, for example, shows the specifier of Y.

(14)



Here I show the X-bar theoretic structure for the purpose of exposition. (14) is represented as (15) with brackets.

(15) $[_{XP}\ [_{YP}\ \mathrm{SY}\ [_{Y'}\ \mathrm{Y\ CY}]]\ [_{X'}\ \mathrm{X}\ [_{ZP}\ \mathrm{SZ}\ [_{Z'}\ \mathrm{Z\ CZ}]]]$

Applying the bare mapping rule (5), we get the following representation:

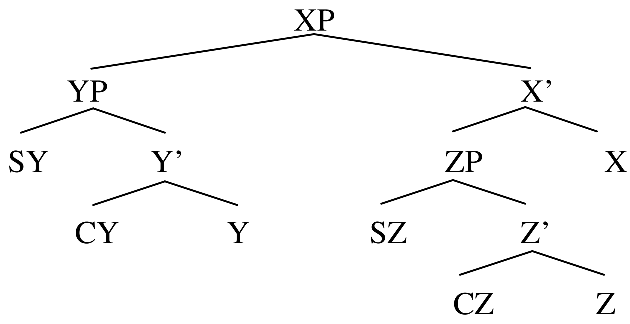
(16) // SY / Y CY /// X / SZ / Z CZ ///

The number of boundaries between CY and X is three. CZ also has three boundaries to its right. CY is in the right edge of YP, and CZ is in the right edge of ZP and XP. On the other hand, the number of boundaries between X and SZ is one. SY has two boundaries to its left. SZ is in the left edge of ZP, and SY is in the left edge of YP and XP. Thus, the bare mapping theory predicts more boundaries at the right edge of a maximal projection than at the left edge in right-branching structure.

(17) // SY / Y CY /// X / SZ / Z CZ ///
$[_{XP}\ [_{XP}$ SY Y CY] X $[_{XP}$ SZ Z CZ]]

Next, consider the syntactic structure of head final languages. For example, look at the following (partly) left-branching structure where I assume that specifiers are merged at the left of the intermediate projection of heads:

(18)



(18) is represented as (19) with brackets.

(19) $[_{XP}\ [_{YP}\ SY\ [_{Y'}\ CY\ Y]]\ [_{X'}\ [_{ZP}\ SZ\ [_{Z'}\ CZ\ Z]]\ X]]$

Applying the bare mapping rule (6), we get the following representation:

(20) // SY / CY Y //// SZ / CZ Z // X //

The number of boundaries between Y and SZ is four. SY has two boundaries to its left. SZ is in the left edge of ZP, and SY is in the left edge of YP and XP. On the other hand, the number of boundaries between Z and X is two. X has two boundaries to its right. SZ is in the left edge of ZP, and SY is in the left edge of YP and XP.

(21) // SY / CY Y //// SZ / CZ Z // X //

$[_{XP}\ [_{YP}$ SY CY Y] $[_{ZP}$ SZ CZ Z] X]

The position between Y and SZ corresponds to both the right edge of YP and the left edge of ZP in (21). Things are not as clear as in the right-branching case. However, if the position between Z and X, which corresponds to the right edge of ZP, does not block any prosodic rule, we may conclude that the left edge of a maximal projection is more relevant in prosodic phrasing than its right edge. This is what the end based theory predicts in right-branching structure. The bare mapping can make the same prediction as shown in (21).

3.1.3 Shanghai Chinese

The most difficult problem for the bare mapping theory is to explain the different parameter values among Chinese dialects. As the lists in (3) and (4) show, the value of the edge parameter is reported to be right in the Xiamen dialect and left in Shanghai Chinese (Chen 1987, Selkirk and Shen 1990). These languages share the basic properties of grammar, especially word order and phrase structure. Both languages are syntactically head initial. If we are trying to derive the edge parameter from the syntactic head parameter, how can we make prosodic difference in two languages that have the same parametric value of the syntactic parameter?

A solution is to suppose that the sensitivity to boundaries is different in two languages. Let us first consider the rules and the data of these languages. Chen (1987: 131) argues that Tone Group Formation in Xiamen can be formulated as (22).

(22) Mark the right edge of every XP with #, except where XP is an adjunct c-commanding its head.

The rule (22) correctly predicts the tone group boundary (#) in (23a) and (23b).

(23) a. yi tsiong hit pun ts'eq # sang hoo tang-oq

he Obj-marker that Cl book give to schoolmate

'He gave that book to his schoolmate.'

b. yi kap tang-oq # kai-siao tsit e lu-ping-yu

he to schoolmate introduce one Cl girlfriend

'He introduced a girlfriend to his schoolmate.'

hit pun ts'eq in (23a) and *kap tang-oq* in (23b), which are XPs, have a tone group boundary to its right.

Selkirk and Shen (1990:320, 332) argues that Shanghai Chinese has syntax-phonology mapping rules with the parameter setting in (24) and (25).

(24) Shanghai Chinese Prosodic Word Rule: (p. 320)

Prosodic Word: {Left, Lex^0}

where Lex^0 stands for word belonging to the lexical categories N, V, A.

(25) Shanghai Chinese Major Phrase Rule: (p. 332, p. 328)

Major Phrase: {Left, Lex^{max}}

Shanghai and Xiamen has almost the same syntax, but their edge parameter values seem to be different. First, let us examine the data shown by Selkirk and Shen (1990).

(26) a. ‘zaw ‘mo

toward horse

(LH) (LH)

‘toward the horse’

b. peq ‘mo tshaw

give horses vegetables

(MH)(LH) (MH)

(26a) is a crucial example. Prepositions and their objects make their own prosodic phrase. This is also the case with verbs and their objects as shown in (26b). Remember that in Xiamen prepositions or verbs and their objects are grouped into the same prosodic phrase as shown in (23a) and (23b).

How can we derive the prosodic phrasing in Shanghai if we eliminate the phonological edge parameter? First, note that the prosodic domain in Shanghai is smaller than that of Xiamen or Taiwanese. Consider the following examples cited from Yip (2002:118):

(27) V-NP

a. Taiwanese: One tonal domain

[$_{V}$ pang] [$_{NP}$ hong-ts’e] ‘fly kite’

fly kite

b. Shanghai: Two tonal domains

[$_{V}$ taN] # [$_{NP}$ ‘niN] ‘hit people’

hit people

As shown in (24), Selkirk and Shen (1990) also assume that the prosodic domain in question is prosodic word in Shanghai unless there is focus effect. Then, if we assume that the variable *n* in the boundary deletion rule (8) is relatively small, say *n*=1, in Shanghai, we can explain the data in (26).

(28) a. $[_{PP}$ $[_{P}$ ‘zaw] $[_{N}$ ‘mo]]

toward horse

‘toward the horse’

b. // ‘zaw // ‘mo //

c. / ‘zaw / ‘mo / (*n*=1)

d. (LH)(LH)

(29) a. $[_{VP}$ $[_{V}$ peq] $[_{N}$ ‘mo] $[_{N}$ tshaw]]

give horses vegetables

b. // peq // ‘mo // tshaw //

c. / peq / ‘mo / tshaw / (*n*=1)

d. (MH) (LH) (MH)

The phrase structures in (28a) and (29a) are interpreted as (28b) and (29b) by the mapping rule (5) and some of their boundaries are deleted as in (28c) and (29c) to make the prosodic domain shown in (28d) and (29d).

Second, Yip (2002:121) suggests that the prosodic domain in Shanghai is determined partly by rhythm. Look at the following examples from Selkirk and Shen (1990:321):

(30) a. $[_{VP}$ $[_V$ ‘z] $[_{PP}$ $[_P$ ‘laq] $[_N$ ‘zawNhe]]

live at Shanghai

(L H) (L H) <- LH ~~LM~~ LH ~~MH~~

‘live in Shanghai’

b. $[_{VP}$ $[_V$ tsou] $[_{PP}$ $[_P$ taw] $[_N$ ‘noetsiN]]

walk to Nanjing

(M H) (L H) <- MH ~~MH~~ LH ~~HL~~

As each of the sequences in (30) has four morphemes, it might be natural to divide them into two for rhythmic reasons. I would like to argue that the following structure is the input to the syntax-phonology mapping rule:

(31) a. $[_{VP}$ $[_V$ ‘z] $[_{PP}$ $[_P$ ‘laq] $[_N$ $[_N$ ‘zawN] $[_N$ he]]]]

b. $[_{VP}$ $[_V$ tsou] $[_{PP}$ $[_P$ taw] $[_N$ $[_N$ ‘noe] $[_N$ tsiN]]]]

‘zawNhe and ‘noetsiN are a kind of compounds as is evident from the fact that they consist of two chinese characters (Shanghai 上海 is ‘up-sea’ and Nanjing 南京 is ‘south-city’.) Then the mapped phonological representations are (32a) and (32b).

(32) a. // ‘z /// ‘laq /// ‘zawN // he ////

b. // tsou /// taw /// ‘noe // tsiN ////

It is not unnatural to put a prosodic boundary between the preposition and its object because there are three boundaries (the maximum sequence in the example) there as well as between the verb and the preposition. Eurhythmic consideration may well choose the actual phrasing. This is also the case with a personal pronoun embedded as a possessive in a post verbal noun phrase as in (33).

(33) a. [$_{VP}$ [$_{V}$ taN] [$_{DP}$ [$_{D}$ ‘ngu] [$_{N}$ ‘njitsz]]]

hit 1SG son

(M H) (L H) <- MH ~~LH~~ LH ~~MH~~

b. [$_{VP}$ [$_{V}$ taN] [$_{DP}$ [$_{D}$ ‘ngu] [$_{N}$ [$_{N}$ ‘nji] [$_{N}$ tsz]]]]

c. // taN /// ‘ngu /// ‘nji // tsz

The alleged structure (33a) can be analyzed as (33b) where *‘nji-tsz* 儿子 is a compound.

However, there are some data which need careful examination. The following examples have four words, which are grouped into three and one:

(34) a. [$_{VP}$ [$_{V}$ taw] [$_{NP}$[$_{QP}$ [$_{Q}$ ʔiq] [$_{CL}$ pe]] [$_{N}$ ‘zo]]]

pour one cup tea

(M H L) (LH) <- MH ~~MH~~ ~~HL~~ LH

‘pour a cup of tea’

b. [$_{VP}$ [$_{V}$ ‘ma] [$_{NP}$[$_{QP}$ [$_{Q}$ tsi] [$_{CL}$ po]] [$_{N}$ taw]]]

buy how many knife

(L H L) (HL) <- LH ~~MH~~ ~~MH~~ HL

‘buy some knives’

Selkirk and Shen (1990) argue that focus or semantic weight is involved in Shanghai. I speculate that this is the case in these examples as well. Duanmu (1992:74) argues that *ʔiq* and *pe* in (34a) are ‘function words’ which do not carry stress. *Taw* and *‘zo* are stress bearing units. Then these cases are the matter of phonology, and not the matter of syntax-phonology interface. This claim can be supported by the following examples also from Selkirk and Shen (1990):

(35) a. [$_{VP}$ [$_{V}$ taw] [$_{NP}$[$_{QP}$ [$_{Q}$ ʔiq] [$_{CL}$ pe]] [$_{N}$ ‘zo]]]
pour one cup tea
(MH) (M H) (LH) <- MH MH ~~HL~~ LH
‘pour one cup of tea’

b. [$_{VP}$ [$_{V}$ ‘ma] [$_{NP}$[$_{QP}$ [$_{Q}$ tsi] [$_{CL}$ po]] [$_{N}$ taw]]]
buy how many knife
(LH) (M H) (HL) <- LH MH ~~MH~~ HL
‘how many knives ... buy?’

In (35a) and (35b), *ʔiq* and *tsi* are used as quantifiers. These words can start their own domains if they have semantic content. This fact is not surprising if we assume the bare mapping theory.

(36) a. // taw //// ʔiq // pe /// ‘zo ///
b. // ‘ma //// tsi // po /// taw ///

If we apply the boundary deletion rule with *n*=2, we get the right phrasing:

(37) a. taw // ʔiq pe / ‘zo /
b. ‘ma // tsi po / taw /

Thus, we can explain the data in Shanghai which might be a problem in deriving the prosodic edge parameter from the syntactic head parameter. It is quite a progress, I believe, that we can dispense with the edge parameter which is not without problems.

3.1.4 Clitics and Function Words

Before we move on to the next section, let us consider how to deal with clitics (and function words) in our mapping theory. There seems to be at least three alternative hypotheses: (i) clitics are invisible to syntax-phonolgy mapping rules; (ii) clitics have a syntactic boundary only on their left or right; (iii) clitics have a boundary on their left and right. The hypothesis (i), proposed by Selkirk (1984) among others, makes nice predictions in cases such as (34) above. However, clitics make difference in prosody and the length of the sequence.

(38) a. It's a long story.

b. I don't get it.

The rhythm and the length of (38a) and (38b) are different from that of *long story* and *get.* Function words *it's, a, I,* and *it* make their own syllable and make the sequence longer. The hypothesis (iii) cannot distinguish between independent words and clitics. For example, the prosody in (39a) is different from that in (39b).

(39) a. Línda plays ténnis.

b. She pláys it.

Thus, I would like to explore the hypothesis (ii) here. I assume that (39b), for example, has the syntactic representation (40a) which is interpreted as (40b).

(40) a. [She [pláys] it]

b. / She / plays / it /

As shown in (40a), I assume that clitics and unstressed function words have a boundary on only one side of them. *She* has a boundary on its left and *it* has a boundary on its right. Thus we can distinguish (39a), which has the representations (41a) and (41b), from (39b).

(41) a. [[Linda] [[plays] [tennis]]]

b. // Linda /// plays // tennis ///

If we apply the boundary deletion rule with *n*=1 to (40b) and (41b), we get (42a) and (42b), respectively.

(42) a. She plays it

b. / Linda // plays / tennis //

(42) correctly predicts that the whole sentence in (42a) makes a prosodic category (a prosodic word) and that each word in (42b) makes its own prosodic category.

Now let us go back to the problematic examples including clitics and function words, which we have seen in section 2.2.1, note 10. With the idea that clitics have a syntactic boundary only on one side, we can explain why (43a) and (43b) behave in the same way (cf. Inkelas and Zec 1995).

(43) a. [Ánnemarìe áte it]$_{\phi}$

b. [Ánnemarìe áte]$_{\phi}$

The syntactic structure of (43a) and (43b) are (44a) and (44b), respectively.

(44) a. [[Ánnemarìe] [áte] it]]

b. [[Ánnemarìe] [áte]]

The mapping rule applies to them to give (45a) and (45b).

(45) a. // Ánnemarìe // áte / it //

b. // Ánnemarìe // áte //

Note that the number of the boundaries between the subject and the verb in (45a) is the same as that in (45b). The deletion rule applies to (45a) and (45b) with n=2. Then there are no boundaries in the output (46a) and (46b) which correctly predicts that the entire sentences are contained in a prosodic phrase as shown in (43a) and (43b).

(46) a. Ánnemarìe áte it (n=2)

b. Ánnemarìe áte (n=2)

Thus, we can explain the behavior of clitics and function words with the bare mapping theory straightforwardly.

3.2 Deconstructing prosodic hierarchy

In this section, I will point out a number of problems with the prosodic hierarchy theory (cf. Selkirk 1984 among others) and will propose an alternative theory based on prosodic boundaries.[2] I will also show that the boundary-based theory explains phonological data straightforwardly and that it allows us to derive the effects of the strict layer hypothesis (Selkirk 1984) from a simple phrasing rule.

3.2.1 Problems with the Prosodic Hierarchy Theory

Based on the idea that there are categories in prosody as well as in syntax, Selkirk (1984:26) shows the list of prosodic categories for English as in (47).

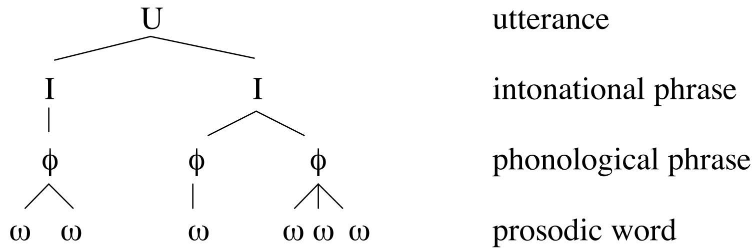
(47) a. intonational phrase (IP)
b. phonological phrase (PhP)
c. prosodic word (PWd)
d. foot (Ft)
e. syllable (Syl)

To illustrate the hierarchy, let us consider the following sentence:

---

[2] This section is a modified version of Tokizaki (2002).

(48)



In Pakistan, Tuesday is a holiday

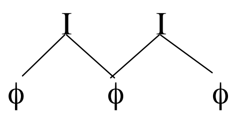
In (48), the node U immediately dominates Is, the next level category, which in turn immediately dominate ϕs. Selkirk (1984) proposes the strict layer hypothesis (SLH) to the effect that a category of level *i* in the hierarchy immediately dominates (a sequence of) categories of level *i*–1. Here I show the formulation of SLH by Nespor and Vogel (1986:7) (cf. Hayes 1989:204, Ladd 1996:238).

(49) a. A given nonterminal unit of the prosodic hierarchy, X^{P}, is composed of one or more units of the immediately lower category, X^{P-1}.

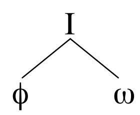
b. A unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part.

The hierarchy in (48) conforms to the hypothesis. Ladd (1996:239) argues that if this hypothesis is correct, the following types of hierarchies cannot be allowed in the prosodic hierarchy theory.

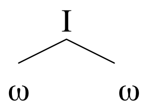
(50) a. Multiple domination



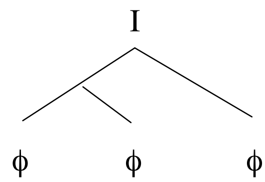
b. Heterogeneous sisters



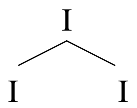
c. Skipping of levels



d. Unlabelled nodes



e. Recursion



The structures in (50b), (50c), (50d), and (50e) violate (49a), and (50a) violates (49b).

The prosodic hierarchy theory has been widely accepted in the literature.[3] However, a number of questions arise with the nature and kinds of prosodic categories. First, what is the status of each prosodic category? Ladd (1996:237) compares those categories with Halliday's (1967b) prosodic units. We can understand that Selkirk's utterance, intonational phrase, foot, and syllable correspond to Halliday's utterance, tone group, foot, and syllable, respectively. However, the rest of the categories listed in (47), i.e. phonological phrase and prosodic word, do not have their corresponding categories in Halliday's list of prosodic units.

Second, how many kinds of prosodic categories are necessary? To the list (47), Nespor and Vogel (1986) add clitic group as a further prosodic category between phonological phrase and prosodic word. Condravdi (1990) and Kanerva (1990) also claim that further prosodic domains such as minimal phrase and focal phrase are needed. Pierrehumbert and Beckman (1988) and Selkirk and Tateishi (1988, 1991) refer to prosodic phrases in Japanese as Major Phrase and Minor Phrase. However, we should

---

[3] For a critical review of the theory, see Inkelas and Zec (1995:548) among others.

avoid the proliferation of categories if possible. A still more attractive assumption is that there are no prosodic categories per se in the theory. I would like to pursue the idea below.

In the subsequent sections, I will show that the bare mapping theory gives us simple answers to the questions above. Prosodic categories are the derived notion of the strings demarcated by prosodic boundaries. We do not have to posit different prosodic categories between intonational phrase and foot. The theory claims that application of phonological rules is blocked not by prosodic categories but by prosodic boundaries themselves.[4]

Besides the conceptual problems mentioned above, the prosodic hierarchy theory has some empirical problems. First, Gussenhoven and Jacobs (1998:245) show the following examples:[5]

(51) a. $_{I}$(Why don't you sell Jane<u>t</u>, <u>y</u>our Honour?)$_{I}$ [t j] * [tʃ]

| | |

L_I H*L H_I

b. $_{I}$(Why don't you sell Jane<u>t</u> <u>y</u>our honour?)$_{I}$ [t j] *or* [tʃ]

| | |

L_I H*L H_I

The noun phrase *your Honour* in (51a) is vocative, and *your honour* in (51b) is the direct object of the verb *sell*. Assimilation [tʃ] is possible in (51b) but not in (51a). Notice that

---

[4] Akasaka and Tateishi (1999) also suggest that prosodic categories may only be boundary markers.

[5] L shows low pitch, H high pitch, and H*L bitonal tone with pitch accent. See Pierrhumbert and Beckman (1988) and Ladd (1996) for the details of notation.

the whole sentence is included in an intonational phrase in (51a) as well as in (51b), as shown with the intonation contour (L_I H*L H_I). In (51a) there is no boundary between intonational phrases to block assimilation between [t] in *Janet* and [j] in *your*. Thus we cannot explain the phonological difference between (51a) and (51b) in terms of Intonational Phrase.

Gussenhoven and Jacobs also argue that we cannot ascribe the difference in assimilation between (51a) and (51b) to phonological phrase, either. They mention that the phonological phrase tends to correspond to syntactic phrase (p.245). If this is also the case with (51), then NP *Janet* and *your Honour/honour* each make their own phonological phrases. Thus there is a boundary between the phonological phrase containing *Janet* and that containing *your* both in (51a) and (51b).

(52) a. … Janet $_{Ph}$) ($_{Ph}$ your Honour $_{Ph}$)

b. … Janet $_{Ph}$) ($_{Ph}$ your honour $_{Ph}$)

Thus there is no difference in the phonological phrase between (51a) and (51b).

Alternatively, we could postulate a further prosodic constituent, say ?P, between intonational phrase and phonological phrase in order to explain the difference in assimilation between (51a) and (51b). We could argue that (51a) is divided into two ?Ps and that (51b) is included in one ?P, as shown in (53a) and (53b).

(53) a. ($_{?P}$ Why don’t you sell Janet $_{?P}$) ($_{?P}$ your Honour $_{?P}$)

b. ($_{?P}$ Why don’t you sell Janet your honour $_{?P}$)

However, it is ad hoc to propose a new category to explain just one case, and it leads to proliferation of prosodic categories without sufficient grounds.

In fact, Gussenhoven and Jacobs suggest that instead of postulating a further constituent between phonological phrase and intonational phrase, we should define intonational phrase on the basis of syntactic and length criteria. They also show similar examples in (54), where the left periphery of the sentence is in question.

(54) a. $_I$(Tonight your name will be on everybody's lips)$_I$ [t j] *[tʃ]

b. $_I$(Write your name on this envelope)$_I$ [t j] *or* [tʃ]

In (54a) *tonight* is a sentential adverb and *your* is a constituent of the subject. In (54b) *write* is a verb and *your* is a constituent of its direct object. Assimilation between the word final [t] and the word initial [j] may occur in (54b) but not in (54a). If we assume that phonological phrase corresponds to syntactic phrase (XP) as above, we have a phonological phrase boundary just before the NP *your name* in (54b) as well as in (54a).

(55) a. ($_{Ph}$ Tonight $_{Ph}$) ($_{Ph}$ your name $_{Ph}$) …

b. ($_{Ph}$ Write $_{Ph}$) ($_{Ph}$ your name $_{Ph}$) …

Thus we cannot explain why assimilation is possible not in (55a) but in (55b). Again it is ad hoc and is undesirable to postulate a new prosodic category higher than phonological phrase and lower than intonational phrase, as shown in (56).

(56) a. ($_{?P}$ Tonight $_{?P}$) ($_{?P}$ your name $_{?P}$) …

b. ($_{?P}$ Write your name $_{?P}$) …