博士論文

Word prosodic structure in Japanese: A cross-dialectal perspective

（日本語の語レベルの韻律構造：通方言的な視点から）

ポッペ　クレメンス　ピーター
Clemens Pieter Poppe

WORD PROSODIC STRUCTURE IN JAPANESE: A CROSS-DIALECTAL PERSPECTIVE

by

Clemens Pieter Poppe

A dissertation presented to the
Graduate School of Arts and Sciences of
The University of Tokyo
in partial fulfillment of the requirements for the degree
Doctor of Philosophy

2015

Dissertation Committee:
Chair: Dr. Shin-ichi Tanaka (The University of Tokyo)
Dr. Takane Ito (The University of Tokyo)
Dr. Shuichi Yatabe (The University of Tokyo)
Dr. Shin'ichi Tanaka (Kobe University)
Dr. Timothy J. Vance (NINJAL)

Voor mijn moeder

Abstract

In this study, a number of issues in the word prosodic structure of Japanese are investigated from a cross-dialectal perspective. The main goal of the thesis is to account for the accentual and tonal patterns in a number of different dialects by examining the role of word-internal prosodic constituents above the syllable.

The study is couched within Correspondence Theory (McCarthy and Prince 1995), a sub-theory of Optimality Theory (Prince and Smolensky 1993/2004). A new definition of the Indirect Reference Hypothesis (Inkelas 1989; Bermudéz-Otero 2012) is proposed, according to which a phonological constraint may not refer to morpho-syntactic information unless (i) it forces alignment between designated prosodic units and the exponents of designated morphosyntactic nodes or (ii) it forces correspondence between forms that share morphosyntactic or lexical features. In this way, it is possible to restrict direct reference to morphosyntactic and lexical features, while at the same time allowing for analogical effects.

Based on data from a total of six different dialects the following claims are made about tone and prosodic structure.

(i) As for tone, it is shown that H tone is not simply a feature of the postlexical phonology or phonetic implementation. Rather, this tone must be assumed to be active in the lexical phonology. Furthermore, it is demonstrated that in at least two different dialects the interaction of H tone with vowel height cannot be reduced to the interaction of H tone and vowel height with metrical structure, which is an interesting conclusion from a typological point of view.

(ii) As for prosodic structure, it is shown that, besides the prosodic word, the prosodic stem and the foot play an important role in the prosodic grammar. Evidence is presented for the co-existence of trochees and iambs, covert feet, as well as monomoraic feet. Evidence for the prosodic stem comes from accentual alternations in the verbal paradigms of five different dialects. Importantly, however, not only prosodic structure, but also paradigm uniformity is shown to be necessary in order to account for the accentual alternations. What is more, it is demonstrated that a foot-based analysis of accent in verbs has some advantages over a purely tonal analysis. The default accent pattern in all dialects can be formulated in terms of a trochaic foot that is right-aligned with the prosodic stem. The evidence for the co-existence of these two constituents is important for prosodic theory, because both units must be assumed to be dominated by the prosodic word in the Prosodic Hierarchy.

Acknowledgements

First of all, I am very grateful to my supervisor Shin-ichi Tanaka for his constant support and encouragement. I would also like to give explicit thanks to the internal committee members Takane Ito and Shuichi Yatabe from the University of Tokyo, as well as the external members Shin'ichi Tanaka from Kobe University, and Timothy Vance from the National Institute for Japanese Language and Linguistics.

Special thanks are due to Jeroen van de Weijer for his constant support and help. Although he was not an official member of my dissertation committee, his comments helped me a lot.

Many people have helped me by commenting on my work or discussing phonology with me at conferences, at more informal occasions, or through email. I would like to thank all of these people, especially Yuki Asahi, Jeroen Breteler, Marco Fonseca, Carlos Gussenhoven, Daiki Hashimoto, Jeffrey Heinz, Ryota Horikawa, Ayat Hosseini, Larry Hyman, Chiyuki Ito, Junko Ito, Adam Jardine, No-Ju Kim, Haruo Kubozono, Seunghun J. Lee, Armin Mester, Kuniya Nasukawa, Marc van Oostendorp, Thomas Pellard, Gábor Pintér, Hidetoshi Shiraishi, Naoya Watabe, Yasunori Takahashi, Rudy Toet, and Kyoko Yamaguchi.

Some of the findings presented in this thesis were only possible because I had direct access to the speech of native speakers. I would like to express my gratitude to my informants of Tokyo, Matsue, and Ichihara Japanese, and to the people who introduced the speakers of the latter two dialects to me.

Many thanks also to IJsbrandt 't Mannetje, Edwin Michielsen, and Celine Wijsman for sending me material that I could not find in Japan. Special thanks go to Gijs van der Lubbe for the interesting discussions about language in general, and Japanese and Ryukyuan in particular.

Most of all, however, I would like to thank all my friends and my family, especially my mother, Sayuri, and Sophie.

Table of contents

Key to abbreviations and symbols

ABL	= ablatative
COND	= conditional
CONT	= continuous
DAT	= dative
IMP	= imperative
NEG	= negative
NOM	= nominative
NONPAST	= nonpast tense
PAST	= past tense
N1	= first member of a compound
N2	= second member of a compound
Rt	= root
St	= morpho-syntactic stem
m-stem	= morpho-syntactic stem
p-stem	= prosodic stem (phonological stem)
p-word	= prosodic word (phonological word)
p-phrase	= prosodic phrase (phonological phrase)
ʊ	= utterance
ι	= intonational phrase
ϕ	= prosodic phrase (phonological phrase)
ω	= prosodic word (phonological word)
ψ	= prosodic stem (phonological stem)
F (Ft)	= foot
σ	= syllable
μ	= mora
ˈ	= pitch accent
0	= unaccented
H	= H tone
Hˈ	= H toned TBU resulting from an accent
L	= L tone
TBU	= tone-bearing unit

1. Introduction

1.1. Aims of the study

The goal of this thesis is to investigate a number of topics in the word prosodic phonology of Japanese which have not received the attention they deserve in the Japanese and cross-linguistic literature. While all of the topics are related with each other, they can be summed up as follows.

i. Word-internal prosodic constituency
ii. The relative roles of tone and metrical structure in the accent system(s)
iii. The interaction between tone and vowel quality in certain dialects
iv. Accentual alternations in verbal paradigms
v. Micro-variation in word-prosodic systems across Japanese dialects

These topics are related with each other in that they all deal with word prosodic structure. By investigating these topics, I hope to fill some gaps in the existing literature on Japanese word-level prosody. In a cross-dialectal approach, I examine what kind of accentual and/or tonal representations, and what different types of prosodic domains (syllable, foot, prosodic stem, prosodic word) are necessary to account for the word-prosodic systems of Tokyo Japanese, Shizuoka Japanese, the Matsue variety of Izumo Japanese, the Ichihara variety of Boso Japanese, and Kanazawa Japanese.

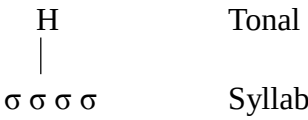
There are at least two reasons for taking a cross-dialectal approach. First of all, non-standard dialects have not been given the attention they deserve in more theoretically oriented studies of Japanese phonology. Second, in many cases, it is simply impossible to choose among different analyses when looking at a single system. By taking into account what consequences the choices we make have in terms of cross-dialectal typology, as well as word-prosodic typology in general, it may become easier to choose among different possible analyses for the different dialects. Thus, at a more general level, this thesis can be seen as a case study of what role phonological microvariation (in this case, variation among closely related dialects; see van Oostendorp 2008) may play in choosing between different possible phonological analyses for one or more varieties of a language.

1.2. Prosodic structure

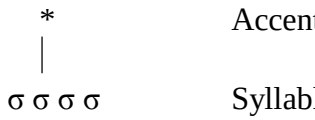
Before we take a closer look at the different topics that are taken up in this thesis, I would like to clarify what I mean with word prosodic structure.

Fox (2002) and Gussenhoven (to appear, a) approach prosodic structure as the organization of phonological representations in terms of multiple interconnected linear

"tiers" and a hierarchy of constituents. Tiers are representational levels that function as a host for particular features that interact with each other. The theory of tiers and their interaction is known as Autosegmental Phonology (Goldsmith 1976; 1990) and has become part of "Basic Grammatical Theory" (Hyman 2012). The idea of separate interacting tiers is exemplified in (1a), where a high tone (H) residing on the tonal tier is linked to the syllable (σ) tier by means of an association line to the second syllable. The two tiers together form a so-called plane (Clements 1985), in this case the tonal plane. A similar representation is possible if abstract accent marks are treated as autosegmental (see Hagberg 1993/2006), as in (1b).

(1) a.


Tonal tier / Syllable tier

b.


Accent tier / Syllable tier

Turning to hierarchical constituency, a distinction can be made between two types of prosodic structure: prosodic constituents that are in general assumed to be the result of the mapping between morpho-syntax and phonology, and genuine phonological units like the mora, the syllable, and the foot. Despite the different nature of the two types of consituents, they are usually assumed to be part of a single hierarchy called the Prosodic Hierarchy (Selkirk 1978; Nespor & Vogel 1986). The consituents of the Prosodic Hierarchy are considered to function as the domains to which phonological rules or constraints refer. The version of the Prosodic Hierarchy in (2), is fairly generally accepted in that it does not contain the more controversial units (e.g. the Clitic Group) that have been proposed.

(2) *Prosodic Hierarchy*

υ	Utterance
ι	Intonational phrase
ϕ	Phonological phrase
ω	Phonological/Prosodic word
F	Foot
σ	Syllable
μ	Mora

As the Prosodic Hierarchy has been claimed to be universal, a lot of research has been devoted to the search for evidence in favor of or against certain levels within the hierarchy across languages. Interestingly, some languages have been argued to require two instantiations of one category. For instance, for Tokyo Japanese a minor and major phonological phrase have been proposed (Martin 1952; McCawley 1968; Kubozono 1988; Selkirk & Tateishi 1988; cf. Pierrehumbert & Beckman's (1988) "accentual phrase" and "intermediate phrase"), whereas the Sino-Tibetan language Limbu has been claimed to exhibit multiple prosodic word domains (Schiering et al. 2010). On the other hand, some languages have been argued to lack certain categories. For instance, according to Schiering et al. (2010), Vietnamese has no prosodic words.

There have been broadly four types of responses to these problems. The first one may be called relativized universality: all levels are universally available for all languages, but not all languages make use of the available categories (Ito & Mester 1992; Green 1997). A solution proposed by Ito and Mester (2007, 2009, 2013) is to restrict the number of categories, and to allow recursion of these basic categories to form "prosodic subcategories". Bickel et al. (2009) and Schiering et al (2010) take what may be called the opposite approach; they propose that prosodic domains such as the prosodic word are not universal, but emergent, the result of which is that a language may lack a certain phrasal domain while at the same time it makes use of multiple slightly different word-level domains (see also Hyman 2008). Finally, in a number of proposals the existence of hierarchical structure in phonological representations is completely denied (Neeleman & van de Koot 2006; Scheer 2004, 2008; Samuels 2011). However, as shown by van Oostendorp (2013), the arguments of these so-called "linearists" are not convincing.[1]

In this thesis, it is assumed that prosodic structure exists. The question then is what prosodic structure looks like, what its functions are within the grammar, and to what extent variation among word-prosodic systems is the result of differences in prosodic structure (as opposed to differences in lexical specifications of pitch accents and other features).

Among the consituents of the Prosodic Hierarchy, a distinction can be made between two types of prosodic structure: prosodic constituents that are the result of the mapping between morphosyntax and phonology, and genuine phonological units of a metrical

[1] Note that the problems related to the universality of metrical (and prosodic) categories are theory-dependent. For instance, in tree-based versions of Metrical Phonology (Liberman & Prince 1977), metrical structure is built in a bottom-up fashion, with only labels that refer to metrical strength attached. Furthermore, in grid-based approaches of Metrical Phonology, (Prince 1983; Halle & Vergnaud 1987; Halle & Idsardi 1995), as well as in Dependency Phonology (Anderson & Ewen 1987), no reference to constituent labels is necessary.

nature. Ito & Mester (2011) distinguish between the two by calling the former "interface categories" and the latter "rhythmic categories", and this distinction will be adopted in this thesis.

In most work in which the Prosodic Hierarchy plays a role it is assumed that the interface and rhythmic categories are part of the same hierarchy. According to Inkelas (1989, 1993) and Downing (1998a/b, 2006), however, it is necessary to distinguish between two types of hierarchies: the Prosodic Hierarchy (consisting of interface constituents) and the Metrical Hierarchy (consisting of rhythmic constituents). This is shown in (3), where the interface categories above the level of the prosodic word are left out.

(3)a. *Prosodic Hierarchy*

…
|
ω Prosodic word
|
St Prosodic stem
|
Rt Prosodic root

b. *Metrical Hierarchy*

F Foot
|
σ Syllable
|
μ Mora

Inkelas (1989, 1993) and Downing (1998a/b, 2006) give a number of different arguments for setting up two different hierarchies. The most important argument is the existence of interface categories below the level of the word in a number of languages.[2] The other arguments given by Downing (2006) can be summarized as follows. In the first place, as their names suggest, interface units and metrical units are defined in totally different ways: interface domains are defined in relation to morpho-syntactic units, while metrical units are defined in terms of what they parse. Moreover, according to Downing (2006), interface units are the domains of phonological processes, whereas metrical units are not (cf. Nespor and Vogel 1986). Downing's (2006) final argument is that the absence of a clear relation between the prosodic word and the stress foot in many languages is a problem from the point of view of a universal theory of Prosodic Hierarchy, the goal of

[2] Evidence for the prosodic stem comes as opposed to the prosodic word comes from Bantu languages (Myers 1987; Hyman 1993; Mutaka 1994; Downing 1999), Salishan languages (Czaykowska-Higgins 1998), and North-Kyungsang Korean (Kim 1997, 2010). According to Downing (2006), evidence for the p-root comes from Kambera (Klamer 1998). However, in Klamer's own analysis (1998) reference is made to a minimally bimoraic, maximally trimoraic (disyllabic) foot (which also is relevant in terms of stress), and not to a prosodic root in the sense of Downing (2006).

which lies in the formalization of universal or near universal patterns observed in the languages of the world. In this thesis, it will be shown that in Japanese the foot and the prosodic stem co-exist, and that they are both dominated by the prosodic word.

Metrical Theory (Liberman 1975; Liberman & Prince 1977) is a theory of relative prominence and linguistic rhythm,[3] and proposes that not only arboreal structure, but also metrical grids are part of the prosodic grammar. Metrical grids are hierarchies of prominence in which rhythmic beats are distributed. The example of the grid in (4) is taken from Hayes (1995).

(4) Constituent structure in a bracketed grid (Hayes 1995: 38)

Bracketed Grid
(x)
(x x)
(x x)(x)
(x x)(x x)(x)
Mississipi mud

The bracketed grid in (4) represents both prominence in terms of grid marks and constituent structure in terms of brackets. In Chapter 3, a critical review of grid-based approaches to Japanese accent is given.

In this thesis, prosodic structure refers to phonological representations that include any of the devices introduced in this section: prosodic features (tone or accent) that reside on their own autosegmental tier, and the prosodic constituents to which they are sensitive.

1.3. Issues in word-prosodic structure in Japanese

In this section, the first four of the five topics listed above are introduced in more detail. The fifth topic, the study of micro-variation in word-prosodic systems across Japanese dialects, speaks for itself: by focusing on dialectal variation observed for the same phenomenon, we may gain a better understanding of the language as a whole.

1.3.1. Word-internal prosodic constituency

Typological studies of Japanese dialects have shown that all Japanese dialects make use of relative levels of pitch to mark prosodic domains, and that not only the positions within

[3] Coleman (1998: 112) points out that Liberman and Prince's (1977) paper on Metrical Phonology was inspired by work on syllable structure by Kahn (1976), who in turn was inspired by formalisms of the Dependency Phonology model proposed by Anderson & Jones (1974).

these domains, but also the domains themselves, differ among dialects (Uwano 1977, 1984b, 1999, 2012b; Igarashi 2012; Kubozono 2012). However, within these studies, apart from discussions on the relative roles of the mora and the syllable (see Sibata 1962, insufficient attention has been paid to the role of prosodic constituency at the level of the word and below. Igarashi (2012, 2014) focuses on prosodic structure across dialects above the level of the word. Kubozono (2012) does discuss differences with respect to the role of "basic prosodic units" (mora, syllable) and the "domain" of accent or H tone assignment across dialects (i.e. "word" vs. "minimal syntactic phrase" (*bunsetsu*)), but the role of the prosodic word and the constituents between the prosodic word and the syllable is unclear. In a footnote Kubozono (2012: 1396) points out that the use of the term "prosodic word" is avoided because "its definition may vary from one dialect to another". This also suggests that finding answers to questions like how prosodic words and other prosodic constituents differ across dialects is what is necessary to gain a better understanding of the differences observed among dialects.

Most studies on word prosodic typology within Japanese are preoccupied with the number of accentual or tonal classes that exist for nouns, while inflected forms like verbs and adjectives are ignored. The reason for this is probably that in most Japanese dialects, inflecting forms have at most two different classes, often referred to as "accented" and "unaccented". However, there is a lot of variation as to where in verbs and adjectives the accents or tones are assigned, and if we are interested in the difference between attested variation and unattested variation, this is certainly an area that needs to be studied. Since verbs and adjectives are morphologically complex forms, they may also exhibit prosodically complex structure. The architecture of this structure is an important topic that has not received the attention it deserves.

A word-level prosodic constituent that has received more attention is the foot. Since the work of Poser (1990), Japanese is considered to be a language that makes use of bimoraic feet. Poser (1990) gives evidence on the basis of eight different phenomena that involve foot structure: (i) hypocoristic formation; (ii) kinship terms; (iii) geisha (bargirl) client names; (iv) rustic girls' names; (v) verbal reduplication; (vi) mimetics; (vii) a secret language used among jazz musicians; (viii) compound accentuation. Kubozono (1995; 1997) revised Poser's (1990) use of the foot to account for compound accentuation, and extended the generalizations to loanword accent (Kubozono 2011). The accumulation of apparent evidence for foot structure prompted Labrune (2012a: 171) to go so far as to state that "the evidence for its relevance is massive and uncontroversial". However, many things are unclear with respect to the details of foot structure in Japanese, and the evidence in favor for it has never been seriously scrutinized. Therefore, in this thesis issues

concerning foot structure are thoroughly investigated.

One of the central issues in this thesis concerns the role of word-internal prosodic structure in the accentual/tonal grammar of Japanese. Some new proposals concerning the prosodic word will be made, but the focus will lie on two word-internal constituents: the foot and the prosodic stem.

1.3.2. The relative roles of tone and metrical structure in the accent system

The role of word-internal prosodic structure is not only interesting from the viewpoint of phonological theory, but also from the viewpoint of prosodic typology. The Japanese language exhibits an enormous amount of variation in terms of prosody (Uwano 1999b; Kubozono 2012).

An often recurring question in studies of language with "restricted tone systems" (Schadeberg 1973; Voorhoeve 1973) in which the distribution of tones shows similarities with that of stress in accentual languages, is how the lexical prominent positions are marked in the lexicon. The word-prosodic system of languages like Tokyo Japanese has been shown to be amenable to both accentual and tonal analyses (see e.g. van der Hulst 2011; Kubozono 2011a, 2012; Uwano 2012b). The status of pitch accent languages is one of the perennial questions in word prosodic typology. Metrical structure is potentially relevant for determining the lexical status of prosodic features: an abstract accent mark, possibly in the form of a foot-head, lexical tone, or a combination of tone and foot structure (as proposed by Gomez-Imbert & Kenstowicz 2000 for the Tukanoan language Barasana; see van der Hulst 2011 for a general discussion of the lexical representation of pitch accent).

Interestingly, the claims from the theoretically oriented literature that Tokyo Japanese pitch patterns are sensitive to foot structure do not figure in the discussions on the place of the language in (word) prosodic typology. The main reason for the lack of reference to foot structure in word prosodic typological studies probably lies in the fact that alternative analyses are often available. A cross-dialectal approach may shed light on the accent vs. tone controversy. If feet are relevant for Tokyo Japanese accent assignment, feet are expected also to play a role in closely related dialects. By comparing a number of closely related accentual systems, evidence may be found in favor of certain underlying or surface representations that are relevant for prosodic typology.

One of the dialects discussed in this thesis, Maisaka Japanese (Chapter 4), has a counting system in which the accent patterns partly depend on whether a word has an even or an odd number of moras or syllables. Such counting suggests that there metrical structure plays a role in the accent system. Therefore, Maisaka Japanese may provide us

with new insights into the nature of foot structure in Japanese.

1.3.3. The interaction between tone and vowel quality

For a number of Japanese dialects, descriptive generalizations have been proposed that suggest there is a difference in phonological weight between syllables with high vowels that are immediately followed by syllables with a non-high vowel (*CI.CA*), and syllables with high vowels immediately followed by a syllable with another high vowel (*CI.CI*) (Uwano 1984a, 2012; Nitta 2001).[4] In these dialects, high tones avoid association with high vowels followed by non-high vowels. According to Odden (2001), such interaction is only attested in Japanese (Odden 2001). Therefore, how to account for this interaction is an issue that is important for cross-linguistic typology and phonological theory.

What is more, de Lacy (2007) – who does not mention the attested patterns in Japanese dialects – argues that the interaction between tone and sonority may only be indirect, i.e. mediated by syllable or foot structure. This makes these dialects interesting test cases in a pursuit of foot structure in Japanese; if we should find metrical structure in Japanese, we may expect to find evidence in these dialects. In this thesis, three such dialects are investigated: Ichihara Japanese (Kinda'ichi 1942; Fujiwara 1972), Matsue Japanese (Hiroto & Ohara 1952; Uwano 1981; Nitta 2001), and Kanazawa Japanese (Uwano & Nitta 1982, 1983; Nitta 1989, 2001; Uwano 1996, 1997).

1.3.4. Accentual alternations in verbal paradigms

Since the advent of output-based approaches to phonology in the 1990s, studies on the pitch patterns of morphologically complex words in Tokyo Japanese have mainly focused on nominal compounds (Kubozono 1993, 1995, 1997; Kubozono et al. 1997; Tanaka 2001, 2005; Ito & Mester 2007, 2010). Accentual alternations in other morphologically complex forms such as verbs, adjectives and noun-particle combinations form an area that is relatively understudied. In more recent years, however, a number of studies on this subject have appeared (Nishiyama 2010; Yamaguchi 2010; Poppe 2012; Oshima 2014). Still, the accent patterns of inflected forms in dialects other than Tokyo Japanese have received almost no attention at all. While some studies mention some of the alternations in verbal and adjectival paradigms and provide partial analyses (Haraguchi's 2001 study of Tsuruoka Japanese; Yoshida and Zamma's 1999 study of Kyoto Japanese), no attempt has been made at a systematic analysis of accentual alternations in inflected forms across Japanese dialects. In Chapter 6 of this thesis an attempt is therefore made at clarifying the principles behind accentual alternations in basic verbal forms across a number of dialects,

[4] C = any consonant, I = high vowel, A = non-high vowel.

examining the relative roles of prosodic domains and paradigmatic relations among inflected forms are examined.

1.4. The choice for an output-based approach

The view of prosodic structure taken in this thesis is heavily based on autosegmental, metrical and prosodic hierarchy-based approaches to phonology. Theories of representations, however, say nothing about how two representations are mapped onto each other. For this, we need a theory of computation. In this thesis, a non-derivational,[5] output-based model of phonological computation is assumed. In this section, the reasons for choosing an output-based approach are discussed.

Before going into the details of this approach, I would like to justify the choice for a non-derivational as opposed to derivational approach which allows intermediate levels of mapping. The reason for this is twofold. First, non-derivational approaches have become mainstream in the field of phonology. Second, and in the light of the topic of this thesis more importantly, a non-derivational approach is inherently more restrictive than a derivational one. As this thesis focuses on representational issues, allowing intermediate levels of representation would multiply the number of possible analyses, making it harder to choose among different possible represenations.[6]

The particular theory adopted in this thesis is Correspondence Theory (McCarthy & Prince 1995), the central idea of which is that relations of identity hold between different levels of representation. Correspondence Theory is couched within Optimality Theory (OT; Prince & Smolensky 1993). In OT, constraints on the well-formedness of output forms determine whether a form is grammatical or not. There are broadly two types of constraints: "markedness" constraints and "faithfulness" constraints. Markedness constraints militate against "marked" representations (e.g. ONSET: "A syllable should have an onset"), and are grounded in phonetics and/or typology. The job of faithfulness constraints is to check whether two representations are faithfully mapped onto each other. Examples of this are input-output (IO) faithfulness, where input forms – which in many cases are identical to the underlying forms proposed in rule-based models – are evaluated in terms of their correspondence with output forms (surface forms), and output-output

[5] Non-derivational in the sense that ordering of mechanisms of repair is not allowed (van Oostendorp & Hermans 1999: 9).

[6] According to some researchers, intermediate levels of representations are necessary to account for opacity in a unified way and thus are a strength of derivational approaches (Vaux 2008). However, I am not convinced by the arguments in favor of derivational approaches to opacity. In many cases they involve morphological irregularities. What is more, as shown by Baković (2011), the argument that in an approach based on ordered rules opacity can be accounted for in a unified way does not hold. Finally, in non-derivational approaches, different types of solutions are available (see again Baković 2011).

(OO) faithfulness, which is about the correspondence between two (or more) independently attested surface forms (see Kager 1999 for a discussion of different types of correspondence). While in classical OT it is assumed that the set of constraints is innate, there are also versions of OT in which it is assumed that constraints are learned on the basis of input data (e.g. Boersma 1998; van de Weijer and Sloos 2013).

Another output-oriented approach to phonology that in recent years has been gaining popularity is known as "usage-based" or "cognitive" phonology (Bybee 2001; Nesset 2008). In usage-based models, the grammar consists of schemas that are abstracted from forms that are stored in the lexicon. Such schemas not only licence the output forms from which they are abstracted, but also function as templates on the basis of which new forms are created. What is more, schemas may be updated, strengthened or faded out depending on new forms a speaker encounters.

An important difference betweent OT and schema-based models lies in what the two models focus on. Whereas OT focuses on typological variation among grammatical systems in terms of rerankable constraints, schema-based models focus on the cognitive mechanisms that regulate the selection of language-specific schemas (see Nesset 2008). As the focus of OT lies on typological variation, it is well suited for a study on phonological variation like the present one. Differences among similar dialects can be easily expressed by means of a set of rerankable constraints, and less so in terms of language-specific schemas.

Another reason for adopting OT is that most theoretical studies on metrical and prosodic structure have been conducted within this framework (Kubozono 1995, 1997; Ito & Mester 2007, 2009, 2013). It is much easier to compare different claims about representations if at least the way in which the output forms are computed is the same across the analyses under comparison.

As noted by Scheer (2010), among others, OT as such is not a theory of phonology or even language, but a theory of constraint interaction, i.e. computation. This means that the nature of representations is as much an issue within OT as it is in any other approach (see van Oostendorp and van de Weijer 2005). The reason for this is that a theory of representations is in principle independent of how the representations are computed. Therefore, I would like to stress that in principle few of the issues regarding representations in this thesis hinge on the fact that phonological computations are modeled in an OT-fashion. At certain points within the thesis, however, I will point out that some issues regarding representations do in fact depend on the theory of computation that one assumes.

1.5. Indirect reference

In the literature, many different views exist on the relation between phonology and morpho-syntax (Scheer 2011). The general theory of grammar that I assume is that of Parallel Architecture (Jackendoff 1997), in which phonological, morpho-syntactic and semantic representations exist in parallel. In such a model, it is natural to adopt the Indirect Reference Hypothesis (Inkelas 1989), according to which phonological rules or constraints may only refer to phonological domains as opposed to morpho-syntactic ones.

(5) Indirect Reference Hypothesis (Inkelas 1989: 10)
Phonological rules refer to only prosodic constituency.

The Indirect Reference Hypothesis sets boundaries on what entities phonological constraints may refer to. The hypothesis is reformulated in terms of constraints by Bermúdez-Otero (2012) as in (6).

(6) Indirect Reference Hypothesis (constraint-based; Bermúdez-Otero 2012: 77)
A phonological constraint may not refer to syntactic, morphological, or lexical information unless to require alignment between designated prosodic units and the exponents of designated syntactic (word-syntactic or phrase-syntactic) nodes.

This can be said to be the strongest possible version of the hypothesis. Bermúdez-Otero (2012) is able to adopt this strong version of the hypothesis because he assumes Stratal OT (Bermúdez-Otero 1999; Kiparsky 2000) in which three levels of derivation are assumed. In this thesis, however, I assume Correspondence Theory with both IO and OO faithfulness – the latter of which is needed to model the cyclic effects that are accounted for by allowing different levels in the stratal model – I propose to revise the definition as in (7).

(7) Indirect Reference Hypothesis (constraint-based, revised)
A phonological constraint may not refer to morpho-syntactic or lexical information unless (i) to require alignment between designated prosodic units and the exponents of designated morpho-syntactic nodes or (ii) to require correspondence between forms that share morpho-syntactic or lexical features.

Thus, I assume that OO constraints may refer to morphological entities like root or affix. Importantly, however, what is compared in the case of OO faithfulness are two

phonological representations that share morphological material. The sharing of morphological material is assumed to be visible at the morpho-syntactic level only, as in (8).

(8)

Morpho-syntax		Phonology
[[Root$_i$]Affix$_\alpha$]	⇆	[ ]
↓↑		↓↑
[[Root$_i$]Affix$_\beta$]	⇆	[ ]

Thus, an OO-correspondence constraint compares the phonological structure of two forms sharing the same root, for example, but the root itself is not visible in the structures that are compared. In this sense, OO faithfulness observes the Indirect Reference Hypothesis.[7]

For some phenomena possible alternative analyses that do not conform to the Indirect Reference Hypothesis are considered. Analyses in which direct reference to morpho-syntactic structure is allowed are explicitly presented as such.

1.6. Data

The data that form the basis of the analyses developed in this thesis comes from different types of sources. Most of the data from Tokyo Japanese comes from papers written in English by several scholars, accent dictionaries (NHK 1998; Kinda'ichi and Akinaga 2001), as well as from checks with native speakers. The data from the other dialects, on the other hand, comes from dialectological material and papers written by Japanese-style accentologists, and for some of the dialects (Matsue, Ichihara) from personally conducted fieldwork.

1.7. Romanization

In this thesis, I adopt a slightly revised Kunrei romanization in the examples of linguistic forms, and the Revised Hepburn system of romanization otherwise. The difference with the original Kunrei system is that long vowels are written as two vowels instead of using a circumflex.

[7] It may be necessary to allow constraints on prosodic subcategorization (Inkelas 1989) that specifiy the prosodic properties of the base to which morphemes attach. Technically, these are morphological constraints that refer to phonological information. See Trommer (2011: 64), who calls the combination of the Indirect Reference Hypothesis and prosodic subcategorization the "Extended Indirect Reference Hypothesis".

1.8. Organization of the thesis

The structure of this thesis is as follows. In Chapter 2, an overview of prosodic structure in Japanese is given. This overview contains both a discussion of the nature of Japanese pitch accent, as well as an overview of the word-level prosodic constituents that have been claimed to play a role in Japanese.

In Chapter 3, metrical analyses and a purely tonal analysis of Tokyo Japanese accent are compared. After critically reviewing the existing metrical approaches, a purely tonal approach is developed. I show that although default loanword accent and lexical compound accent can be accounted in a purely tonal approach, without reference to foot structure, it is necessary to look for explanations outside phonology (morphology, lexical frequency) to account for accent in phrasal compounds and unaccentedness. It is argued that although this may speak in favor of the metrical approach, there are many unresolved problems, and additional evidence is necessary.

In Chapter 4, 5, and 6, I present such additional evidence. In Chapter 4, I demonstrate that in the Maisaka dialect, which has an odd-even counting system, only a metrical approach can account for the accent patterns in a convincing way. Maisaka Japanese also provides us with answers to a number of questions concerning Japanese foot structure that have remained unanswered in the literature.

Chapter 5 deals with three dialects in which tone and vowel height interact. For all three dialects, it is shown that constraints that directly relate tone and vowel height with each other are necessary. Furthermore, evidence is found for the interaction between tone and foot structure.

In Chapter 6, accentual alternations in the verbal paradigm of five different dialects (Tokyo, Shizuoka, Ichihara, Matsue, and Maisaka Japanese) are analyzed. It is shown that both the prosodic stem as well as paradigm uniformity play a crucial role in the phonology. Furthermore, additional evidence for foot structure is presented.

Finally, the main findings of this thesis and directions for future research are summarized in the conclusion.

2. Word-prosodic structure in Japanese: an overview

In order to set the stage for the issues to be treated in this thesis, in this chapter an overview of word-level prosody in Japanese is presented. The chapter is divided into two parts. In the first part, different views on the phonological representation of Japanese pitch accent are introduced and critically examined. In the second part, the evidence for word-level prosodic units that have been proposed for Japanese is discussed.

2.1. Lexical and surface representations of Japanese pitch accent

Many introductions to Japanese pitch accent have appeared in the literature (Haraguchi 1988, 1999; Kubozono 2008, 2011, 2012), and the topic has been debated within studies of a more typological nature (Beckman 1986; Fox 2000; Hyman 2006, 2009; van der Hulst 2011). Within these studies, two related questions play a central role. The first one is whether apart from tone and stress languages, a separate category of pitch accent languages exists. Hyman (2006, 2009) has pointed out that a property-based approach to word-prosodic systems is preferred over an approach in which languages are simply labeled as being a tone language, a stress language, or a pitch-accent language. He argues that there is no single category of pitch-accent languages, and no canonical pitch accent system. According to Hyman, we should not ask whether a language is "a pitch accent language" or a "tone language", but we should ask what langugaes do with tone or with metrical structure (Hyman 2009: 234). On the other hand, van der Hulst (2011) makes a case for recognizing a class of accentual languages that includes languages with pitch accent.

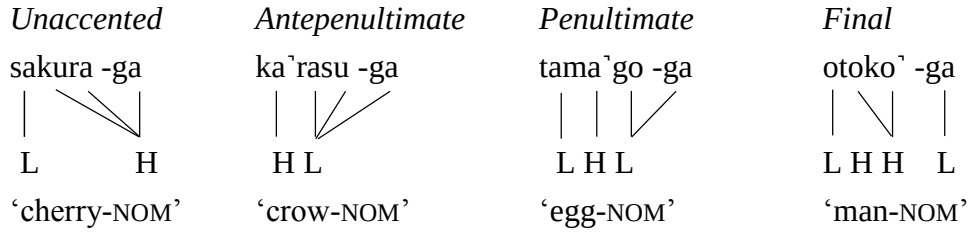
The second question is what the lexical and surface representations of the pitch accent are. It is this question that is taken up in this section, where the lexical and surface representations that have been proposed for Japanese pitch accent are introduced, and the motivations for these representations are critically evaluated. We will start with a discussion of the representations that have been proposed in the autosegmental and autosegmental-metrical approaches. Here the focus will lie on the "accentual" vs. "tonal" analyses of the cross-linguistic literature. Afterwards, the dominant view in the Japanese school of accentology is introduced and its merits and demerits will be scrutinized.

2.1.1. Representations of Japanese pitch accent

The first comprehensive study of pitch accent in the framework of generative phonology was that by McCawley (1968). Okada (1975) extended the approach taken by McCawley to a number of dialects other than Tokyo Japanese. The first comprehensive study in terms of autosegmental theory was conducted by Haraguchi (1977). Haraguchi (1977) analyzes

a large number of dialects based on data from the Japanese accentological literature. As pointed out by Shibatani (1979), an important achievement of his work is that he typologizes the different dialects in terms of their tonal melodies and the rules by which they are assigned to words, rather than in terms of underlying accents. Representations of trimoraic words in Tokyo Japanese that are representative of his work are given in (1). Rather than by a star " * " as in Haraguchi (1977), accent marks are indicated by a superscript hook " ˺ " that is placed immediately after the accented mora/syllable.[8] The accent is phonetically manifested by a sharp drop in pitch, which in the autosegmental framework is the result of a drop from H to L tone.

(1) Autosegmental analysis of Tokyo Japanese accent (based on Haraguchi 1977, 1999)



In Haraguchi's analysis, the H and L tones of the basic HL pattern that is assigned to the accented syllable (or, in case of unaccented words, to the final syllable) spread to the left and right edges respectively, after which a dissimilation rule ("initial lowering") introduces a L tone that replaces the initial H tone in words in which the initial syllable is not accented.

In the Autosegmental-Metrical (AM) model (Ladd 1996), Pierrehumbert and Beckman (1988) argue against a classical autosegmental analysis in which tones spread, and instead argue for an analysis in which tones are assigned to nodes in a prosodic tree and in which tone bearing units (TBUs) may be underspecified for tone even on the surface level. Tonal contours are the result of interpolation between local tonal events, so-called F_0-targets (Bruce 1977). The surface representations of Tokyo Japanese are presented in (2).

[8] This hook is used as the mark indicating pitch accent throughout this thesis. In cases in which the tonal surface representations are relevant, the hook will in many cases also be added to tonal representations in order to distinguish between H-toned moras that are accented (or have an underlying H tone) and those that are not. The use of this diacritic is for the sake of clarity rather than based on the idea that Japanese pitch accent should be represented by a diacritic rather than by lexical tone.

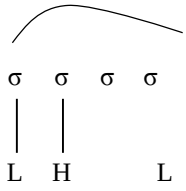
(2) Autosegmental-metrical analysis of Tokyo Japanese[9]

Unaccented	*Antepenultimate*	*Penultimate*	*Final*
sakura -ga	karasu -ga	tamago -ga	otoko -ga
L%H% L%	L%HL L%	L%HL L%	L%H%HL L%
'cherry-NOM'	'crow-NOM'	'egg-NOM'	'man-NOM'

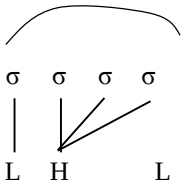
Pierrehumbert and Beckman (1988) argue that their underspecified representations better reflect the actual F_0-curves. However, Sugahara (2003) presents evidence for two types of speakers: those who make use of interpolation, and those who spread H tones to the right in unaccented words. The difference between the two types of speakers for unaccented words of four syllables is graphically illustrated in (3).

(3) Both LHH(H) and LHØ(Ø) are attested (i.e. two types of speakers; Sugahara 2003)

a. Interpolation



b. Spreading



For the underlying representations of pitch accent in Tokyo Japanese, at least the following four types of representations have been proposed: the underlying accent (plus tone-melody) analysis (Haraguchi 1977), the H-only analysis (Poser 1984), the L-only analysis (Backley and Nasukawa 2013), and the AM-analysis (Pierrehumbert and Beckman 1988). As the star is part of the generally accepted representation of pitch accents in the AM-model, here the star is not replaced by the hook " ˀ ".

(4) Different analyses of the underlying specification of accent in Tokyo Japanese

Example	*Accent*	*H-only*	*L-only*	*H*L*
karasu(ga)	μˀμμ(μ)	HØØ(Ø)	ØLØ(Ø)	H*LØ(Ø)
tamago(ga)	μμˀμ(μ)	ØHØ(Ø)	ØØL(Ø)	ØH*L(Ø)
otoko(ga)	μμμˀ(μ)	ØØH(Ø)	ØØØ(L)	ØØH*(L)
sakura(ga)	μμμ(μ)	ØØØ(Ø)	ØØØ(Ø)	ØØØ(Ø)
	(Haraguchi 1977)	(Poser 1984)	(B&N 2013)	(P&B 1988)

[9] Note that the trailing L tone of the HL pitch accent is not associated in the analysis of Pierrehumbert & Beckman (1988). Gussenhoven (to appear, b), however, presents evidence in favor of association of this tone. See Chapter 3 for a revised autosegmental-metrical representation.

Several scholars have pointed out that while accentual analyses of pitch accent languages or restricted tone languages are possible, it is always possible to present a tonal alternative (Pulleyblank (1983/1986; Poser 1984; Hyman 2006, 2009). In Chapter 3, it will be shown that in the case of Tokyo Japanese, it does not really matter which of the input representations we assume, as long as the lexical phonology has access to both the H and the L tone.

While it cannot be denied that tonal analyses of pitch accent languages are always possible, van der Hulst (2011) points out that even if we adopt Hyman's definition of a tone language as a language "in which an indication of pitch enters into the lexical realization of at least some morphemes" (2006: 229), it may still be the case that we can distinguish between accentual languages on the one hand, and tone languages on the other. According to van der Hulst (2011), if we accept that accents are not necessarily obligatory, and not necessarily culminative, both languages with atonal or unaccented words (such as Tokyo Japanese), as well as languages with multiple non-adjacent H tones in which we need only a single tonal value (such as Kirundi; Goldsmith & Sabimana 1989) could be analyzed as accentual.[10]

While one may get the impression that the above discussion is mainly one of terminology, I would like to point out that if we agree that differences among languages in their phonological representations and the relation between phonology and phonetic implementation are important, the issue of tone vs. accent is of both typological and theoretical importance.

Apart from the importance of the difference between phonological representations and phonetic implementation, van der Hulst (2011) also discusses another important point that deserves more attention than it has typically received in typological discussions of pitch accent: the rules or constraints that refer to accent and/or tone. Regrettably, in most typological studies, representational issues are discussed as if not related to computational issues. That is, an examination of the nature of rules or constraints that must be assumed to refer to either accent marks, metrical structure, or tones, in most cases is not conducted.

[10] Hyman (2006) explicitly rejects the possible analysis in which the multiple H tones in Kirundi are analyzed in terms of what he calls "stress accent". Rather, he argues that we are dealing with an instance of another function of metrical structure, namely "a well-known counting function used to locate or assign phonological properties within a string" (Hyman 2006: 242). Thus, he takes the position that tonal languages simply may have metrical structure to which tones are sensitive, as in Leben (1996). However, the point is that if we follow van der Hulst and "push the use of accent to its limits" (van der Hulst 2011: 1017), we can analyze words with multiple H tones as having multiple accents. In this approach, then, stress accent is simply a case of accent that is both obligatory and culminative. (Whether in such an approach an analysis in terms of multiple accents would be chosen for Kirundi, is a separate question).

It was Pulleyblank (1983/1986) who attacked the diacritic approach to pitch accent not only by appealing to the fact that diacritics are to be avoided if they can be replaced by non-diacritic representations – which in the case of pitch accent, in many cases is a H tone – but also by alluding to the fact that if diacritics and tones coexist in the phonological grammar, rules may refer to both diacritics and tones, or to both at the same time. Indeed, in many of the analyses proposed by Haraguchi (1977), rules that refer to diacritic "stars" co-exist with rules that refer to tones.

Summarizing, the choice for an accentual as opposed to a tonal analysis is at least partly a matter of theoretical assumptions. In this thesis, it will be assumed that TBUs may be unspecified for tone even in surface forms, as in the AM-model. Comparisons between the diacritic and the tonal approach will be made where relevant. It will be shown that in a two-level non-derivational analysis, in certain cases we need some kind of diacritic to distinguish accentual tones from non-accentual (boundary) tones. However, the focus will lie on the relative roles of tone and metrical structure rather than the distinction between accent and tone.

2.1.2. Pitch accent in Japanese accentology

Among the indigenous approaches to Japanese accent, the approach in terms of "accent kernels" has proved to be the most successful one. In his historical summary of approaches to Japanese accent, Poser (1984) attributes the accentual analysis of Japanese to Miyata (1927, 1928), who argued that the surface pitch patterns can be abstracted away from by positing an accent at the point before which the pitch drops. Arisaka (1941) refined the accentual view by making a distinction between words with an "underlying" final accent and unaccented words, which have highly similar pitch patterns on the surface in isolation (because the distinct fall in pitch is not (completely) realized word-finally), but are realized differently when followed by for example the nominative particle. Finally, the most important contribution of Hattori (1961, 1967), according to Poser (1984), was that he distinguished between the "accent kernel" and the "prosodeme", which corresponds to an "accentual unit" (Uwano 2012b) that forms the domain of phonetic interpretation. Uwano (2012b) states that the definition of "accent kernel" in its present use is due to Hattori (1954/1955), "who defined the kernel as the High (and tense) portion which immediately precedes a Low (and weak) portion".

While in Hattori's approach the phrase-initial rise is still attributed to the word, Kawakami (1961) presents evidence for the phrasal nature of the rise. What is more, Kawakami (1961) argues in favor of a dynamic approach to accent, rather than one in terms of static H and L tones. This approach was adopted by Uwano (1977, 1999, 2012),

who has applied the dynamic approach in an enormous number of descriptive and historical analyses dialects of Japan and the Ryukyuan language, and developed a typology of accent systems (Uwano 1999b, 2012).

In this approach, the location of the accent and the direction of pitch movement are specified in both the lexical and the surface representation, without the need for any static tones (Kawakami 1961; Uwano 1977, 1999b, 2012).[11] As we can see from the data in (5), both the pitch accent as well as the phrase-initial rise are represented by brackets. These brackets are not as abstract as the diacritics used to indicate accent in the autosegmental literature. Rather, they are mnemonic for the direction of pitch movement: "]", the "lowering kernel", indicates a drop in pitch, and "[" indicates a rise in pitch.

(5) Lexical and surface represention of accent in the Japanese school of accentology

a. Lexical representation:	ka]rasu	'crow'
	koko]ro	'heart'
	otoko]	'man'
	sakana	'fish'
b. Surface representation:	[ka]rasu	'crow'
	ko[ko]ro	'heart'
	o[toko]	'man'
	sa[kana	'fish'

The most important difference between the Japanese accentological approach and the AM approach is that the tonal specifications in the former are thought to be of a dynamic nature (i.e. in terms of tonal movement) rather than of a static nature (i.e. in terms of H and L tonal targets) as in the AM model.

The differences between the dynamic and static approaches are summarized by Uwano (2012b). One of the arguments in favor of the dynamic approach is that not all drops in pitch are from a H tone to a L tone. Thus, in an utterance that contains several accents, as in the example taken from Uwano (2012b: 11) in (6), it is not the case that all syllables bearing an accent "]" are realized on the same H pitch. As pointed out by Uwano (2012b), this follows naturally from the dynamic representation, while in the approach based on static representations, it is necessary to evoke a pitch-lowering process like downstep or catathesis.

[11] Uwano (2012b) notes that although the notion of "accent kernel" (*akusento-kaku*) originates in the work of Hattori (1954), the idea of a dynamic approach to accent kernels (and phrasal tones) was put forth by Kawakami (1961).

(6) ‘A man with black glasses ran away’

a. kuro]i me]ganeno otoko]ga ni]geta lexical representation (dynamic)

b. {ku[ro]i me]ganeno otoko]ga ni]geta} surface representation (dynamic)

According to Uwano (2012b), catathesis or downstep “is not a matter of pitch range reduction of fall, but simply the absence of rise”. However, it is doubtful whether this view of downstep can be applied to the numerous cases of downstep that have been reported for both tonal and intonational languages in the literature. There is evidence for downstep from languages spoken in different parts of the world, especially in sub-Saharan Africa (Connell 2011). If downstep is necessary to account for other languages in the world, there is no reason why we could not apply it to Japanese. What is more, as recognized by Uwano (2012b) himself, before a pitch drop resulting from an accent, “a small rise in pitch” can be observed. According to Uwano (2012b:11), this rise is not obligatory and should be interpreted as “a mechanical anticipation of the following fall in pitch”. However, we could also argue that the rise is the result of the phonetic implementation of an associated H tone, and that in some cases the phonetic implementation is not complete. In any case, the occurrence of multiple accents and the different phonetic values of phonological H tones forms no evidence against the static approach.

While the Japanese accentological approach has proved to be of enormous value for the study of Japanese prosody, the approach is less useful from a cross-linguistic point of view. Also, adopting the kernel-based approach would make a theoretical investigation from the perspective of modern phonological theory and typology very difficult if not impossible.[12] Therefore, in this thesis tone will be represented in terms of H and L targets.

2.2. Prosodic units in Japanese

We will start our overview of prosodic units in Japanese with the rhythmic categories of the mora, the syllable, the foot, and the colon. Because the foot plays a major role in the analyses in the following chapters, it is discussed more extensively than the other categories.

2.2.1. Mora

Among the three rhythmic units that have been proposed for Japanese, the mora is the least controversial one. The mora is described by McCawley (1978:114) as “something of which a long syllable consists of two and a short syllable of one”. This immediately

[12] See Kisseberth (2001) for a discussion of some problems of the Japanese accentological approach.

shows that it is difficult to discuss the mora without also making reference to the syllable. Interestingly, however, Japanese linguists have not always felt the need to refer to what presently are known as syllables. This is probably related to the psychological salience that the mora has (and the syllable lacks) for speakers of most varieties of Japanese.[13]

The evidence for the mora is uncontroversial, and excellent summaries of the evidence exist (Vance 1987, 2008; Kubozono 1999; Labrune 2012a). In (7), the different kinds of evidence discussed in these studies are summarized.

(7) The role of the mora in Japanese phonology
a. Temporal unit (Bloch 1950; Han 1962a, 1994; Homma 1981; cf. Beckman 1982).
b. Counting unit and tone-bearing unit in the accent system (McCawley 1968)
c. Word formation (morphological blending; Kubozono 1990)
d. Unit of segmentation (Kubozono 1985, 1995, 1996; Otake et al. 1993, 1996)
e. Language games/Secret jazz language (Haraguchi 1991; Ito et al. 1996)
f. Text setting in traditional verse (*haiku*, *tanka*) and music (Vance 1987)
g. Orthography (see Labrune 2012a)

In the cross-linguistic literature, the role of the mora as a temporal unit was already pointed out by Trubetzkoy (1969[1939]), who distinguished between mora languages and syllable languages. While the phonetic details of mora timing are still debated, the psychological reality of the mora as a temporal unit is evident (see Warner & Arai 2001 for a state-of-the-field summary). Moras in Japanese are not only counted for rhythmic purposes, but also for the purpose of accentuation. According to McCawley (1968), the default accent in Tokyo Japanese is placed on the syllable that contains the third mora from the right (see Chapter 3).

The mora has also been shown to play a role in morphological blending (Kubozono 1990), in speech segmentation (Kubozono 1985, 1995, 1996; Otake et al. 1993, 1996), as well as in languages games (Haraguchi 1991) and a secret jazz language (Ito et al. 1996; see below).

The use of the mora is especially salient in text setting in poetry and music. For instance, a haiku consists of three lines of five, seven, and again five moras, respectively. The role of the mora in Japanese orthography is also evident; in the *kana* "syllabaries", the so-called "moraic consonants" and the second part of a long vowel or diphthong are all written with a single character.

[13] In some dialects, such as Kagoshima Japanese and Tohoku Japanese, there is no clear evidence for moras. Dialects like this are known as "syllabeme dialects" in the Japanese literature (Sibata 1962).

In Japanese phonology, a distinction can be made between "regular moras" and what in the Japanese literature are called "special moras" (*tokushu haku/tokushu mōra*; see Vance 1987, Labrune 2012a). In Japanese phonology, the special moras are usually represented as in (8). Examples of words containing these moras are given in (9). Hyphens indicate mora boundaries.

(8)
/Q/:	moraic obstruent; first part of a geminate
/N/:	moraic nasal
/R/:	second half of a long vowel
/J/:	second part of a rising diphthong ending in /i/.

(9)
/Q/:	ji-Q-ke-N (ji-k-ke-n)	'experiment'
/N/:	o-N-i-N-ro-N (o-n-i-n-ro-n)	'phonology'
/R/:	ho-R-ge-N (ho-o-ge-n)	'dialect'
/J/:	ga-Q-ka-J (ga-k-ka-i)	'conference'

In functional terms, special moras are special for a number of reasons. To start with, they are avoided by accentual and phrasal H tones. An accent on a special mora is not allowed in general. Furthermore, when in Tokyo Japanese the second mora of the phrase is a special mora, the phrase starts with a [HH.. (in the case of /J/, /R/, and /N/) or [LL.. pattern in the case of /Q/] rather than the default [LH attern (see below in 2.2.2.).

Both within and across dialects, not all special moras show the same behavior. Uwano (1984) proposes an "independency hierarchy" (Tanaka 2000) which orders the special moras in terms of their ability to bear accent, which may differ per dialect. The independency hierarchy proposed by Uwano (1984) based on the behavior of special moras across dialects is as shown in (10).

(10) /J/ > /R/ > /N/ > /Q/

Kubozono (1999) and Tanaka (2008) add to the evidence of the hierarchy in (10), and seek to explain the hierarchy in terms of sonority. According to Kubozono (1999), the special moras that behave as a vowel, i.e. /J/ and /R/, are more independent than those that behave as a consonant, i.e. /N/ and /Q/, due to their higher sonority. Apart from the degree of independence, Kubozono (1999) proposes another way of measuring the degree of 'specialness' of special moras, namely in terms of their stability. This concept is also based on sonority; the greater the distance in sonority between the head nucleus and the

following special mora, the greater the stability of the special mora. However, as pointed out by Nasu (2009), if the difference in sonority were the relevant criterion to measure the stability of moras, we would expect the "stability hierarchy" to be as in (11a). However, Nasu (2009) points out that this is not the case, the actual stability hierarchy based on an analysis of speech errors by Kubozono (1999) being the one in (11b). Note that both hierarchies are different from that in (10).

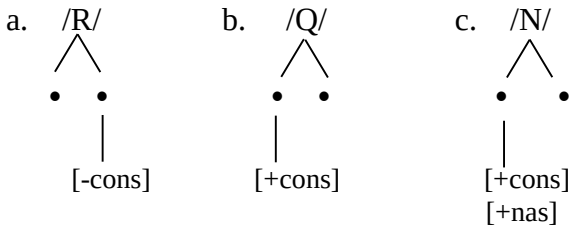
(11)a. /Q/ > /N/ > /J/ > /R/
b. /N/ > /J/ > /R/ > /Q/

Nasu (2009) makes an interesting proposal concerning the definition of moraic stability. He points out that while /R/ and /Q/ can be defined purely structurally in that they consist of simply a vowel (V) and a consonant (C) respectively, /J/ and /N/ have the features [+high] and [+consonantal, +sonorant, + nasal] respectively. Nasu (2009) strengthens his analysis by providing data from abbreviated loanword compounds that show the frequency with which special moras are not parsed into clipped words follows the stability hierarchy in (11b).

Tanaka (2008: 69) proposes a similar stability hiearchy on the basis of the number of musical notes assigned to special moras in Japanese songs. Tanaka's results suggest that the stability of /J/ is slightly higher than that of /N/, but apart from that, his findings match those of Nasu (2009). Furthermore, Tanaka (2008) points out that the interaction of the independency hierarchy and the stability hierarchy can be seen as the cause behind the slightly different findings of different scholars.

The approach to special in Labrune (2012a/b) resembles that of Nasu (2009) in that the moras are special in structural terms. That is, they are incomplete. This is shown in (12): /R/ (and /J/, which is left out) lacks an onset, whereas /Q/ and /N/ are characterized by the absence of a vocalic nucleus. Thus, Labrune (2012a/b) defines special moras as "deficient prosodemes" that contain an empty position.

(12) Japanese mora structure (Labrune 2012a/b)



Apart from the traditional special moras, Labrune (2012a/b) proposes to treat devoiced vowels and epenthetic vowels, which are also considered to be prosodically "weak", in the same way. The representations she proposes for these two types of moras are given in (13). Note that the representation of moras that include epenthetic vowels implies that the derivational history is visible in the representation.

(13) Japanese mora structure (Labrune 2012a/b)

a. devoiced vowel b. epenthetic vowel

Labrune's (2012a/b) "hierarchy of Japanese prosodemes" is based on the mora's ability to bear accent. Labrune claims that this relative capacity is conditioned by the mora's "intrinsic phonetic prominence and acoustic energy, as well as by the number of filled structural positions it contains" (Labrune 2012a: 170). Labrune (2012a: 170) proposes a basic hierarchy of CV >> •V >> C• (where • stands for an empty position), which is further subclassified as follows.

(14) C*a* > C*o* > C*e* > C*u* > C*i* > •*a* > •*o* > •*e* > •*u* > { •*i* > R > N > Q ; $CV_{epenthetic}$; $CV_{devoiced}$ }

Note that in the hierarchy in (14), vowel height is considered to be related to be a factor in the position on the prosodic hierarchy. While this distinction is only marginally relevant in Tokyo Japanese (Tanaka 2008), there are a number of dialects in Japanese in which high vowels and non-high vowels show clear differences in H tone-bearing potential. In Chapter 5 we will take a look at how such dialects may be analyzed.

2.2.2. Syllable

While work by Hattori (1954/55), Martin (1967), McCawley (1968) and Kubozono (2003) has shown the need to recognize syllables besides moras in Japanese, lately the syllable has come under attack. For example, Uwano (2003) claims that the syllable is not necessary to account for accentuation in dialects that exhibit moras, and Labrune (2012a/b) argues that syllables are non-existent in Tokyo Japanese.

To illustrate the difference between syllables and moras in Japanese, let us take a look

at how syllable parsing differs from moraic parsing. The words in (15) are repeated from those in (9).

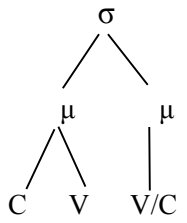
(15)	/Q/:	[ji-Q-]$_{\sigma}$[ke-N]$_{\sigma}$	'experiment'
	/N/:	[o-N-]$_{\sigma}$[i-N-]$_{\sigma}$[ro-N]$_{\sigma}$	'phonology'
	/R/:	[ho-R-]$_{\sigma}$[ge-N]$_{\sigma}$	'dialect'
	/J/:	[ga-Q-]$_{\sigma}$[ka-J]$_{\sigma}$	'conference'

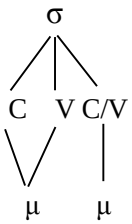
As can be seen in (15), a sequence of a regular mora and a special mora forms one syllable. In words with only regular moras, the number of syllables equals the number of moras. In other words, the difference between moras and syllables is only relevant in words with special moras.

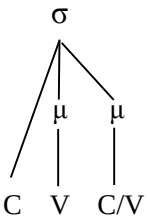
Before we take a look at the evidence in favor of syllables, it should be pointed out that there is no consensus on what the Japanese syllable looks like even among proponents of the syllable. Labrune (2012a: 149) gives four different types of syllable structure that have been proposed for Japanese in the literature. Three of these are presented in (16).[14] Furthermore, I added the proposal by Inaba (1998) given in (16d). The structure in (16a), which was proposed by Kubozono (1995), resembles the original mora model of Hyman (1985). The proposal by Hyman (1985) can be seen as a moraic version of the "Body-Core" approach (McCarthy 1979). The proposal in (16b), which was put forward by Terao (1992) and Kubozono (1994, 1998), seems to be based on the idea of a distinct "prosodic plane" for moras (see Dresher and van der Hulst 1998). The representation in (16c) is the familiar representation proposed by Hayes (1989) within moraic theory. The structure in (16d) is proposed by Inaba (1998), who adopts the X-bar approach to syllable structure (Levin 1985) in which "C" stands for "complement", "N" for "nucleus", and "S" for specifier. In this approach, moras are not subconstituents of syllables.

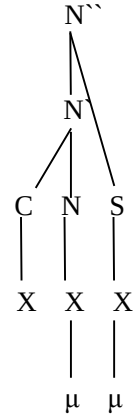
[14] As it is not clear what role the mora plays in the model proposed by Haraguchi (2003), his approach is not discussed here. His analysis is basically the same as that in (16a), with the difference that the lefthand mora is replaced by the "core", and the righthand mora by the "coda".

(16) Proposals for Japanese syllable structure

a. 

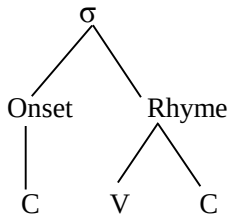
b. 

c. 

d. 

As Labrune (2012a/b) stresses, it is striking that in the proposals for syllable structure in Japanese, the syllabic sub-constituent of the rhyme (17), which plays a major role in classic approaches to syllable structure, plays no role whatsoever. Labrune (2012a) suggests that this casts doubt on the existence of the syllable in Japanese.

(17) The "classic model" as in Fudge (1969)



However, the point is that in (Tokyo) Japanese, the onset and the nucleus form a unit that has a stronger cohesion than that between the nucleus and the coda, which means that we have a mirror image of the rhyme in Japanese. Furthermore, Japanese is not the only language for which this type of syllable structure has been proposed. For instance, there is evidence from language games that Korean syllable structure is similar to that of Japanese, i.e. CV-C rather than C-VC. Also, recent experimental studies on sub-syllabic constituency in Korean show that speakers of Korean show a clear tendency toward left-branching (Lee 2006; Lee and Goldrick 2008; Berg and Koops 2014).[15]

[15] According to Lee and Goldrick (2008: 166), the left-branching effect emerges from "statistical properties of segment combinations". In Berg and Koops (2014), however, it is argued that the left-branching effect is not due to phonotactics but due to the relatively high degree of coarticulation between the onset and the nucleus.

The evidence for syllable structure in (Tokyo) Japanese that has been proposed in the literature comes from a range of phenomena (Kubozono 1999, 2003). McCawley (1968) analyzed the default accent rule in Tokyo Japanese as placing an accent on the syllable containing the antepenultimate mora. According to Uwano (2003) and Labrune (2012a/b), however, we can do without the syllable if we simply state the rule as accenting the third mora from the right, but that if this mora is a special mora, the accent moves one mora to the left. While such a reformulation is certainly possible, by itself the rule does not explain why the accent would move to the left rather than the right.[16] Labrune (2012a: 156) also argues that it is "not empirically true to state that the syllable is accented, because only the first part of it bears a high pitch". The evidence she gives is that in a word like *kyo˺oto* 'Kyoto', the pitch falls from H on the first half of the long /oo/ to L on the second half of the long /oo/. This of course does not even come close to evidence against the syllable as the bearer of accent; it is perfectly possible that the H portion of the accent is only allowed to associate to the head of the syllable. Actually, this is a natural assumption to make. While it is true that in for instance Kagoshima the whole syllable to which the H tone is associated is pronounced on a high pitch, this does not mean that this should be the case in all other dialects or languages in which the syllable plays a role.

Acording to Labrune (2012a: nother piece of evidence against the syllable as the bearer of accent is that in words like *obaasan˺-kko* 'child cherished by his grandmother', which consists of the noun *obaasan* 'grandmother' and the suffix *-kko* 'child', the accent falls on the moraic nasal rather than the vowel preceding it. However, this is not necessarily a problem. Although it is known that superheavy syllables are generally avoided in Japanese (Kubozono 1999), there could be exceptions to this constraint (see Ito and Mester 2015b). Alternatively, it could be that the moraic nasal and first part of the geminate form a single syllable in order to satisfy a constraint that forces the alignment of the accent with a syllable head, as in (18). This would speak in favor of the syllable as a constituent rather than against it.

(18) $[o]_{\sigma}.[baa]_{\sigma}.[sa]_{\sigma}.[n˺Q]_{\sigma}.[ko]_{\sigma}$ 'child cherished by his grandmother'

Even if we not all accents on special moras can be analyzed in terms of a syllable headed by a special mora, this does not form evidence against the syllable. It is perfectly possible that there are forces stronger than the constraint against accents on syllable non-heads that are responsible for exceptions.

[16] In Labrune's (2012a/b) approach, the reason for this would be that an accent on the final foot is avoided (NONFINALITY-Ft').

Evidence from the accent system in favor of syllable structure comes from the accent patterns of nominal compounds (see Kubozono 1999), final accent deletion before *no* (the geminate particle; see Vance 1987), the accent patterns of nominal compounds (see Kubozono 1999), and some morphoaccentual processes involving names (Kubozono 1999). Although Labrune (2012a) notes that the final accent deletion before *no* involves many exceptions, and that the evidence from compound nouns and names is complicated by the interaction with morphology, taken all together it seems difficult to ignore the role of syllable structure in these morphoaccentual processes.

The association of the phrase-initial LH tones (in the literature also kown as initial lowering or dissimilation; Haraguchi 1977), provides us with a stronger another piece of evidence for syllables. The data consisting of unaccented words in (19) is adapted from Uwano (2009). For the sake of convenience, the pitch patterns are given in "quasi-phonetic" (or "systematic phonetic") patterns (a rough approximation of the surface tonal patterns that does not take tonal underspecification into account).

(19) Phrase-initial pitch patterns in Tokyo Japanese (adapted from Uwano 2009)

	Segmental structure	*Moraic structure*	*Pitch*	*Gloss*
a.	sakana	μμμ	LHH	'fish'
b.	koori	μRμ	HHH	'ice'
	kenka	μNμ	HHH	'fight'
	kaisya	μJμ	HHH	'company'
c.i.	kappa	μQμ	LLH	'river-child'
ii.	asi̥ta	μV̥μ	LLH	'tomorrow'

According to Uwano (2009), the variation in the phrase-initial rise in Tokyo Japanese can not be explained from a synchronic point of view, even if one refers to the syllable. Uwano (2009) states that if the syllable really were involved, we would expect all special moras to behave as a dependent and accordingly assimilate in terms of tone to the head mora within the same syllable. In other words, we would expect all special moras to behave like the one in *kaQpa* (19c-i), in which the tonal value of the second mora, which is a special mora, assimilates to that of the first mora, rather than the opposite pattern observed in (19b), in which the head mora assimilates in terms of tone to the special mora. While Uwano (2009) without doubt is right that the way the patterns came into existence must be explained diachronically, his arguments against an analysis based on the syllable only make sense if we assume the rising pattern is assigned as a unit, after which rules of modification take place. However, as will be shown immediately below, an analysis in

which the phrase-initial rise is decomposed into a sequence of two tones, L and H, is perfectly possible.

Hattori 1954/55 was the first to argue in favor of syllables on the basis of the differences in the phrase-initial tones between words starting with a heavy syllable and those starting with a light syllable. That is, initial lowering does not apply in words starting with CVV (shorthand for CVR and CVJ) or CVN. In terms of phonological constraints, the argument for syllable structure on the basis of the pitch patterns in (19) can be summarized as follows.

The default pattern is the one in (19a), in which the phrase-initial L tone is associated to the first mora (= first syllable), and the phrasal H tone to the second mora (= second syllable), following the general constraints that force phrase-initial tones to associate from left to right, one by one (as in Goldsmith's (1976) "Association Convention"), starting at the beginning of the phrase. The L tone can be thought to be inserted in order to satify a constraint that demands phrases to start with a L tone, and the H tone in order to satisfy a constraint that forces accentual phrases to have a H tone.

Words with initial heavy syllables (19b) and (19c-i), as well as forms with devoiced vowels in the second syllable do not surface with the initial LH pattern.[17] This can be explained if we assume that the initial heavy syllable attracts the H tone. As shown by Gordon (2006), H tones are attracted by heavy syllables in many languages. Interestingly, in most tone languages, only heavy syllables in which the second mora contains sonorant material is treated as heavy (Gordon 2006). Related to this, Tanaka (2008) shows that syllables including the moraic obstruent *Q* count as light, as opposed to heavy syllables whose second mora dominates sonorant material, which count as heavy (see Chapter 3). Therefore, it makes sense that only heavy syllables of the form CVV and CVN attract the phrasal H tone in Tokyo Japanese. Let us assume that the attraction of H tones to heavy syllables with sonorant second moras is the result of the following constraint.[18]

(20) ALIGN(H, $CVX_{[+son]}$): Align a H tone with (the head of) a syllable of the form $CVX_{[+son]}$ (i.e. CVV and CVN).

Because the constraint in (20) only targets syllables of the form CVV and CVN, it is no surprise that words starting with a syllable of the form CVQ do not get a H tone in phrase-

[17] The tone pattern LHH is allowed for the words in (19b) in very careful speech. This suggests that in very careful speech, dependent moras are parsed into a syllable.

[18] Gussenhoven (2004) proposes that the initial-H pattern in heavy syllables is the result of the constraint NORISE according to which syllable-internal rises in pitch are not allowed (Gussenhoven 2004: 146).

initial position. Instead, the H tone skips the moraic obstruent – avoiding the association of a H tone with a syllable non-head, or alternatively, the association of a H tone with a non-sonorant mora – and associates to the following syllable. The advantage of assuming that the H tone avoids non-sonorant moras is that it also works for the phrase-initial tonal patterns of words with a devoiced vowel in the second mora, which show the same behavior as words with a moraic obstruent in the second mora. The point is that a difference in weight between syllables with sonorant second moras and those with non-sonorant second moras can be observed, which forms evidence for the syllable.

Labrune (2012a/b) criticizes the syllable-based account of the different phrase-initial tone patterns. Apart from referring to the difference between heavy syllables with sonorant second moras and non-sonorant second moras for which two different possible explanations were offered above, Labrune (2012a/b) points out that a word like *ko-i-kuchi* 'strongly flavored' may actually be realized with the LH-pattern. While this is a correct observation, this is not a problem if we assume that the high vowel in the second mora constitutes a separate syllable, which is a natural assumption because, as pointed out by Tanaka (2013), the two moras belong to different morphemes.[19]

Labrune (2012a/b) also cites work by Poser (1984) and Pierrehumbert and Beckman (1988) in which it is shown that even in phrases starting with a long vowel in what can be seen as the first heavy syllable, a rising contour can be observed. However, the fact that the phrase-initial L is not associated with the first mora does not necessarily imply that it is not there; it may well be aligned with the left edge of the phrase without being associated. In fact, this is in line with the proposals by Pierrehumbert & Beckman (1988) and Gussenhoven (2004).

All in all, Labrune's arguments against the use of the syllable in the phrase-initial rise are not convincing. I have sketched a possible account by which differences in phrase-initial tonal patterns of words with different types of syllable structure can be explained. Labrune (2012a/b), however, does not offer any explanation for the difference in behavior between sonorant and special moras on the one hand, and non-sonorant special moras and moras with devoiced vowels on the other.

Following this discussion of the role of the syllable in phrase-initial lowering, let us proceed with two more arguments given by Labrune (2012a) against the syllable. The first of these concerns the absence of resyllabification in response to onset maximization. Kawahara (2014) argues that this argument can be dismissed on the basis of onset

[19] According to Kubozono (2008), at least /ai/ sequences form a single syllable. The status of other vowel sequences is less clear. In his study on compound accent, Kubozono (1995, 1997) says that the /oi/ sequence in the word *ototoi* 'day before yesterday' is parsed into two syllables.

maximization across morpheme boundaries in verbs. For example, when the verbal root *nak-* 'cry' is followed by the negative affix *–anai*, the resulting form is *na.ka.nai*. While this is true if we adopt a morpheme-based analysis, in a stem-based analysis in which all stems end in a vowel or a moraic consonant, the data presented by Kawahara (2014) does not form conclusive evidence in favor of onset maximization. In any case, Kawahara's (2014) observation that the lack of onset maximization in forms like *an.i* 'easy' (consisting of the Sino-Japanese roots *an* 'safe' + *i* 'easy') is expected from the point of view of morpheme-syllable alignment is enough to counter Labrune's argument concerning onset maximization.

Finally, as shown by Kawahara (2014), Labrune's argument that there is no phonetic and psycholinguistic evidence for the syllable in Japanese is simply false. Kawahara (2014) discusses a range of studies ranging from phonetic measurements and psycholinguistic experiments to text-setting in Japanese songs and verse, which all suggests that syllables and syllabification play a role in Japanese phonology.

As the evidence in favor of the syllable is robust, in this thesis I assume that the syllable is part of the prosodic organization of at least Tokyo Japanese. Furthermore, there is no reason to doubt about the role of the syllable in the other dialects discussed in this thesis.

2.2.3. Foot structure

Since the influential work by Poser (1984, 1990), the bimoraic foot has gained popularity in analyses of Japanese, especially among formally oriented linguists. The evidence presented by Poser is based on the following phenomena: (i) hypocoristic formation; (ii) kinship terms;[20] (iii) geisha (bargirl) client names; (iv) rustic girls' names; (v) verbal reduplication; (vi) mimetics; (vii) a secret language used among jazz musicians; (viii) compound accentuation. Ito (1990) presents evidence for the foot from truncation and word minimality. Furthermore, studies on the compound and loanword accent in OT have also claimed that the foot is indispensable.

Labrune (2012a:171) writes that the evidence for the bimoraic foot in Japanese is "massive and uncontroversial". While the above list looks impressive, as noted by Vance (2013:171), the evidence in favor of foot structure in Japanese deserves more scrutiny than it typically receives. Furthermore, the role of foot structure in dialects other than Tokyo Japanese has been left largely unexamined. One of the goals of the present thesis is to find answers to the many remaining questions concerning the role of the foot in Japanese accentuation.

[20] Note that it is very doubtful that the kinship terms discussed by Poser (1990) are derived forms synchronically.

Foot-based analyses of Japanese accentuation are examined in Chapter 3. In this section, I will only discuss studies that are not (exclusively) related to accentuation. Thus, in this section the focus lies on evidence for foot structure in Japanese from (i) poetry, (ii) experiments and phonetics, and (iii) prosodic morphology. We start with the evidence from poetry.

2.2.3.1. Evidence from poetry

As discussed above, one of the pieces of evidence for the mora comes from traditional Japanese verse, in which lines of five moras and seven moras constitute poems: 5-7-5 in *haiku*, 5-7-5-7-7 in *tanka*. According to Bekku (1977), the moras of Japanese verse are organized into four units of two beats. That is, all lines consist of four binary units. In terms of moras, this boils down to saying that the lines consist of four bimoraic feet. Sakano (1996) further groups the four bimoraic feet into two colons (or "cola"), which makes all constituents binary. The evidence given for the idea that lines consist of eight moras comes is based on the length and positions of pauses within the verses. In the *haiku* by Matsuo Bashō (1644-94) in (21), the first and third lines would each have one pause of the length of three moras, and the second line a pause of a duration of one mora.

(21)	line 1	ki no moto ni	(5)	'Underneath the tree
	line 2	siru mo namasu mo	(7)	the soup and the fish salad
	line 3	sakura ka na	(5)	[full of] cherry blossoms'

As for the location of the pauses, there is no consensus among scholars. Asano (2002) gives a summary of the literature, and proposes a new analysis based on experimental data. She concludes that the position of the pauses may occur line-initially, line-internally, or line-finally, depending on the structure of the text. Asano (2002) also claims that the location of the pauses is a consequence of foot structure. For instance, the fact that line-initial monomoraic words are followed by a pause is taken as evidence for the parsing of this word into a prosodic word that consists of a branching foot of which the second mora is silent or covert. However, the pause following a line-initial monomoraic word could simply be a way to stress that there is a word boundary after the first mora. Note that colons are argued to be necessary in order to ensure the structure of a haiku is 8-8-8 rather than 6-8-6.

Now, if feet are really relevant in Japanese verse, an important question is how these feet relate to consituents that are relevant in the spoken language. While the foot has been argued to host the pitch accent in Tokyo Japanese, as we will see in Chapter 3, the feet

that have been proposed in the analyses of accent are not always of the same nature as those proposed to account for verse. In this light, clarifying the role that accent plays in the acceptability of poems as conforming to the prescribed rhythm may reveal more about the relation of bimoraic units in poems and possible metrical structure in the accent system. An interesting study by Iwai (1992), whose analysis is also foot-based, suggests that accent location does exert influence on the degree of acceptability of certain *haiku*. Still, more research is necessary.

2.2.3.2. Evidence from experiments and phonetics

Experimental work in order to shed light on the psychological reality of foot structure in Japanese has also been conducted. Kurisu (1994) presented words of even and odd numbers of moras to native speakers of Japanese, and asked them to draw a segmentation line at the "first segmentation point" from the start of the word. The subjects in nearly two-third of the cases drew a segmentation line before an odd-numbered mora, and in slightly more than one-third of the cases before an even-numbered mora. Kurisu (1994) concludes that this forms evidence for the idea that subjects parse words into bimoraic feet.

Tajima (1998) and Tajima and Port (2004) present evidence in favor of foot structure from "speech cycling" tasks. Speakers were asked to repeatedly produce a short text while listening to fixed types of rhythm. Tajima (1998: 60) reports that in a first experiment, "phrases with an odd number of syllables were repeated with longer inter-repetition pauses than were phrases with an even number of syllables." Furthermore, odd numbers of syllables tended to occur at harmonic phases of the repetition cycle, which suggests the feet were trochaic. In a second experiment, in which foot and word boundaries coincided, speakers did show relative temporal stability in terms of bimoraic units, with foot-initial syllables occurring at the downbeats of a three-beat rhythm. However, in a third experiment, in which foot and word boundaries did not coincide, there was more variability. It might be the case that words or morphemes rather than feet may are involved here. Tajima (1998: 111) does admit that it might have been the case that speakers may have been influenced by their awareness of "statistical tendencies in the morpho-lexical system of their language". Among the words of a list of 34,086 words analyzed by Tajima (1998), most words were quadrimoraic (44%), followed by trimoraic (27%). Moreover, 82% of the quadrimoraic words (36% of the total word list) consisted of two bimoraic morphemes (Tajima 1998: 15). The question thus is whether speakers abstract bimoraic feet or simply a bimoraic morphological template from these tendencies.

Kondo and Arai (1998) and Kondo (2006) present evidence for foot structure based on

vowel-to-vowel coarticulation. Both anticipatory and carryover coarticulation effects in the F2 values of /a/ were stronger within foot boundaries than across foot boundaries. However, the stimuli used in the experiment all consisted of quadrimoraic nonce words, which most probably were interpreted by the subjects as consisting of two bimoraic morphemes. Therefore, again it is not clear whether the foot or some other constituent (the morpheme, or the prosodic stem) formed the relevant domain of coarticulation.

Poser (1984, 1990) cites a study by Teranishi (1980) in which it is stated that, counting from the left, the even-numbered mora of every two moras tends to have a greater length than odd-numbered moras. Also, odd-numbered "stray moras" (i.e. moras that cannot form a unit with a following mora) are said to be longer in slower speech. If these observations are valid, this would suggest that Japanese has iambic feet rather than trochaic feet. This contradicts the prevalent view of Japanese foot structure being trochaic (Labrune 2012a).

Apart from the above studies, evidence for feet has been presented based in research that made use of TEMAX (in which amplitude contours are obtained through spectrographic analysis; Kitazawa et al. 1996; Ayusawa et al. 1998) and the modulation spectrum ("a 1/3-octave band, long-term average spectrum obtained from intensity contours"; Komatsu and Arai 2009). However, in these studies it is not clear under what conditions feet are built, and what evidence language learners may use in order to learn to build such feet. For instance, it is not clear whether the bimoraic units that are reported in these studies are sensitive to morpheme and word boundaries.

2.2.3.3. Feet in prosodic morphology

In this section, the role that feet have been claimed to play in the prosodic morphology of Japanese is discussed. Rather than repeating all the evidence presented in the literature, a number of phenomena through which the most important issues can be illustrated are discussed. In order to facilitate the discussion, a brief overview of theoretical issues in prosodic morphology related to the status of the foot is given.

Prosodic morphology can be defined as "the study of morphological processes that are crucially characterized in terms of a (relatively) constant output prosodic shape rather than (relatively) constant output segments" (Downing 2006: 1). The use of metrical consituents in prosodic morphology has become part of the mainstream of phonological theory, especially since the work of McCarthy and Prince (1986, 1993). The central tenet of the theory of prosodic morphology is the Prosodic Morphology Hypothesis, according to which templates referred to in morphology are defined in terms of prosodic units like the mora, the syllable, and the foot (McCarthy and Prince 1986).

There are two different approaches to templatic morphology in which reference is made to some form of the Prosodic Hierarchy. The first approach is the original version of Generalized Template Theory (GTT; McCarthy and Prince 1994), called "Prosodic Hierarchy-based Generalized Template Theory" by Downing (2006). The second approach is Downing's (2006) "Morpheme-based Generalized Template Theory".

In both versions of GTT, constraints are argued to be formulated in non-arbitrary ways. For instance, many languages are subject to a minimality condition according to which a root, stem, or word must be made up of at least two moras or two syllables. This makes sense if a root, stem, or word is subject to a constraint that ensures the mapping of the relevant morphological category onto a prosodic word (p-word), in tandem with a constraint that says a p-word must consist of at least one foot. The constraints are defined in (22) (slightly adapted from Downing 2006: 25).

(22)a. STEM=P-WORD: Stems are mapped onto p-words.
b. HEADEDNESS: Every p-word must contain a foot.

On top of this, a constraint that requires feet to be binary in terms of moras or syllables, FOOT-BINARITY (FT-BIN; Prince and Smolensky 1993) will make sure that, depending on the language, a p-word consists of at least two moras or two syllables.

Because prosodic morphemes are defined in terms of morphological units, rather than phonological ones, construction-specific restrictions on the size of morphemes are not allowed, which makes the theory restrictive.

The most important difference between the Prosodic Hierarchy-based and morpheme-based versions of GTT in the context of this thesis is that in the former a single Prosodic Hierarchy is assumed, while in the latter the co-existence of the Prosodic Hierarchy and the Metrical Hierarchy discussed above is assumed.

In her discussion of Prosodic-Hierarchy-based GTT, Downing (2006) gives a number of reasons for why collapsing the two hierarchies into one hierarchy in which the p-word dominates the foot is problematic. Here I will only mention the most important problems.

One of the problems concerns evidence for the existence of interface categories below the category of the p-word. As mentioned in the Chapter 1, for certain languages it has been claimed that it is necessary to distinguish between a morphological stem and a prosodic stem. Still, even if we accept this evidence, it is not necessarily true that all languages have interface categories below the p-word.

Another problem raised by Downing (2006) concerns the central role played by the foot in accounting for minimality effects across languages. She points out that in many

languages, a minimal word does not correspond to a minimal foot (see Hayes 1995; Garrett 1999). Related to this, if bimoraic and disyllabic minimality requirements are analyzed in terms of foot structure, we do not expect languages without foot structure in their prosodic systems to be subject to minimality conditions (see Garrett 1999). Also, in certain languages, different types of words have different minimality conditions, e.g derived vs. underived forms (e.g. (Tokyo) Japanese; Ito 1990). This is not expected if minimality is an automatic consequence of the fact that a p-word must dominate a branching foot.

Downing (2006) develops an alternative theory based ons the idea that lexical heads must branch (see Dresher and van der Hulst 1998), where branching is satisfied not only when the constituent in question branches, but also when the constituent that it dominates branches. The definition of prosodic branching is given in (23).

(23) PROSODIC BRANCHING (Downing 2006: 122)
A constituent branches *iff* it or its daughter contains more than one daughter.

The constraint is adapted by Downing (2006) from Ussishkin (2000: 43), who refers to a similar constraint proposed by Ito and Mester (1992), who call the principle "Hierarchical Locality" (see below).

The point made by Downing (2006) is that in a language in which there is no evidence for feet, minimality constraints can be defined in terms of the branching of morpho-prosodic consituents such as the the p-word or prosodic stem (p-stem). However, this seems to imply that an interface category (the p-stem) dominates a rhythmic category (a syllable), which means the hierarchies are in some way collapsed into one after all. If so, it is not clear why an approach in which the p-word directly dominates the foot would be problematic.

Now that we have discussed some theoretical issues in prosodic morphology, let us turn to data from Japanese and investigate whether we can find evidence for the nature of the prosodic constituents in question (projected from morpho-syntactic constituents or purely rhythmic categories for which there is independent evidence), and the exact relations among the different constituents.

The work of Ito (1990), Poser (1990), and Ito and Mester (1992) has shown that a common feature of the different prosodic-morphological phenomena in Japanese is that (i) stems consist of at least two moras, and (ii) words consist of at least two syllables. In (24), which is adopted from Ito (1990: 222), "F" stands for "foot".

(24) a. Minimal Stem: F = (μμ)
b. Minimal Word: > σ

An important observation made by Ito (1990) is that the bimoraic minimality requirement only holds for derived forms. Evidence for this comes from the forms in (25), which are taken from Ito (1990). While the underived words in (25a) are grammatical, in the hypocoristic form in (25b), the stem that is followed by the diminutive suffix *-tyan* must be bimoraic, and a word derived by truncation as in (25c) must be disyllabic.

(25) Minimality in derived words only

a.	ya_{μ}			'arrow'
	yu_{μ}			'hot water'
b.	*ma_{μ}-tyan	$ma_{\mu}.ri_{\mu}$.-tyan		'Mariko-DIM'
c.	*bi_{μ}	$bi_{\mu}.ru_{\mu}$	< bi.ru.din.gu	'building'

The difference between stems and words can be illustrated by the forms in (26), which are again taken from Ito (1990). Because the form in (26a) is followed by a suffix, and thus is a stem, it only needs to consist of two moras. The form in (26b) needs to stand by itself, and therefore must consist of at least two syllables.

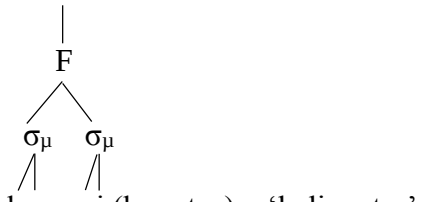
(26)	Base	Truncation	
a.	saiko	sai_{σ}-$tyan_{\sigma}$	'Saiko-DIM'
b.	saikederikku	*sai_{σ}, $sai_{\sigma}ke_{\sigma}$	'psychedelic'

The reason that the minimal word requirement in (24) is not defined as a disyllabic foot $[\sigma\sigma]_F$ by Ito (1990) lies in the fact that there is no rhythmic constituent that corresponds to this structure, the Japanese foot being assumed to be bimoraic. This raises the following question: why does the minimal stem correspond to a rhythmic constituent, while the minimal word does not? Ito and Mester (1992) recognize this problem, and propose that the obligatory disyllabicity is the result of a more general constraint on branching they call "Word Binarity", which states that words derived in the prosodic morphology must be prosodically binary. This constraint not only explains why words are minimally disyllabic, but also why words in Japanese prosodic morphology maximally consist of four syllables; the longest possible word consists of two bimoriac feet. The three types of structure that satisfy Word Binarity are presented in (27).

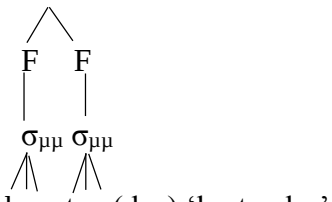
(27) a. $_{\omega}[_{F}(\sigma_{\mu}\sigma_{\mu})]$
b. $_{\omega}[(F)\ \sigma_{\mu}]$
c. $_{\omega}[(F)(F)]$

Interestingy, the three types of structure that satisfy Word Binarity given in (27) are exactly the three types of words that are attested in as output forms of truncations or "clippings" (28).

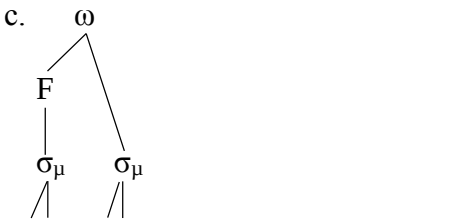
(28) a.



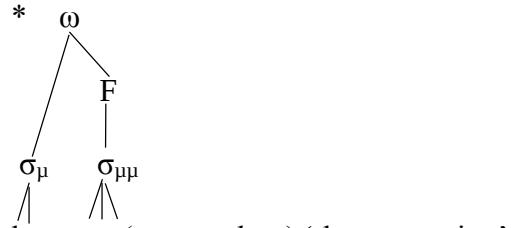
b.



c.



d. *



The first three types of structure in (28a/b/c/) are all grammatical, but the fourth type of structure in (28d) is ruled out by an alignment constraint that forces the left edge of a p-word to be aligned with a foot ("Left Edge Matching"; Ito and Mester 1992).

The reason why in derived words a single heavy syllable is not allowed to stand as a free word lies in the Hierarchical Locality constraint. According to this constraint, which is identical to the Prosodic Branching constraint in (23), access to structural information is only available at the level of the constituent in question itself or the next level immediately dominated by this constituent. Thus, a p-word has access to the internal structure of feet, but not of syllables. A foot, however, does have access to the internal structure of syllables.

Let us now turn to the question whether the bimoraic units to which minimal stems in Japanese prosodic-morphology correspond are necessarily feet. If the minimality requirements are defined in terms of branching anyway, in a Prosodic Hierarchy of the form {p-word → p-stem → syllable → mora}, we may entertain the possibility that what must consist of two moras is a p-stem, rather than a foot. This proposal makes sense when we realize that in a construction like the hypocoristic form, *-tyan* can be assumed to be

specified in the lexicon as a suffix that attaches to a morpho-syntactic stem which maps onto a p-stem. If a p-stem directly dominates syllables, it is branching if it consists of at least two moras. The branching constraint is satisfied in (29a), but not in (29b).

(29)	*Input*	*Morphology*	*Mapping onto p-stem*
a.	saiko	sai-tyan	$[(sa_{\mu}i_{\mu})_{\sigma}]_{\text{p-stem}}$tyan
b.	saiko	sai-tyan	$*[(sa_{\mu})_{\sigma}]_{\text{p-stem}}$tyan → violation of minimality condition

In (30), the alternative p-stem-based approach to truncation is schematically illustrated. P-stems are indicated by "ψ".[21]

(30) a.	$_{\omega}[_{\psi}[\sigma_{\mu}\sigma_{\mu}]]$	e.g.	$_{\omega}[_{\psi}$[heri]	'helicopter'
b.	$_{\omega}[[\psi]\ \sigma_{\mu}]$	e.g.	$_{\omega}[_{\psi}$[kon]bi]	'combination'
c.	$_{\omega}[[\psi][\psi]]$	e.g.	$_{\omega}[_{\psi}$[baa]$_{\psi}$[ten]]	'bartender'

What may speak in favor of the foot-based analysis is that in the p-stem based analysis it is not clear why the option in (30c) is unmarked. That is, it is not clear why the output of a truncation would consist of two p-stems, the prosodic counterpart of two morphological stems (m-stems) rather than simply one. A possible answer to this question could be that the template in (30c) is the result of analogical extention of the pattern of (p-)stem-(p-)stem compounds such as *rimo-kon*, which is derived from *rimooto kontorooru* 'remote control'.[22]

There is even a third possibility, namely that both p-stems and feet are involved. Actually, this third possibility seems to correspond to the approach taken in Ito (1990), where the stem is also referred to as a "prosodic stem", although the reason for this might have been practical (i.e. to distinguish between stems in general and stems in the prosodic morphology) rather than formal.

The question whether we are dealing with feet, p-stems, or both, is not unimportant from a theoretical point of view. For instance, the analysis of minimality by Ito and Mester (1992) is based on the premise that within the Prosodic Hierarchy, the p-word dominates the foot, the foot dominates the syllable, and the syllable dominates the mora. However, if it turns out that the p-stem is also necessary in Japanese phonology (see Section 2.3.3.

[21] Note that the option of replacing the foot by the p-stem is available for the following phenomena described by Poser (1990) that involve minimality conditions: hypocoristic formation, kinship terms, geisha (bargirl) client names, rustic girls' names, verbal reduplication, and mimetic root minimality.
[22] Ito (1990) proposes a (p-)stem-(p)-stem analysis of such compound truncations, in which the (p-)stem is subject to a constraint that it must consist of a bimoraic foot.

and Chapter 6), the organization of the Prosodic Hierarchy needs to be reconsidered.

Furthermore, there is the issue of Syllable Integrity. According to Syllable Integrity, foot boundaries must coincide with syllable boundaries (see Prince 1976; Everett 1997). However, in a number of studies on Japanese accentuation, it is assumed that foot boundaries may straddle syllable boundaries (Suzuki 1995; Inaba 1998; Mutsukawa 2009; Tanaka 2013; Kager and Martínez-Paricio 2014; see also Chapter 3). Under the assumption that syllables are dominated by feet, the principle of Proper Bracketing (Nespor and Vogel 1986), according to which "a given constituent cannot be simultaneously part of two or more higher prosodic constituents"; Ito and Mester 1992: 8) would be violated. If Proper Bracketing is a universal principle, which is the way in which it seems to be treated in the literature, violations of Syllable Integrity are only possible if the foot directly dominates the mora rather than the syllable. If the arguments in favor of violations of Syllable Integrity are valid, this would have serious consequences for the approach to minimality based on the foot. The reason for this is that if there is no foot level between the level of the p-word and the level of the syllable, it is not clear why the p-word has no access to the internal structure of syllables.[23] If p-stems, instead of feet, are situated between p-words and syllables, however, minimality effects could be accounted for (see Downing 2006 for the application of this idea to a range of languages).

Possible (but inconclusive) evidence for the violability of Syllable Integrity in Japanese comes from mimetic forms. According to Nasu (2002: 55), (at least derived) mimetic forms are subject to a "template satisfaction condition" (see McCarthy & Prince 1986). That is, mimetic output forms must fill a quadrimoraic template (31).[24]

(31) $[\mu\ \mu\ \mu\ \mu]_{\omega}$

Whenever this template is filled, the quotative particle *-to*, which can be optionally used in mimetic constructions, is kept outside the prosodic word, the result of which is the absence of a geminate.[25] All words in (32) are based on the root *pika*, and mean something like 'brightly' or 'in a flashing manner'.

[23] Note that if we assume the Prosodic Hierarchy is not fixed within a language, but erected in the computation of output forms instead, the problem of Syllable Integrity could be dealt with by allowing different versions of the hierarchy for different phenomena.

[24] According to the template satisfaction condition, "all elements in a template are obligatorily satisfied" (McCarthy and Prince 1986: 6).

[25] Note that the template is only a minimum requirement; emphatic forms like *pikapikaQto* involve spreading from the particle to the base, which implies the particle is p-word-internal. The incorporation of the particle into the p-word enables the emphatic geminate to be realized.

(32)a. [μ μ μ μ]$_{ω}$
| | | |
piˀkapika to

b. *[μ μ μ Ø]$_{ω}$ → [μ μ μ μ]$_{ω}$ *or* [μ μ μ μ]$_{ω}$
| | | | | | |
pikari pikaˀri-to pi Qkaˀri (to)

A quadrimoraic template without internal constituency is suspect from a formal point of view; phonology is assumed not to be able to count beyond two (McCarthy 1986).[26] Interestingly, Nasu (2002) does not argue that this template consists of two bimoraic feet, even though he uses feet in other parts of his analysis. The reason for this seems to be related to this proposal that in a form like *pikaˀQto* 'brightly', which also exists, the accent is realized in a bimoraic foot that comprises the moraic obstruent: [pi (kaˀQ) to]$_{ω}$. This means that if we analyze the quadrimoraic template as consisting of two feet, we have a problem, as shown in (33).

(33)a. [(μ μ) (μ μ)]$_{ω}$ Foot structure that would be needed for a quadrimoraic template
b. [μ (μˀ μ) μ]$_{ω}$ Foot proposed by Nasu (2002) to account for accentuation

If we define the template in terms of the foot, in forms that have a moraic obstruent as their third mora, we either need to assume the moraic obstruent forms a syllable by itself (34c-i), or allow violations of Syllable Integrity (34c-ii),.

(34) a. [(piˀ.ka.)$_{F}$(pi.ka)$_{F}$]$_{ω}$
b. [(pi.kaˀ.)$_{F}$(ri.to)$_{F}$]$_{ω}$
c.i. [(pi.<kaˀ>$_{σ}$)(<Q>$_{σ}$.to)$_{F}$]$_{ω}$ → moraic obstruent heading a syllable
c.ii. [(pi.<kaˀ)(Q>$_{σ}$. to)$_{F}$]$_{ω}$ → violation of Syllable Integrity

[26] Psychological research suggests that amounts of two and maybe three do not require actual counting. There is clear evidence for a sharp distinction between the status of three and four in terms of psychological processing (Dehaene 2011). This means that a quadrimoraic template ought to be defined in terms of smaller subconstituents. In generative phonology, counting beyond two has been claimed to be impossible (McCarthy and Prince 1986; Odden 1994; cf. Marlo et al. 2014). This is one of the reasons why default antepenultimate accent in Tokyo Japanese, as well as in several other languages, has been analyzed in terms of a non-final foot, with the final mora (and/or syllable) being "invisible" or extrametrical (see Chapter 3).

Both options in (34c) may be said to be controversial representations, but neither of the two can be ruled out in principle. The form in (18) above was proposed to involve a syllable consisting of a moraic obstruent and a moraic nasal, with the former heading the syllable. In that sense, a representation as that in (34c) may be allowed in order to fill the template in cases in which the moraic obstruent is used. On the other hand, violations of Syllable Integrity have been argued to be necessary for other languages, the most well-known case of which is Pirahã (Everett 1997).[27]

What we can conclude from the above discussion is that there are reasons why the foot-based approach to minimality effects may need some rethinking. Importantly, assumptions about foot structure made in analyses of accentual phenomena may be relevant for issues in the prosodic morphology.

The strongest case for a role of foot structure in the prosodic morphology of Japanese (as opposed to the p-stem) comes from a topic that is only marginally part of the grammar of Japanese: the secret jazz language or argot called "zuuzya-go" (henceforth ZG; Tateishi 1989; Ito et al. 1996; Kurisu 2005). In this secret language, different types of constituents of the base are metathesized or reversed. Before we delve into the generalizations and possible analyses, the relevant data is introduced.

The data in (35a-e) is from Tateishi (1989), and that in (35f) from Ito et al. (1996).

(35)a.	hi	→	ii.hii	'fire'
	he	→	ee.hee	'fart'
b.	hai	→	ii-haa	'lungs'
	pan	→	nn-paa	'bread'
	ha.ra	→	raa.haa	'stomach'
c.	on.na	→	naa.on	'woman'
	toon	→	nn-too	'tone'
	pa.taan	→	tan.paa	'pattern'
	piyano	→	yanopi	'piano'
	kusuri	→	suriku	'drug'
	tukue	→	kuetu	'desk'
d.	kya.ra.me.ru	→	me.ru.kyara	'caramel'
	koo.hii	→	hii.koo	'coffee'

[27] Kager and Martínez-Paricio (2014), who make use of recursive feet, argue that apart from "conservative Tokyo Japanese" (see Chapter 3), Gilbertese and Dihivo Macedonian violate Syllable Integrity. In their approach, violations of Syllable Integrity are allowed if the PARSE constraint is defined in terms of moras rather than syllables.

e.	ma.nee.zyaa	→	zyaa.ma.ne	‘manager’
	ko.maa.sya.ru	→	sya.ru.ko.ma	‘commercial’
f.	i	→	-	‘stomach’
	oo	→	-	‘king’
	suu.paa.man	→	-	‘superman’

The data in (35f) is interesting, as they show words consisting of a single vowel and words of more than five moras have no ZG form. Words of five moras may only have a ZG form if they start with a sequence of light-heavy syllables, with the second heavy syllable containing a long vowel (Ito et al. 1996: 233).[28]

Tateishi (1989) reports that for some bases, variant ZG forms exist, as in (36).

(36)	yama-da	→	daayama, dayama	‘Yamada (surname)’
	koohii	→	hiikoo, hikoo, hiiko	‘coffee’
	onna	→	naaon, naon	‘woman’
	pataan	→	tanpaa, tanpa	‘pattern’
	pan	→	nnpaa, npaa, nnpa	‘bread’
	maneezya	→	zyaamane, zyamane (*zya(a)manee)	‘manager’

The data given by Tateishi (1989) and Ito et al. (1996) does not always correspond. It seems that in Tateishi (1989) trimoraic output forms are avoided more than in the data presented in Ito et al. (1996). Ito et al. (1996) attribute the variation to different ZG “dialects”.

Tateishi (1989) was the first study of ZG that presented a comprehensive analysis. An abbreviated version of his analysis is given in (37).

(37) a. Copy all melodies of the rightmost largest constituent of a word that does not cover the whole string of the base to the bimoraic template. […]
b. Suffix a bimoraic template to the newly-created template, and copy the melody from the melody still linked to the base.

Simplified versions of sample derivations that can be found in Tateishi (1989) are summarized in (38). The rightmost largest constituent structures are the morpheme (or

[28] The reason for this is that a quadrimoraic output form is possible for these words without erasing featural material (Ito et al. 1996).

foot) in (38a), the foot in (38b), the syllable in (38c), the mora in (38d),[29] and the segment in (38e).

(38) a.	yama-[guti]	→	[guti]	→	[guti][yama]
b.	pi[yano]	→	[yano]	→	[yano][pi(i)]
c.	ha[ra]	→	[raa]	→	[raa][haa]
d.	pa[n]	→	[nn]	→	[nn][paa]
e.	h[i]	→	[ii]	→	[ii][hii]

Apart from the possibility to view the "rightmost largest constituent of a word" in (38a) and (38b) as a foot, the bimoraic template itself can of course also be analyzed as a foot.

Ito et al. (1996) present additional (and variant) data and make the generalizations summarized in (39), where parentheses indicate foot boundaries.

(39)	Base		ZG
	μ	→	(μμ)μ
	μμ	→	(μμ)μ
	μμμ	→	(μμ)μ
	μμμμ	→	(μμ)(μμ)[30]
	μV:μμ	→	(μμ)(μμ)

Ito et al. (1996) formalize their analysis in terms of Correspondence Theory (McCarthy and Prince 1995). The most interesting part of the analysis is their treatment of trimoraic words, for which they report two different patterns.

Ito et al. (1996) show that not all trimoraic words behave like the ones in (35c). They do analyze words like *piyano* as having a final bimoraic foot, as in Tateishi (1989). However, they argue that words that show this pattern are exceptions. In regular forms, the trimoraic output is reversed, as in (40). The call this type of reversal "segmentally non-exhaustive reversal".

[29] In Tateishi (1989), the mora does not include the onset.

[30] Ito et al. (1996) state that inputs with ligh-heavy-light (lhl) syllable structure have no ZG counterpart if the heavy syllable does not consist of a long vowel. In this respect such forms thus resemble lhl-words of five moras.

(40) a.	tenisu	→	sunite	‘tennis’
b.	bando	→	donba	‘band’
c.	sakana	→	nakasa	‘fish’
d.	kokoro	→	rokoko	‘heart’

The forms in (40) can be analyzed in different ways. Ito et al. (1996) mention that we could assume that the forms are simply read from right to left (in moraic units) instead of left to right. However, Ito et al. (1996) opt for another analysis: the mora in the middle stays in place, and the first and final moras are swapped. This idea is formalized in terms of constraints that enforce “prosodic role faithfulness”. A distinction is made between foot heads and foot non-heads, or “foot tails”. The foot tails, which are underlined in (41), are forced to stay where they are because of “prosodic role preservation”.

According to Ito et al. (1996), the foot structure in (41) is due to a constraint against a head foot (F′) in p-word-final position (NONFINALITY(F′)).

(41)a.	(te<u>ni</u>)su	→	(su<u>ni</u>)te	‘tennis’
b.	(ba<u>n</u>)do	→	(do<u>n</u>)ba	‘band’
c.	(sa<u>ka</u>)na	→	(na<u>ka</u>)sa	‘fish’
d.	(ko<u>ko</u>)ro	→	(ro<u>ko</u>)ko	‘heart’

The analysis in terms of foot heads and foot tails also makes the right predictions for the trimoraic forms of (40c), which have the following foot structure.

(42)a.	pi(ya<u>no</u>)	→	(ya<u>no</u>)pi	‘piano’
b.	ku(su<u>ri</u>)	→	(su<u>ri</u>)ku	‘drug’
c.	tsu(ku<u>e</u>)	→	(ku<u>e</u>)tsu	‘desk’

The only difference between the two sets is that the foot structure of the base must be assumed to be known in the input. This is of course only possible if speakers have evidence for differences in foot structure of these words from surface forms in Japanese. Without such evidence, the foot-based analysis is problematic.

Although the foot structures in (41a) and (41b) can be said to be motivated because of the initial accent in *te’nisu* ‘tennis’ (42a) and the dependent mora in second position in (41b), the foot structure of the forms in (41c) and (41d) can only be based on the assumption that final head feet are banned. However, if this is the case, it remains a mystery why the feet in (42) are word-final. The answer to this problem given by Ito et

al. (1996) is that words like *pi(yano)* 'piano' must be lexically stored including foot structure. However, the question is on the basis of what evidence this foot structure is deduced in the first place.

There may be a possible answer to why the foot structure of the words in (41) and (42) differs, though. The initial moras of the words in (42) have (i) a voiceless consonant and (ii) a high vowel. This means that these moras may often be devoiced, especially in casual speech.[31] In other words, the initial syllables of these words are "weak". While this may be a possible answer to the question of why the two sets of words would have different metrical structure, there is one caveat; the word *kokoro* 'heart' is a famous example of a word in which a word- (and phrase-)initial non-high vowel devoices. The question thus is why devoicing in *kokoro* would not result in the alignment of the foot to the right edge of the word whereas it does in for example *kusuri* 'drug'. The answer may lie in the accent pattern with which the informant(s) pronounce the word *kokoro*. If it is *kokoro*ꜜ, we could assume the footing *(koko)(ro*ꜜ*)*, the ZG output form of which is indeed *(roko)(ko)*. On the other hand, if it is *koko*ꜜ*ro*, we would expect *ko(ko*ꜜ*ro)*.[32]

Whether the difference in foot structure between the two sets of words can be inferred from the data is a rather important question. The reason for this is that a distinction similar to that between a foot head and a foot tail is absent in interface constituents like the p-stem. In that sense, the absence of a distinction between head and tail in p-stems makes it tempting to accept the foot-based analysis, and thereby adopt a foot-based approach to Japanese prosodic morphology in general. Still, as pointed out above, it could well be that both consituents are involved.

In summary, although the evidence for foot structure from prosodic morphology looks convincing at first sight, there are some issues that need to be resolved. Related to the above discussion is the question whether a constituent refered to in the prosodic morphology should be independently motivated. According to the principle of metrical coherence (Lahiri 2001), we expect metrical constituents in prosodic morphology to be identical to metrical constituents that are independently motivated. Lahiri (2001: 1361) observes that "almost always the foot used for stress is the same as that used for prosodic morphology". In her discussion of putative counterexamples, Lahiri (2001) suggests that alternative analyses may be available, and that furthermore metrical coherence tends to

[31] The initial high vowel in the words in (42b) and (42c) are very likely to get devoiced even in more careful speech.

[32] The NHK accent dictionary gives two forms: *koko*ꜜ*ro* and *kokoro*ꜜ. As far as I know, the latter is the innovative variant of the two pronunciations. Note that if the non-ZG form is *(ko)(ko*ꜜ*ro),* the ZG output form would be identical to the Japanese form in terms of segmental structure, because the two foot heads that are switched are both ko. If the foot structure *(ko)(ko*ꜜ*ro)* is correct, this may be another for why rokoko is chosen as the output form.

be about preferences rather than absolutes. On the other hand, Yip (1995) argues that isolated instances of prosodic categories are expected if prosodic categories are universally available; some languages may simply make more use of certain categories than others. We will come back to issues concerning metrical foot structure in Chapter 3, where the evidence for foot structure from Tokyo Japanese accentuation is critically examined. Furthermore, in Chapters 4 and 5 the possible role of foot structure in other dialects is investigated. Finally, in Chapter 6, the role of prosodic structure in verbal paradigms across a number of dialects is investigated.

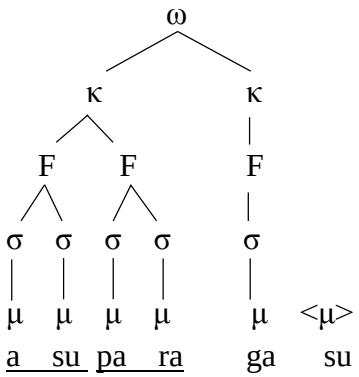
2.2.3. Colon and superfoot

The colon or superfoot is a rhythmic category that dominates the foot. In other words, a colon or superfoot consist of one or multiple feet. Evidence from accentuation for this category is discussed in Chapter 3. In this section, Hashimoto's (2012) arguments for the colon from loanword truncation are evaluated.

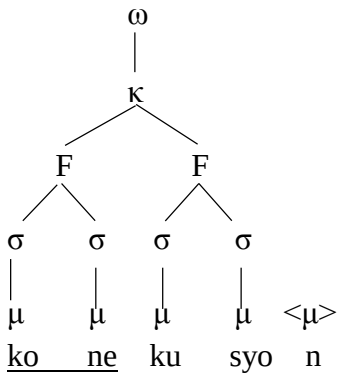
Hashimoto (2012) argues that the colon is necessary to account for loanword truncations. His analysis can be summarized as follows. Under the condition that the word-final mora be extrametrical, a truncated form consists of the left-hand part of the highest-level constituent that branches starting from the p-word (modulo phonotactic repairs). Hashimoto's (2012) analysis is able to account for 93 out of 109 forms.

To give an example, if the p-word itself is the first constituent in the hierarchy that branches, then the left colon is selected as the truncated form (*asupara* < *asuparagasu* 'asparagus' in (43a)). If, on the other hand, the first branching unit is the colon (κ), then the leftmost foot of this colon is selected (*kone* < *konekusyon* 'connection' in (43b)). Note that feet are assumed to be built from left to right. Underlining indicates the truncated portion of the segmental string.

(43)a.



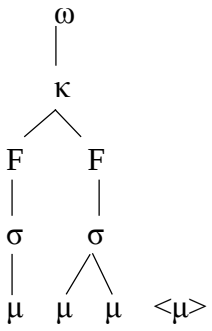
b.



While highly original and interesting, Hashimoto's (2012) approach is not without problems. To start with, the truncation *ibe* (< *ibento* 'event') is the result of the combination of material from two feet, *(i)* and *(ben)*, in which the final moraic nasal is deleted due to a constraint against word-final heavy syllables (*H#). Thus, the supposedly optimal form consists of two degenerate feet. This is counterintuitive, however, and an analysis in which the first two moras are mapped onto some bimoraic unit seems more natural.

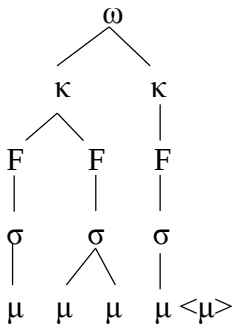
The reason why in Hashimoto (2012) *ibe* is not a single foot is illustrated by (44) and (45). In (44), which is the actual analysis, feet are assumed to dominate syllables, rather than moras, and the constraint Syllable Integrity may not be violated.

(44)a.



i be n to 'event'

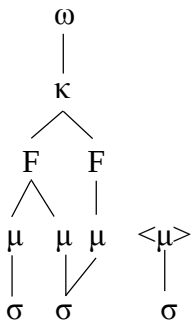
b.



ri ha a sa ru 'rehearsal'

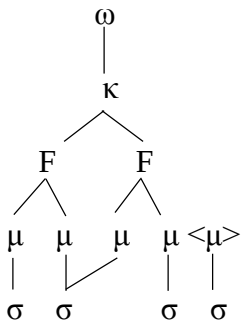
If Syllable Integrity were violable, as in (45), we would get **rihaasa* as the truncated form of *rihaasaru*, which is incorrect (45b). Thus, in Hashimoto's (2012) analysis we cannot but accept that the form *ibe* is not a single foot.[33]

(45) a.



i be n to 'event'

b. *



ri ha a sa ru 'rehearsal'

Apart from the counterintuitive nature of the analysis, adopting the colon is problematic

[33] Note that although Syllable Integrity is assumed to be inviolable in the foot-building process, at the same time it is assumed that the final mora of a heavy syllable can be extrametrical.

from the viewpoint of minimality effects in other types of words. Although the colon may enable us to account for loanword truncations, the consequence of adopting the colon is that according to Hierarchical Locality, a p-word is now only considered to be branching if it consists of two colons or two feet, which is not in line with the facts. Of course we could redefine Hierarchical Locality in such a way that a constituent branches if itself, its daughter, or its granddaughter branches. However, such a definition is problematic from the viewpoint of locality. Therefore, I conclude that rather than adopting the colon, a template-filling approach to loanword truncations along the lines of Ito (1990) and Ito and Mester (1992) is to be preferred.[34]

2.3. Interface categories

Most of the research on interface categories in Japanese has focused on the larger phrasal domains that are relevant to intonation, in particular the "accentual phrase" (Pierrehumbert and Beckman 1988), also known as "minor phrase" (McCawley 1968; Poser 1984; Kubozono 1988), and the "intermediate phrase" (Pierrehumbert & Beckman 1988), which has also been called "major phrase" (McCawley 1968; Poser 1984; Kubozono 1988). In this thesis, the terms minor phrase and major phrase are adopted. However, where it is not necessary to distinguish between the two, I will use the term phonological or prosodic phrase (p-phrase) or simply "phrase".

The importance of the phrase in Tokyo Japanese has cast a shadow on the prosodic word, which has not received a lot of attention. In order to facilitate the discussion of the prosodic word, we will take a look at the minor phrase first.

2.3.1. Minor phrase

The minor phrase is known as the domain of (i) culminativity of pitch accent, and (ii) initial lowering, i.e. the assignment of the phrase-initial LH tones (McCawley 1968; Poser 1984; Pierrehumbert & Beckman 1988; Kubozono 1988). As such, the domain has a very prominent role in the tonal grammar, and therefore can be relatively easily identified. The accentual phrase is dominated by the major phrase, which is the domain of catathesis or downstep, i.e. pitch range compression following a pitch accent.

[34] Other alternatives that should be mentioned are the analyses by Labrune (2002) and Kubozono (2010). Labrune's (2002) analysis is based on the hypothesis that "the base is truncated immediately before the accented mora" (Labrune 2002: 106). Kubozono (2010) criticizes Labrune's (2002) approach and presents an alternative analysis that makes use of pseudo-constituents, the boundaries of which are determined by a principle according to which a word is cut in half in the middle of the word. A problem of his pseudo-compound analysis is that it is not clear what units are used in order to count length. Kubozono (2010) argues that pseudo-compounds are morphologically simplex, but prosodically complex. If so, it should be spelled out what prosodic constituents we are dealing with.

In terms of its morpho-syntactic counterpart, the minor phrase is related to the XP. However, not every XP is mapped onto a minor phrase. Thus, the minor phrase may consist of multiple words, as long as only one of them carries an accent.[35]

While minor phrases may consist of multiple words, Warner et al. (2010) demonstrate that the initial rise actually functions as a reliable cue to word boundaries. In their experimental data, 40% of all words onsets showed clear ("non-catathesized") pitch rises, and in 19.5% "catathesized but clear rises" could be observed. This means that in almost 60% of all lexical words, the word boundary coincided with a minor phrase boundary. Furthermore, a word-spotting experiment showed that listeners indeed use the phrase-initial rise as a cue for word boundaries. The study by Warner et al. (2010) thus shows that the minor phrase bears a close relation to morpho-syntactic words. Together with the fact that the domain of accent culminativity is the minor phrase rather than the prosodic word, it is not surprising that the prosodic word has been relatively neglected in the literature.

2.3.2. Prosodic word

According to Nespor & Vogel (1986: 107), the prosodic word (or "phonological word"; henceforth "p-word") can be defined as "the lowest constituent of the prosodic hierarchy which is constructed on the basis of mapping rules that make substantial use of nonphonological notions". In the ideal case, in which a one-to-one mapping between morpho-syntactic consituents and corresponding phonological constituents is assumed, the p-word would simply be the phonological counterpart of the morpho-syntactic word. In other words, as long as no other phonological constraints interfere, every morphosyntactic category X (N, V, or A) would correspond to a prosodic word. In a recent version of Prosodic Phonology known as Match Theory (Selkirk 2009, 2011), this is exactly what is claimed to hold true for languages universally. Thus, morpho-syntactic words map onto p-words, morpho-syntactic phrases map onto phonological phrases, and morpho-syntactic clauses map onto intonational phrases.[36] Although there seems to be no doubt that the phonological word is related to the morpho-syntactic word, the question is how morpho-syntactically complex constructions are mapped onto p-words. For instance, two different analyses have been proposed for Japanese compounds (46a). In the first analysis (46a-i), compounds generally consist of a single phonological word. In this analysis, the default compound accent that can be observed in words in which the

[35] See Kubozono (1988/1993) for a discussion of the relative roles of syntactic structure and phonological constraints in phrase formation.
[36] Ito and Mester (2013) call this highest level of structure "phonological clause".

accent of the second member (N2) is not parsed is analyzed as a morpheme-boundary marker (Kubozono 1995, 1997; Tanaka 2001; Ito & Mester 2003; Nishimura 2013). In an alternative analysis (46a-ii), the different members of a compound comprise different p-words, with the accent serving as a marker of a p-word boundary (Ito and Mester 2007). Ito and Mester (2007) mention that the absence of an accent in the first member of a compound (N1) has been taken as evidence for the phrasing of compounds into a single p-word (46a-i), as opposed to English (46b), in which both members of a compound are accented and must therefore be assumed to form separate p-words (see Tanaka 2001; Ito and Mester 2003).

(46)

a.i.	[gogaku kyoˀosi]$_{\omega}$	‘language intructor’
	[yama noˀbori]$_{\omega}$	‘mountain climbing’
ii.	[[gogaku]$_{\omega}$ [kyoˀosi]$_{\omega}$]$_{\omega}$	‘language intructor’
	[[yama]$_{\omega}$ [noˀbori]$_{\omega}$]$_{\omega}$	‘mountain climbing’
b.	[[lánguage]$_{\omega}$ [instrùctor]$_{\omega}$]$_{\omega}$	
	[[móuntain]$_{\omega}$[clìmbing]$_{\omega}$]$_{\omega}$	

Bearing in mind that a minor phrase may be associated with only a single accent, and may even be unaccented, the argument from the lack of accent in N1 of course does not hold water. It may well be the case that the compound consists of multiple p-words but only a single minor phrase.[37] In the absence of evidence for the single p-word approach, is there evidence for an approach in terms of multiple p-words? According to Ito and Mester (2007), from the point of view of the Indirect Reference Hypothesis, the “junctural accent” is better analyzed as aligned with a p-word boundary.

Other evidence for recursive p-words comes from the behavior of Sino-Japanese morphemes in compounding. Ito and Mester (1996) demonstrate that gemination across p-word boundaries is not allowed, which explains the difference between the words *bes-seki* and *toku-betu-seki* in (47).

(47) a.	[betsu-seki]	→	[beQseki]$_{\omega}$	‘different seat’
b.	[[toku-betu]seki]	→	[[toku-betsuˀ]$_{\omega}$ [seki]$_{\omega}$]$_{\omega}$	‘special seat’

Note that additional evidence for a difference in prosodic structure comes from the difference in accentuation between the two types of words; the word in (47a), which

[37] Note that some compounds actually consist of multiple phrases. See Chapter 3 for discussion.

consists of a single p-word, is unaccented, whereas the recursive p-word in (47b) is assigned a junctural accent at the final mora of N1.[38]

Interestingly, the restriction on the alternation of /p/ with /h/ in Sino-Japanese compounds is constrained in the same way.

(48) a.	[en-hitu]	→	[en-pitu]$_{\omega}$	'lead-pen; pencil'
b.	[[man-nen]hitu]	→	[[manne˜n]$_{\omega}$ [hitu]$_{\omega}$]$_{\omega}$	'10,000years-pen; fountain pen'

Ito and Mester (2007) distinguish between minimal and maximal projections of p-words. Maximal p-words dominate other p-words, and are dominated by a prosodic unit of another type (the p-phrase, which encompasses both the minor and major phrase in Ito and Mester's (2007, 2011, 2013) theory). The junctural accent is aligned to a p-word boundary contained by the maximal p-word.

Rendaku, or "sequential voicing", is another morphophonological process that may occur in compounds.[39] Rendaku involves the alternation of a voiceless word-initial obstruent (h, t, k, s) to a voiced obstruent (b, d, g, z) when it appears as the second member of a compound (see Vance 1987, 2015; Rosen 2001, among many others). A well-known constraint on rendaku is that it may not apply in right-branching compounds (Otsu 1980; Kubozono 1988). The examples in (49) are taken from Otsu (1980) and Kubozono (2005).

(49)

a.	[nuri[hasi-ire]]	→	nuri-hasi-ire	'lacquered case for chopsticks'
	[nise[tanuki-siru]]	→	nise-tanuki-ziru	'raccoon dog soup that is fake'
b.	[[nuri-hasi]ire]	→	nuri-basi-ire	'case for lacquered chopsticks'
	[[nise-tanuki]siru]	→	nise-danuki-ziru	'soup made from a fake raccoon dog'

Ito and Mester (2007) reformulate this constraint such that rendaku may only apply in "minimal ω-projections". In other words, rendaku may only occur on the boundary of a p-word that does not dominate another p-word.

[38] See Chapter 3 for a discussion of the constraints on compound accent location.

[39] Although rendaku is generally considered as productive, it is also known for its many exceptions and irregularities (see Vance 2015 for a state-of-the-field review). For the sake of the argument, it is assumed here that rendaku applies productively.

(50)

a.	[nuri[hasi-ire]]	→	$[nuri]_{\omega} + [[hasi]_{\omega}\,[ire]_{\omega}]_{\omega}$
	[nise[tanuki-siru]]	→	$[nise]_{\omega} + [[tanuki]_{\omega}\,[\underline{z}iru]_{\omega}]_{\omega}$
b.	[[nuri-hasi]ire]	→	$[[nuri]_{\omega}\,[\underline{b}asi]_{\omega}]_{\omega} + [ire]_{\omega}$
	[[nise-tanuki]siru]	→	$[[nise]_{\omega} + [\underline{d}anuki]_{\omega}]_{\omega} + [\underline{z}iru]_{\omega}$

In (50a), the p-words *hasi-ire* and *tanuki-ziru* dominate lower-level p-words, and therefore are not subject to rendaku. Because *hasi* and *tanuki* do not dominate any other p-word in (50b), they undergo rendaku and surface as *basi* and *danuki*, respectively.

Further evidence for a role for p-words in rendaku comes from the behavior of the word *hon* 'book' when heading a compound (Kubozono 2005). As can be seen in the examples in (51), taken from Kubozono (2005), when the entire compound consists of four moras or less (51a), rendaku does not apply. However, when the entire word consists of more than four moras, the rendaku allomorph *-bon* shows up (51b).

(51)

a.	e-hon	'picture book'
	ero-hon	'erotic book'
b.	manga-bon	'comic book'
	tankoo-bon	'independent book'

Kubozono (2005) points out that the distinction between four moras or less and five moras or more is relevant in other areas of the phonological grammar of Japanese as well. Apart from the blocking of gemination and the /p/-/h/ alternation in Sino-Japanese compounds that we already looked at earlier in this section, Kubozono (2005) also mentions the accent patterns of mimetics and telephone numbers, as well as truncation patterns and metathesis in the secret jazz language. His conclusion is that for all of these phenomena a fundamental distinction exists between words of four or less and words of five or more moras, which suggests that the optimal p-word consists of four moras.

Leaving the topic of compounds, the way in which suffixes and particles are phrased in relation to p-words has not received much attention in the literature. Sato (2008: 14) defines the p-word as "as base plus certain prefixes, suffixes, and/or intensifier infix", where "base" means "the part of the word remaining when all affixes have been removed". However, he does not present convincing evidence in favor of his proposal, and neither does he discuss the behavior of particles.

Poser (1984) states that while the minor phrase is an important prosodic domain, the p-word certainly plays a role in the grammar. He argues that the p-word is "the domain of accentual dominance", and the domain of "final deaccenting" (Poser 1984: 144-145). His examples of accentual dominance are summarized in (52).

(52) The p-word as the domain of accentual dominance (based on Poser 1984)

a.	/miˈ-, -maˈsu/	→ [[mimaˈsu]$_{\omega}$]$_{\varphi}$	'see (polite)'
		(*miˈmasu, *miˈmaˈsu)	
b.	/yoˈn-de, miˈ-, -maˈsu/	→ [[yoˈnde]$_{\omega}$ [mimasu]$_{\omega}$]$_{\varphi}$	'try reading (polite)'
		*[yonde mimaˈsu]$_{\varphi}$	
c.	/yoˈn-de, miˈ-, -maˈsu/	→ [[yoˈnde]$_{\omega}$]$_{\varphi}$ [[mimaˈsu]$_{\omega}$]$_{\varphi}$	'read and see (polite)'

The form in (52a) shows that the lexical accent of the verbal root is overridden by the dominant accent of the politeness suffix *-maˈsu*. In (52b), in which the gerund of the verb to read (*yoˈn-de*) is followed the polite form of 'to see' (*mi-maˈsu;* here used in its grammaticalized sense of 'to try'), the accent of *-maˈsu* only has dominance over the root of 'to see', and not over the preceding form *yoˈnde*. This is argued to form evidence for a domain of accent dominance that is smaller than the p-phrase, namely the p-word. The form in (52c) shows what happens if the verb 'to see' is used in its lexical sense rather than the grammaticalized sense of (52b); the two verbal forms form not only independent phonological words, but also independent p-phrases (minor phrases), which are parsed into a single major phrase.

The argument from final accent deletion, which is actually due to McCawley (1968), goes as follows. Assuming the dative particle *-ni* is inherently accented, its deletion in (53a) as opposed to its retention in (53b) can be explained by the fact that *-ni* is final in the p-word in the former but not in the latter.

(53) a.	/tookyoo, -niˈ, sumi-maˈsu/	→	[tookyoo ni]$_{\omega}$ [sumimaˈsu]$_{\omega}$
			(*tookyoo niˈ sumimasu)
b.	/tookyoo, - niˈ, -wa/	→	[tookyoo niˈ wa]$_{\omega}$
			(*tookyoo ni wa)

According to Poser (1984: 210), this data is evidence that the p-word forms the domain of final accent deletion. However, the evidence is not convincing. The problem with the argument is that there is no evidence that the particle *-ni* is inherently accented. As discussed in Poppe (2012), the accent on the first of two consecutive particles is better

analyzed as inserted by the phonological grammar. Thus, in forms like those in (54), the accent that appears on the first of a sequence of two particles is not an inherent property of the accent, but a default boundary marker of a prosodic domain.

(54) a. sakana-ni ‘fish-DAT’
b. sakana-niˀ-wa ‘fish-DAT-FOC’
c. sakana-karaˀ-wa ‘fish-ABL-FOC’

In Poppe (2012), where I remain agnostic about the identity of the prosodic domain involved, I propose that the domains are assigned as in (55), where the accent falls on the domain boundary.

(55) a. [tookyoo-ni]
b. [[tookyoo-niˀ] wa]
c. [[tookyoo-karaˀ] wa]

However, a problem for this approach is that in certain cases the default accent also appears following an accented word, as in (56a). This suggests that particles may form a separate minor phrase (Maekawa and Igarashi 2007). According to the principle of Proper Bracketing (Ito and Mester 1992), phrasal boundaries must coincide with p-word boundaries. Hence, they cannot be part of the p-word that contains the noun. Therefore, I propose that the particle combinations are assigned the structure in (56b) or (56c).

(56) a. kyoˀoto-niˀ-wa ‘Kyoto-DAT-FOC’
b. {{[kyoˀoto]$_{\omega}$}$_{\varphi}$ {[[niˀ]$_{\omega}$ [wa]$_{\omega}$]$_{\omega}$}$_{\varphi}$
c. {{[kyoˀoto]$_{\omega}$}$_{\varphi}$ {[[niˀ]$_{\omega}$ wa]$_{\omega}$}$_{\varphi}$
d. {{[kyoˀoto]$_{\omega}$ [[ni]$_{\omega}$ wa]$_{\omega}$}$_{\varphi}$

An argument in favor of the representation that involves adjunction in (56c) could be that the particle *-ni* is promoted to a p-word as the result of focus, and that the focus particle *-wa* is simply adjoined to this p-word. However, this is mere speculating. Note that the deletion of a junctural accent would be the result of the second (recursive) p-word being incorporated in the phrase together with the noun, as in (56d).

The question that arises then is what happens when a noun is followed by a single particle. Is the particle incorporated into the p-word projected by the noun, is it directly adjoined to the node of the higher level of the p-phrase, or does it project its own p-word?

I propose that single particles are not incorporated into the p-word projected by the noun, but are adjoined to the phrasal node directly instead. The reason for this is that in this way, it is possible to account for the difference between dominant and recessive suffixes, as I will show now.

Building on Poser (1994), Nagahara (1994) argues that p-word structure is related to the differences between the three accentual types of affixes/particles identified by Poser (1984): dominant, recessive, and dependent. The three types of accent are listed in (57). The morpheme that "sponsors" the accent is the morpheme to which an accent "belongs" in the input form.

(57) Three types of morpho-accentual behavior (based on Poser 1984)[40]

a.	Dependent:	Shift of an accent sponsored by the root to(ward) the affix
b.	Recessive:	Realization of an accent sponsored by the affix in the absence of an accent sponsored by the root
c.	Dominant:	Deletion of an accent sponsored by the root (and depending on the affix, the realization of an accent sponsored by the affix)

For the sake of illustration, in (58) examples of each type of morpho-accentual behavior are given.

(58)

a.	-ya 'store'	soˀba + ya	→	sobaˀya
		toofu + ya	→	toofuya
b.	-made 'even'	kyoˀoto + maˀde	→	kyoˀoto made
		tookyoo + maˀde	→	tookyoo maˀde
c.	-guˀrai 'about'	kyoˀoto + guˀrai	→	kyooto guˀrai
		tookyoo + guˀrai	→	tookyoo guˀrai

After pointing out that only recessive suffixes are subject to a rule that can also be observed at the phrasal level which deletes all accents but the leftmost one, Nagahara (1994) proposes that dominant accentuation involves the deletion of all but the rightmost accent within a p-word, and dependent accent the linking of a floating pitch accent within a p-word. A consequence of this analysis is that verbal suffixes with a floating pitch accent

[40] It should be noted that there is variation in the location where the inserted tone is linked (e.g. pre-accenting, post-accenting, etc.). A full account of accentual dominance in Japanese lies beyond the scope of this thesis. See Rosen (2008) for an analysis of the accentual behavior of different types of suffixes in terms of tonal specification.

in his analysis like *-tari*$_{acc}$ (‘alternative’) must be assumed to form a separate phonological word. This cannot be the case, however, as this would mean that in forms like *kaʾttari* (*kat*$_{acc}$*-* + *tari*$_{acc}$) ‘win-ALTERNATIVE’, a geminate consonant would be realized across a p-word boundary, something which is not allowed in (at least Tokyo) Japanese (Ito & Mester 1996). In Chapter 6, an alternative analysis of basic verbal forms including *-tari* is proposed in which the p-stem plays an important role.

Still, there are at least two insights of Nagahara’s (1994) analysis that remain valid. First of all, the idea that depending on whether a particle is parsed in the same p-word as its host, its behavior in terms of accent may differ, remains valid. Thus, particles may show variable behavior in terms of phonological phrasing, which is reflected in the realization of underlying accents. Second, the idea that dominant accentuation involves the parsing of a suffix into a p-word together with its host, which is also proposed by E. Han (1994: 217), can be adopted. Both of these facts can be illustrated by means of the examples in (59).

(59)

a. [[kyoʾoto]$_{\omega}$ [gurai]$_{\omega}$]$_{\varphi}$ → no dominance by particle; leftmost accent in p-phrase survives

b. [[kyooto guʾrai]$_{\omega}$]$_{\varphi}$ → dominance by particle; rightmost accent in p-word survives

If we assume that only suffixes and particles that shown dominant behavior are parsed into the same p-word as the noun, we can account for the distinction between dominant and recessive suffixes and particles. Therefore, it is better to assume that non-dominant affixes are not parsed into the same p-word as nouns.

In Chapter 6, it will be shown that the approach to junctural accent proposed here also works for the accent patterns of certain verbal forms, namely contracted forms of the gerundive followed by the existential auxiliary verb *i-*.

2.3.3. Prosodic stem

In the same way that p-words are the result of the mapping of morphological words to phonological structure, the prosodic stem (p-stem) is the unit of the prosodic hierarchy that corresponds to the morpho-syntactic stem (m-stem). Although in many studies it is assumed that the p-word is the smallest interface category, Inkelas (1989, 1993) and Downing (1998a/b, 1999, 2006) argue that m-stems map onto p-stems, and morpho-syntactic roots (m-roots) may map onto prosodic roots (p-roots).

The p-stem has been proposed for Japanese in a limited number of studies.[41] Adopting the theory of prosodic constituents proposed by Inkelas (1989, 1993), E. Han (1994) argues in favour of two types of compounds in Japanese: those consisting of two p-roots, and those consisting of two p-stems. Compounds consisting of p-roots are words consisting of two Sino-Japanese roots, e.g. *seki-yu*, which literally means 'stone-oil', i.e. 'petroleum'. E. Han (1994) argues that the "fusion" or place assimilation between the coda of a Sino-Japanese root and the onset of the following Sino-Japanese root is only possible within compounds in which the first root is parsed into a p-root (R), as in (60a), and not a p-stem (60b).[42]

(60) a.	[bet] [sit]	→	[[[be<u>s</u>]$_R$[situ]$_R$]]$_\psi$]$_\omega$
	'separate' 'room'		'separate room'
b.	[toku-bet] [kyuu-koo]	→	[[[toku]$_R$-[bet<u>u</u>]$_R$]$_\psi$[[kyuu]$_R$[koo]$_R$]$_\psi$]$_\omega$
	'special' 'express'		'special express'

E. Han (1994) is right that we must be dealing with different domains, but it is not clear whether these are the p-root and the p-stem. The word *toku-betu* 'special' may stand by itself, so it is not clear why this should be considered a stem or p-stem. It is more natural to assume that morphological words are mapped onto p-words. This means that the word in (60b) can be considered a recursive p-word rather than two p-stems. This is the position taken in Ito and Mester (1996, 2015a), who assume that Sino-Japanese roots are mapped onto feet rather than p-roots.[43] The difference between the analyses of E. Han (1994) and Ito and Mester (1996, 2015a) may appear to only involve differences in the names of the different constituents. However, the p-root is not independently motivated, i.e. it does not form a domain to which a certain phonological process exclusively applies. Therefore, an analysis in terms of recursive p-words is preferred.

---

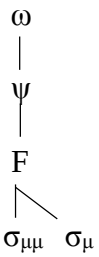
[41] As explained above, Ito (1990) refers to the stems in prosodic morphology as prosodic stems. However, because there are even prosodic suffixes in her approach, what she calls (prosodic) stems seems to refer to stems in the prosodic morphology rather to than a constituent within the Prosodic Hierarchy.

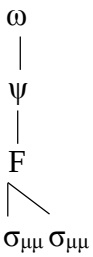
[42] Note that E. Han (1994) assumes that Sino-Japanese roots underlyingly end in a consonant.

[43] In E. Han (1994) the root maps onto a p-root because if it should map on some prosodic constituent in Inkelas' (1989, 1993) theory, it must be the p-root. In Ito and Mester (1996, 2015a) it maps onto a foot due to a constraint that maps Sino-Japanese roots onto feet. The motivation behind this constraint is that Sino-Japanese morphemes may not be larger than two moras. The question remains though why monomoraic Sino-Japanese morphemes are parsed into feet. One possible explanation is that the mapping of bimoraic roots onto a foot has been extended to monomoraic roots by analogy. In fact, the existence of feet in Japanese may be the result of the way in which Chinese loanwords were adapted in the Japanese languages, i.e. as structures consisting of either one light syllable, and in many cases one heavy syllable or two light syllables.

In the discussion of the foot in Section 3.2.3, the possible role of the p-stem in the prosodic morphology of Japanese was briefly discussed. One study in which reference to p-stems is actually made is that by Rodier (1998). Rodier (1998) proposes output parsings of the following type for loanword truncations, in which the p-stem dominates the foot (cf. the representations in (28) above).

(61) a. 

b. 

c. 

According to Rodier (1998), p-words consisting of a single heavy syllable are ruled out by a constraint that forces the head of a p-stem to be "in a rhythmic relation with a non-head constituent at some level of analysis" (i.e. syllable or foot). As it is not made clear what exactly this "rhythmic relation" is, it is interesting to take a look at why Rodier (1998) prefers this constraint above the Ito and Mester's (1992) Word Binarity constraint.

Rodier (1998) says that the Word Binarity constraint proposed by Ito and Mester (1992) is problematic for two reasons. The first reason is that apart from a minimality constraint, it is at the same time a maximality constraint. It is not clear, however, why this would be a problem. The second reason he gives is that in clippings of the rustic girls' names type, in forms like those in (62b), a p-word consisting of a single heavy syllable follows the honorific prefix *o-*, violating Word Binarity.

(62) (Poser 1990: 93)

	Derivative	Base
a.	o(yuki)	(yuki)ko
	o(hana)	(hana)ko
	o(saki)	(saki)ko
b.	o(yoo)	(yoo)ko
	o(kei)	(kei)ko
c.	o(mido)	(mido)ri
	o(kae)	(kae)de

According to Rodier (1998), the prefix cannot be assumed to be parsed into a minimal p-word, as this would entail a violation of Left Edge Matching. However, this is not necessarily a problem if a constraint that regulates the alignment of the prefix to the left

of a base outranks the Left Edge Requirement. While it cannot be ruled out that, as assumed by Rodier (1998), the stem to which the honorific prefix attaches is actually parsed into a p-word,[44] his analysis is based on an ad-hoc constraint, and he does not give conclusive evidence for the parsing proposed in (61). Therefore, we cannot but conclude that there is no conclusive evidence for a role of the p-stem in the prosodic morphology of Japanese.[45]

Rodier (1998) is not the only researcher who has proposed a role for both the foot and the p-stem. Oda (2006) proposes that both constituents play a role in the accentuation of Tokyo Japanese. He adopts the distinction between the prosodic and metrical hierarchies of Inkelas (1989, 1993) in order to explain deaccentuation in compounds with short second members (N2s). His analysis, which is quite complicated, can be roughly summarized as follows.

Oda (2006) starts with the observation that deaccentuation of compounds may only occur when the N2, the head of the compound, is short, i.e. consists of two moras or less. Oda (2006) assumes that the distinction between long and short N2s can be captured in terms of foot structure (as in Poser 1990; see Chapter 3); the head consists of one foot in compounds with short N2s, and of more than one foot in compounds with more than one foot. Oda (2006) proposes that the foot of the metrical plane is mapped onto the p-stem of the prosodic plane by means of the alignment constraint in (61). The alignment is graphically represented in (62). Note that in Oda's (2006) analysis, the p-stem does not (necessarily) correspond to a morpho-syntactic stem; it is used as a device that mediates between feet and the p-word.

(61) ALIGN(F, P-stem)

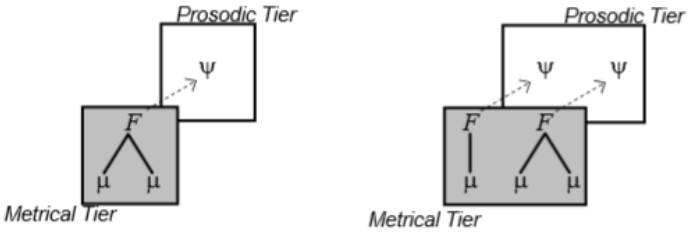
A unit of [the] foot level corresponds with a unit of [the] [p]-stem level.

(Oda 2006: 39)

[44] Note that at least the prefix and the stem together may be assumed to form a p-word, because in the related geisha/bargirl client names in which both *-o* and a suffix *-san* are added, the stem may not end in a geminate (**o-hos-san* vs. *o-hoo-san*, *o-hon-san* < Honda), which suggests that the constraint against geminates across p-word boundaries is at work.

[45] Downing (2006) gives sketches of analyses of simplex and compound loanword truncation in Japanese in which she refers to p-stems, p-roots, and even p-root-compounds. However, she only shows that in her model such analyses are possible, without providing independent evidence for the existence of the proposed constituents.

(62) a. Short N2 b. Long N2 (Oda 2006: 40).



The constraint in (61) only ensures that the foot and the p-stem are mapped onto each other, without a relation of dominance being involved. The constraint only demands that every foot needs to be mapped onto a p-stem at the prosodic tier. Because in compounds with long N2s the mapping process results in two p-stems that exist on the prosodic tier, these two p-stems are parsed into a p-word. In other words, the result of the mapping is an N2 that constitutes a p-word by itself. The consequence of this is that the compound will consist of a recursive p-word, which is marked by an accent. In compounds with short N2s, on the other hand, the result of the mapping process is a single p-stem on the prosodic tier. This p-stem is dominated by a p-word that also dominates the material of the first member of the compound (N1), the result of which is a non-recursive p-word that lacks an accent.[46]

Oda's (2006) analysis is interesting, but it does raise a lot of questions. First, it is based on the idea of strict layering, according to which all moras or syllables must be parsed into feet.[47] It may be clear that independent evidence for this is necessary, especially in the light of the arguments for weak layering by Ito and Mester (1992). What is more, although his analysis provides us with an explanation of why compounds with a long N2 do not appear without an accent,[48] Oda (2006) himself admits that it remains unexplained why only certain short N2s with final accent in their isolation forms cause compounds to be unaccented. Also, it is not clear why for some short N2s compounds only surface without an accent when N1 is bimoraic as well. Thus, we may conclude that more research is necessary to confirm whether the p-stem must be admitted as a part of the phonology of Japanese.

[46] The internal structure of N1 is assumed to be invisible due to head-dependent assymetries (Dresher and van der Hulst 1998).
[47] Oda (2006) explains that it is not clear whether feet dominate moras or an intermediate level of syllables.
[48] As shown in footnote 51 and Chapter 3, compounds with "overlong" N2s may surface without an accent though.

2.3.4. Prosodic compound

The final constituent discussed in this chapter is the prosodic compound (p-comp; see also Peperkamp 1997), a constituent that in general is not considered to be part of the Prosodic Hierarchy.

As Alderete (1999) notes, the accent in Tokyo Japanese simplex nouns may fall on the final syllable, but this is not allowed in compounds.[49] According to Alderete (1999:88), if we assume syllable extrametricality (NONFINALITY; Prince and Smolensky 1993) holds for p-comps but not for p-words, this difference can be explained. It is not possible to restate the generalization in terms of the difference between minimal and maximal projections of the p-word, because the isolation form of a noun may also be parsed into a maximal p-word. However, Alderete's (1999) analysis has a number of weaknesses. The first is that in constructions with accent-shifting suffixes, which Alderete (1999) analyses in terms of recursive p-words, the same NONFINALITY constraint is necessary. Therefore, he proposes that NONFINALITY refers to the highest category in the hierarchy above the innermost p-word (Alderete 1999: 233). Obviously, this weakens the analysis.

One more potential problem for the p-comp-based analysis is that it fails to account for the similarities in accentual behavior between compounds and loanwords (see Kubozono 2006). Although in exceptional cases final accent is observed in loanwords, the non-finality principle seems to hold to a large extent for syllables in loanwords as well (Kubozono 1996). The following chapter contains a detailed discussion of loanword and compound accent in Tokyo Japanese.

2.4. Summary

In this chaper I have given an overview of word-level prosody in Japanese. In the first part of the chapter, a comparison of different approaches to Japanese pitch accent was made, partly in order to justify the autosegmental(-metrical) approach taken in this thesis. Following this, a critical evaluation of the evidence for several prosodic constituents was presented.

In the remainder of this thesis, the mora, syllable, p-word, and p-phrase, for which I showed there is ample evidence, will all play important roles. Furthermore, it will be shown that both the metrical foot and the p-stem play important roles in the accentual grammar of different dialects.

---

[49] Final accent is allowed in compounds with an N2 like *saibansyo*ˀ 'courthouse', e.g. *saikoo-saibansyo*ˀ 'Supreme Court'. The N2 in this word consists of the Sino-Japanese root compound *sai-ban* 'trial' and the dominant suffix *-syo*ˀ 'place'.

3. Metrical and tonal approaches to Tokyo Japanese accent

In this chapter, metrical and tonal approaches to Tokyo Japanese accent are compared. Although in several studies it has been suggested that (Tokyo) Japanese accent can be assumed to be an underlying H or HL tone (cluster) (Poser 1984; Pierrehumbert & Beckman 1988; Hyman 2006, 2009), no serious attempts have been made, apart from the study by Clark (1986), at a purely tonal analysis of the pitch accent system of Tokyo Japanese. In this chapter, a purely tonal account of Tokyo Japanese accent is developed and compared with the metrical analyses proposed in the literature. The extent to which the different approaches are able to account for the phenomena discussed in a theoretically consistent way will also be discussed.

In the next section, relevant data and descriptive generalizations are introduced. Following this, different metrical analyses that have been propoed for (parts of) the Tokyo Japanese accent system are reviewed. Next, it is investigated how far we can get with a purely tonal analysis of Tokyo Japanese. The chapter concludes with a summary of this chapter. It will be concluded that although without reference to metrical structure, some phenomena cannot be accounted for in a unified way, a number of issues concerning the exact role of foot structure and tone remain unresolved.

3.1. Data and descriptive generalizations

3.1.1. Simplex nouns

Tokyo Japanese (TJ) is a well-documented variety of Japanese that has played an important role in studies of accent. As can be seen in (1), the nominal accent system is an example of a system with *n+1* patterns for nouns with *n* moras (or syllables).[50] Unaccented words are indicated with a superscript "0".

[50] Because, in heavy syllables the accent in general is not allowed to fall on a special mora in Tokyo Japanese, the *n*+1 generalization basically holds for syllables as well.

(1) Tokyo Japanese accent system

	1 mora	2 moras	3 moras	4 moras
0	ha^0	kawa^0	sakura^0	tomodachi^0
	‘tooth’	‘river’	‘sparrow’	‘friend’
1	teˀ	soˀra	kaˀrasu	kaˀmakiri
	‘hand’	‘sky’	‘crow’	‘mantis’
2		otoˀ	tamaˀgo	muraˀsaki
		‘sound’	‘egg’	‘purple’
3			otokoˀ	asioˀto
			‘man’	‘footsteps’
4				otootoˀ
				‘younger brother’

As shown in Chapter 2, the Tokyo Japanese accent is realized by a HL tonal complex of which the H part is associated with the accented mora. The pitch patterns can be schematically represented in terms of H and L tones as in (2). However, as pointed out in Chapter 2, in this thesis it is assumed that the surface tonal representations may be underspecified as in (3), where tones marked with “%” are phrasal boundary tones.

(2) Pitch patterns of Tokyo Japanese nouns

kaˀrasu(-ga)	HˀLL(L)	‘crow(-NOM)’
tamaˀgo(-ga)	LHˀL(L)	‘egg(-NOM)’
otokoˀ(-ga)	LHHˀ(L)	‘man(-NOM)’
sakura(-ga)^0	LHH(H)	‘cherry blossom(-NOM)’

(3)

sakura(-ga)	karasu(-ga)	tamago(-ga)	otoko(-ga)
L%H% L%	L%HL L%	L%HL L%	L%H%HL L%

Now, while it is true that there are *n+1* patterns for words with *n* moras (or syllables), the distribution of the different types of words in the lexicon shows clear preferences for certain patterns. This is clear from statistical data on the distribution of accents in (4), where “f” stands for “familiarity”, measured on a seven-point scale (1 = “least familiar”, 7 = “most familiar”). The data is adapted from Ito and Mester (2012), who made numeral approximations of graphs presented in Kitahara (2001).[51] Apart from the distribution of

[51] The data in Kitahara (2001) in turn is based on the lexical database in Amano and Kondo (1999).

the different accent patterns for words of two, three, and four moras, the data also includes the relation between word familiarity and the distribution of the accent patterns. Note that for 2μ words, higher familiarity is correlated with unaccentedness, whereas for 3μ and 4μ words, less familiar words tend to be unaccented compared to more familiar words.

(4) a. Distribution of accent patterns in *2μ-words* with different familiarity rates

	-2	-1	0
f > 3	70.9%	17.5%	11.7%
f > 4	68.9%	18.4%	12.6%
f > 5	62.1%	22.3%	15.5%
f > 6	50.5%	29.1%	20.4%

b. Distribution of accent patterns in *3μ-words* with different familiarity rates

	-3	-2	-1	0
f > 3	36.9%	9.7%	6.8%	46.6%
f > 4	35.9%	10.7%	7.8%	45.6%
f > 5	35.0%	11.7%	8.7%	44.7%
f > 6	36.9%	17.5%	9.7%	35.9%

c. Distribution of accent patterns in *4μ-words* with different familiarity rates

	-3	-2	-1	0
f > 3	8.7%	7.8%	11.7%	69.9%
f > 4	9.7%	7.8%	11.7%	68.9%
f > 5	10.7%	8.7%	12.6%	65.0%
f > 6	13.6%	13.6%	15.5%	53.4%

Leaving word familiarity aside, the data in (4) shows us that there is a strong correlation between word length and accent pattern. To start with bimoraic words, initial accent is the most frequent pattern. Among trimoraic words, on the other hand, both words with antepenultimate accent and unaccented words are frequent. Finally, in quadrimoraic words, the most frequent pattern is clearly that of unaccented words.

For trimoraic words, similar data can be found in Kubozono (2011). The data from Kubozono (2011), which is presented in (5), provides us with information of the distribution of accent patterns across three different lexical strata: native Japanese, Sino-Japanese (SJ), and loanwords. The “popularity” of words with antepenultimate accent in trimoraic loanwords in (5) is striking.

(5) Word type and accent pattern in 3-mora nouns in Tokyo Japanese (slightly adapted from Kubozono 2011: 2885)

		-3	-2	-1	0
Native	(# = 2,220)	17%	9%	3%	71%
SJ	(# = 4,939)	47%	1%	1%	51%
Loanwords	(# = 778)	89%	2%	2%	7%
Total	(# = 7,937)	42%	4%	2%	52%

The productivity of antepenultimate accent in loanwords prompted McCawley (1968), who also refers to Tashiro (1953) for the observation that in "meaningless" strings such as a row of kana symbols as in *ka ki kuˈke ko*, the accent also falls on the antepenultimate mora, to propose a rule according to which the accent is placed on the syllable containing the antepenultimate mora (McCawley 1968: 134). Reference to the syllable is necessary to account for the fact that when the third from last mora is a special mora, the accent is placed on the immediately preceding regular mora (e.g. *erebeˈetaa* 'elevator').

When it comes to loanwords of 4 moras, the situation is more complicated. While in quadrimoraic accented words the accent is indeed assigned on the antepenultimate mora, the unaccented pattern can also be observed frequently. As demonstrated by Kubozono and Ogawa (2004), this is especially true for words consisting of four light syllables. Data on the accent patterns of foreign place names adapted from Kubozono and Ogawa (2004) in (6) exemplifies this. The numbers in parentheses show the actual number of examples.

(6)	σμσμσμσμ	σμμσμσμ	σμσμμσμ	σμσμσμμ	σμμσμμ	Total
+acc	33%	61%	95%	88%	100%	64%
	(26)	(20)	(38)	(30)	(21)	(135)
-acc	67%	39%	5%	12%	0%	34%
	(52)	(13)	(2)	(4)	(0)	(71)

As we can see in (6), in quadrimoraic words that do not end in two light syllables, unaccented words are infrequent. In words ending in two light syllables, on the other hand, there are more unaccented words, and this is actually the dominant pattern in words consisting of a total of four light syllables.

If we add to this the fact that monomoraic words are accented around 70% of the time (NHK 1985, cited by Cutler and Otake 1999), and furthermore that words of five or more moras are accented in most cases (Kubozono 2011), the most frequent accent patterns for words of different lengths can be summarized as in (7).

(7)

μ	→	μˀ
μμ	→	μˀμ
μμμ	→	μμμ^{0} / μˀμμ (depending on the lexical stratum)
μμμμ	→	μμμμ^{0} / μμˀμμ (depending on syllable structure)[52]
≥μμμμμ	→	μμμˀμμ

If we distinguish between "accentedness" and accent location (see Poser 1984; Kubozono 2008, 2011), we may state the generalizations as follows. In words of any length except in quadrimoraic words, having an accent is preferred to having no accent. In words of three moras or more with an accent, the accent is placed on the antepenultimate mora (that at the same time is a syllable head). Finally, because in words of less than three moras there is no antepenultimate mora, the accent falls on the leftmost mora available.

3.1.2. Compound nouns

Next, we turn to the accent patterns of compounds. The examples in (8) are adapted from Uwano (1997b) and Kubozono (2011). The left-hand member of a compound is referred to as N1, and the right-hand member as N2.

(8)

	N1	N2	Compound	Gloss
a.	kaˀbuto	musi^{0}	kabutoˀ-musi	'beetle'
	aˀkita	inuˀ	akitaˀ-inu	'Akita dog'
	niˀngyo	hiˀme	ningyoˀ-hime	'Little Mermaid'
	peˀrusya	neˀko	perusya-neˀko	'Persian cat'
b.	yamaˀ	sakura^{0}	yama-zaˀkura	'mountain cherry blossom'
	iroˀ	otokoˀ	iro-oˀtoko	'playboy'
	usiro^{0}	suˀgata	usiro-suˀgata	'appearance from behind'
c.	onnaˀ	tomodachi^{0}	onna-toˀmodachi	'female friend'
	ne^{0}-	syoogatuˀ	ne-syoˀogatu	'staying at home during the New Year holiday'
	saki^{0}	ototoˀi	saki-ototoˀi	'three days ago'
	edo^{0}	muraˀsaki	edo-muraˀsaki	'royal purple'
	mesuˀ	kaˀmakiri	mesu-kaˀmakiri	'female mantis'

[52] Tanaka (2008) shows that in accented words consisting of four light syllables, the pre-antepenultimate pattern is becoming more frequent. See below in 3.2.2 and 3.3.2, where the influence of syllable structure is also discussed

Compound accentuation in Tokyo Japanese is generally assumed to be determined by the accent pattern of the second member (N2) (Hirayama 1960; McCawley 1968). A well-known descriptive generalization that has been proposed for compound accentuation is based on the difference in behavior between words with an N2 of one or two moras, and those with an N2 of three moras or more (Hirayama 1960; McCawley 1968). However, even among words consisting of two moras or less, as well as among words of three moras or more, different classes can be identified (Uwano 1997b).

First of all, compounds with N2s of one or two moras differ from those with N2s of three or more moras in that in the former but not in the latter, properties of N1 have an influence on accent placement. This situation is summarized in (9), which is adapted from Uwano (1997b: 241).[53]

(9) N2 = 1μ, 2μ

a. N2s that take a single pattern regardless of the accent pattern of N1
   e.g. *-iro*^{0} 'color' (deaccenting), ..μ᾽-*si* 'city' (pre-accenting)
b. N2s that take different accent patterns depending on whether N1 consists of two moras or less, or three moras or more.
   e.g. *i᾽to* 'thread': *katan-i᾽to* 'cotton thread' vs. *kinu-ito*^{0} 'silk thread', *ke-ito*^{0} 'knitting wool'
c. N2s whose accent patterns depends on the length of N1, while the accent pattern of N1 determines the accent pattern of the compound only in cases in which N1 consists of two mora or less.
   e.g. *ha᾽* 'tooth': *itokiri᾽-ba* 'canine tooth', *ma᾽e-ba* 'front tooth', *musi-ba*^{0} 'decayed tooth'[54]

For N2s whose behavior is the same for all compounds, the generalizations in (10) hold (adapted from Uwano 1997b).

(10) a. N2 = 1μ
   (i) N1-final accent in general.
   (ii) Certain N2s with an accent in isolation behave as deaccenting morphemes in compounds.

[53] Uwano's (1997b) discussion contains a lot more information on further subdivisions and variation. However, the generalizations on compounds with N2s of one or two moras in (9) and those with N2s with three or more moras in (10) will do for our purposes.

[54] The bilabial voiced stops here are the result of rendaku (see Chapter 2).

b. N2 = 2μ
   (i) N1-final accent in general.
   (ii) Many N2s with initial accent in isolation keep their accent in a compound. (μˈμ → N1-μˈμ)
   (iii) Many N2s with final accent in isolation behave as deaccenting morphemes in compounds. (μμˈ → N1-μμ^0)

c. N2 = 3μ
   (i) N2s that carry an accent on the second syllable in isolation, keep their accent in a compound.[55]
   (ii) All other N2s are initially accented in a compound.[56]

d. N2 = 4μ
   (i) N2s that carry an accent on the second or third syllable in isolation, keep their accent in a compound.[57]
   (ii) All other N2s are initially accented in a compound.

Finally, compounds with an N2 of five moras or more behave differently from all other compounds (Kubozono, Ito and Mester 1997; Uwano 1999a). The examples in (11) are taken from Kubozono, Ito and Mester (1997).

(11)a. minami^0 + kariforunia^0	→	minami-kariforunia^0	‘South California’
		*minami-kaˈriforunia	
b. nyuˈu + karedonia^0	→	nyuu-karedonia^0	‘New Caledonia’
		*nyuu-kaˈredonia	
c. daˈi + sakusika^0	→	dai-sakusika^0	‘great songwriter’
		* dai-saˈkusika^0	
d. gakusyuu^0 + sankoosyoˈ	→	gakusyuu-sankoosyoˈ	‘study reference
		*gakusyuu-saˈnkoosyo	book’

As we can see from the data in (11), in compounds in which the N2 is longer than five moras, no default compound accent is assigned. Kubozono, Ito and Mester (1997) do

[55] In the speech of younger generations, trimoraic words with penultimate accent in isolation tend to be realized with an N2-initial accent in compounds (Uwano 1997b: 238).
[56] As pointed out by Uwano (1997b: 265, endnote 8), we could also say that N2s with initial or second accent keep their accent, and unaccented words and words with final accent get a default N2-initial accent.
[57] Like trimoraic words with penultimate accent, quadrimoraic words which have penultimate accent, in isolation tend to be realized with an N2-initial accent in compounds (Uwano 1997b: 238).

mention counterexamples to this generalization, e.g. *etiru-a*ꜜ*rukooru* 'ethyl alcohol'. However, they add that there are also speakers who pronounce this word as *etiru-arukooru*^{0}, where the unaccented status of *arukooru*^{0} 'alcohol' is preserved.

This brings us to one serious problem that any analysis of compound accentuation in Tokyo Japanese needs to deal with: the considerable amount of variation between generations and also across individuals. As a result of this variation, it is not always clear to which group a noun belongs. An exhaustive statement of the data, let alone a fully worked out analysis of Tokyo Japanese compound accentuation, lies beyond the scope of this study. Still, when comparisons between two different types of analyses are made, we will try to take issues of variation into account as well.

Summarizing, the accentuation of compound nouns resembles that of simplex nouns in that the observed accent patterns are correlated with word length. On the face of it, this suggests that prosodic constituency might play a role in the accentual grammar. Indeed, several approaches in which the metrical foot plays a role have been proposed. These approaches are evaluated in the following section.

3.2. Metrical approaches to Tokyo Japanese accent

In the literature on Tokyo Japanese accentuation, metrical foot structure has been used to account for five different phenomena.

(12) a. Surface tonal patterns (Abe 1981; Zubizarreta 1982, among others)
b. Default antepenultimate (-3) accent (Yamada 1990, Tanaka 1992; Katayama 1998, Shinohara 2000, Tanaka 2008, Labrune 2012a/b, among others)
c. Compound accentuation (Poser 1990, Kubozono 1995, 1997, 2011, Tanaka 2001, Tanaka 2008, Labrune 2012, among others)
d. Unaccentedness (Ito and Mester 2012);
e. Domain of accent shifts (Haraguchi 1991; Tanaka 1992, 2002, 2005).

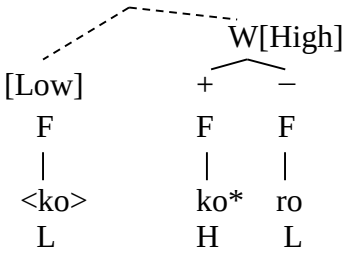
Although the foot-based approach to Tokyo Japanese accent has gained popularity with the advent of OT-based approaches to the phenomenon, in this section foot-based approaches couched within derivational models are also discussed. In order to enable a comparison, the focus will lie on the surface structures rather than how the surface forms are derived. Still, in situations in which the difference between derivational and non-derivational models is of importance, this will be indicated. The phenomena will be discussed in the order of (12).

3.2.1. Deriving the surface tonal patterns from metrical structure

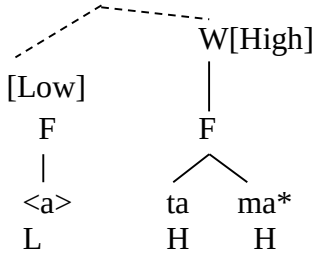
The first metrical analyses of Tokyo Japanese refer to unbounded rather than binary feet. In these analyses, feet consist of a head that dominates an unbounded number of dependents. Abe (1981) and Zubizarreta (1982) propose unbounded feet to account for the tonal patterns of a number of Japanese dialects, reanalyzing data taken from Haraguchi (1977). The analyses by Abe (1981) and Zubizaretta (1982) have a lot in common, and it is the latter that I will dicuss here.[58]

In Zubizarreta's (1982) analysis of Tokyo Japanese, right-headed metrical trees (unbounded feet) are constructed that branch leftward, incorporating all moras except for the word-initial one, which is extrametrical (< >). The material to the left and right of the accented foot is also incorporated into right-headed feet. Some examples of trimoraic words are given in (13).

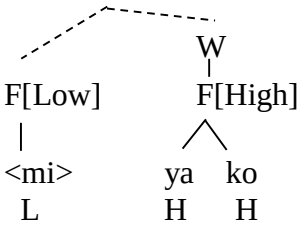
(13) a.



b.



c.



Tonal assignment in this analysis works as follows. In accented words, H tone is assigned at the W(ord)-level. The H tone looks for the head foot (+) of the word, and associates with the material within this foot. The foot indicated with (-) is the polarized counterpart of the head foot, and is subject to the association of the minus value of the H tone [-High]. Furthermore, the initial mora[59] is incorporated in a separate foot which is assigned a L tone. In unaccented words (12c), there cannot be any distinction between head and non-head feet before the initial extrametrical mora is incorporated in the word tree in the final

[58] See Poser (1984) for a critical discussion of Abe (1981).

[59] I refer to moras here for the sake of convenience. Zubizaretta actually does not use this term and states that the leftmost [+syllabic] segment is extrametrical.

stage of the derivation. Hence, H is assigned to the rightmost foot, and again the tone spreads to all material within this foot. As in non-initially accented words, the extrametrical mora and its foot are incorporated in the word tree and assigned a L tone.

Zubizaretta (1982) does not really give arguments in favor of a metrical-tree approach to Japanese accent. Rather, it seems like an exercise in the application of a theory of stress to different dialects of a type of language that lies "between tone and accent" (Zubizarreta 1982: 170). In his discussion of Abe (1981) and similar approaches, Poser (1984) points out that the arguments in favor of a metrical approach based on unbounded feet are not convincing. That is, there are no clear reasons to adopt the proposed tree structure if an autosegmental account is available.

The studies by Idsardi & Kim (1997), Idsardi & Purnell (1997), Purnell (1998), and Kim (1999) should also be mentioned here. These studies assume the simplified bracketed grid (SBG) model developed by Idsardi (1992) and Halle and Idsardi (1995). In this model, two types of foot boundaries (left brackets and right brackets) are assumed, which function more like traditional boundary symbols rather than as an alternative notation for arboreal constituency (see van der Hulst 2009). The theoretical option of open constituents (with a single boundary), which is used for unaccented words in Tokyo Japanese (14b), for instance, illustrates this.

(14)a.

	x	line 1
	x (x x) x	line 0
	kamisori'-ga	'razor-NOM'
	L H H L	

b.

	x	line 1
	x (x x x	line 0
	kamaboko-ga^{0}	'fish pudding-NOM'
	L H	

The H tone is linked to the left of right brackets and to the right of left brackets. Thus, the brackets serve as delimiting the domain of tone association rules. In this sense, the analysis is similar to that of Zubizarreta (1982) discussed above.

The point of the above discussion is that if we assume unbounded feet, we can assume these to function as domains of tone assignment (or, in the case of the SBG model, as boundaries to which tones are aligned). A question to consider then is whether unbounded feet are also strictly necessary to account for the tonal patterns. The answer to this question seems "no", since purely autosegmental and autosegmental-metrical studies have shown that purely tonal accounts suffice. Arguments in favor of unbounded domains then should come from patterns of accent assignment or accent shift. Because such evidence does not exist, however, positing unbounded feet is not an attractive option.

In most studies of Japanese accent that refer to the bimoraic foot, the surface tonal

patterns are not dealt with.[60] In more theoretrically oriented studies, the impression is that the realization of the accent is seen as something secondary, of minor importance to the phonological grammar.

3.2.2. Default antepenultimate accent

Yamada (1990) presents a detailed analysis of Tokyo Japanese using the framework of of Halle and Vergnaud (1987). Based on the idea that the default accent for Tokyo Japanese nouns is antepenultimate, he proposes the parameter settings in (15a). An example in the form of the derivation of the word *miˈdori* 'green' (Yamada 1990: 6) is given in (15b). Parenthesis indicate foot boundaries in general; square brackets indicate the headedness of feet on line 0. In the output form, the accent location is added for the sake of clarity.

(15)a. line 0: binary, left-headed, right-to-left; mora extrametricality (nouns)
line 1: unbounded, right-headed
line 2: head of line 1 is copied; line 0 and 1 are conflated.

b.

				*	.	.	*	.	.		line 2
		*	.	.	(*	.)	.	(*	.)	.	line 1
		[*	*)	.	[*	*)	.	[*	*)	.	line 0

midori→mido<ri>→ mido <ri>→ miˈdo <ri>

The idea of conflation is important in Yamada's (1990) analysis, as it is necessary to erase all stress marks except the ones in the column above the accented syllable. Without this mechanism, it is predicted that secondary stress would surface. This is exemplified in (16), where the only operation on line 3 consists of the copying of the head (left) of line 2. The particle *-guˈrai* is lexically accented, whereas *sakura*^{0} 'cherry blossom' is unaccented.

(16)

				.	.	.	*	.	.	.	.	.	*	.	.	line 3
	*		(.	.	.	*	*	.)	(.	.	.	*	.	.)	line 2
	*		(*	*	.	*)(*	.	)	(.	.	.	*)	.	.	line 1
	*		[*)[**		)[*)[**				*	*	*	***			line 0

sakura-gurai → sakura-gurai → sakura-guˈrai 'at least a cherry blossom'

[60] A notable exception is the study by Asano (2002), who proposes foot structure to account for the location of accents and "tonal domains" to account for the surface tonal patterns. For dialects other than Tokyo Japanese, Lawrence (1990) makes an interesting proposal in terms of foot extrametricality to account for the surface tonal patterns of Ibukijima Japanese. Lawrence (1990) also shows foot structure is partly responsible for the surface tone patterns in the Nakijin dialect of Okinawan, a language that belongs to the Ryukyuan branch of Japonic.

The example in (16) shows that in the case of unaccented words, a lot of structure is built that subsequently needs to be erased again. The only reason for this is that unaccented words cannot be assumed to be exempt from the parameters proposed to account for accented words. However, it may be clear that any analysis that does not need an uneconomic derivation like that in (16) should be preferred.[61] In any case, Yamada (1990) shows that it is possible to account for the predictable antepenultimate accent of Tokyo Japanese by making the final mora extrametrical, and building a trochaic foot.

Tanaka (1992) takes another approach in his analysis that does away with underlying accents. According to Tanaka (1992), accent is assigned on the basis of parameter values for which lexical entries are specified. He derives default antepenultimate accent by making the final foot extrametrical, and building an iambic foot. Thus, the unmarked foot form is a bimoraic right-headed foot, but nouns with left-headed feet exist as well. An example of a derivation that can be found in Tanaka (1992: 196) is given in (17).

(17)

		(. *)	
	(. *) (. *)	(. *)<(. *)>	
murasaki→	murasaki→	muraˀsaki	‘purple’

According to Tanaka (1992), an important argument in favor of the idea that right-headed feet are unmarked comes from the direction of accent shifts due to high vowel devoicing. We will come back to this in section 3.2.5.

With the introduction of OT, more possible analyses have become available. Because of the violable nature of OT constraints, we are not forced to choose for a single type of foot shape (trochaic vs. iambic, bimoraic vs. monomoraic). Furthermore, instead of positing a particular unit of extrametricality (mora, syllable, or foot), the constraint NONFINALITY “No prosodic head of PrWd is final in PrWd” (Prince and Smolensky 1993: 57) can be adopted, making more flexible analyses possible.

One of the first OT-analyses of Japanese accent is that by Suzuki (1995). According to Suzuki (1995), the Japanese foot is a quantitative trochee, consisting of two light syllables

[61] In another paper, Yamada (1992) adopts the SBG approach and proposes that unaccented words are subject to lexical edge marking, with a left boundary projected to the right of the rightmost mora (LRR). This means the representation of *sakura* would be *sakura(* , i.e. with an open constituent to the right of the word. Yamada (1992) invokes what he calls the “Edge Demarcation Convention” to block the further application of any metrical rules. According to this convention, all material outside a bracket is invisible to metrical rules. This means that the whole word is invisible in the case of unaccented words. Even if we allow such representations, the approach has a more serious problem in that a bracket is used as a diacritic that blocks the application of rules, which is not desirable from the viewpoint that brackets form a shorthand notation for arboreal structure (van der Hulst 2009).

or one heavy syllable. In Suzuki's (1995) study, the default accent in trimoraic words with different types of syllable structure is proposed to be the result of building strictly bimoraic trochees that are non-final in the word, i.e. leaving the final syllable unparsed. The metrical parsings in (18) are adapted from Suzuki (1995), with examples added for the sake of illustration.

(18)a.	$(\sigma_\mu\text{ˈ}\sigma_\mu)\sigma_\mu$	(kuˈra)su	'class'
b.	$(\sigma_\mu\text{ˈ}_\mu)\sigma_\mu$	(kaˈa)do	'card'
c.	$(\sigma_\mu\text{ˈ}_{\mu.})\sigma_{\mu\mu}$	(maˈga)zin	'magazine'
d.	$(\sigma_\mu\text{ˈ}\sigma_\mu)_\mu$	(puˈra)n	'plan'

As we can see in (18d), Suzuki (1995) allows foot boundaries to cross syllable boundaries, violating Syllable Integrity. Feet are assumed to be bimoraic and trochaic, and non-final in the word.

Katayama (1998) develops yet another analysis of loanword accent in Tokyo Japanese. In her analysis, the focus lies on variability between the antepenultimate and pre-antepenultimate pattern that can be observed in loanwords. As in Suzuki's (1995) analysis, an important constraint is that the accentual H tone may not fall on the word-final syllable or foot (NONFINALITY). The variation observed in loanwords is claimed to be the result of different constraint rankings, resulting in output forms like those in (19).

(19) a.	$\sigma_\mu\text{ˈ}(\sigma_\mu\sigma_\mu)$	kuˈ(rasu)	'class'
b.i)	$(\sigma_\mu\sigma_\mu\text{ˈ})(\sigma_{\mu\mu})$	(bitaˈ)(min)	'vitamin'
ii)	$(\sigma_\mu\text{ˈ}\sigma_\mu)(\sigma_{\mu\mu})$	(maˈga)(zin)	'magazine'
b.i)	$(\sigma_{\mu\mu})\sigma_\mu\text{ˈ}(\sigma_{\mu\mu})$	(baa)beˈ(kyuu)	'barbecue'
ii)	$(\sigma_\mu\text{ˈ}_\mu)\sigma_\mu(\sigma_{\mu\mu})$	(poˈo)to(ree)	'portrait'
c.i)	$\sigma_\mu(\sigma_\mu\text{ˈ}_\mu)\sigma_\mu$	po(keˈt)to	'pocket'

Interestingly, in Katayama's analysis, an accent may be associated to syllables that are not inside a foot, as in (19a). In cases like this, the constraint ACCENT-TO-HEAD, which forces accents to fall on foot heads, is violated. In Katayama's (1998) analysis, unfooted accents are thus preferred to monomoraic accented feet. However, the problem is that there is no evidence in favor of or against monomoraic feet in Tokyo Japanese.[62] Apart

[62] For the moment, we ignore the assumption that all Sino-Japanese morphemes constitute a foot. It is worth mentioning that if we allow monomoraic feet, antepentultimate accent in trimoraic words could be analyzed in terms of a word-initial monomoraic accented foot, i.e. *(kuˈ)(rasu).* Indeed, this approach can also be found in the literature (Asano 2002).

from this problem, the variation observed does not seem to be predictable at all. This becomes clear when we take a look at the data in (20).

(20)a.	bitaˀmin, *biˀtamin	‘vitamin’
b.	*dokuˀtaa, doˀkutaa	‘doctor’
c.	kaˀadigan, kaadiˀgan	‘cardigan’
	bangaˀroo, baˀngaroo	‘bungalow’

Katayama (1998) proposes that the variation in the assignment of loanword accent is the result of reranking of constraints. The reranking of these constraints makes the right predictions for words consisting of four light syllables. In such words, antepenultimate accent is dominant. According to Katayama (1998), the reason for this is that reranking of the same constraints results in two different optimal outputs which both have antepenultimate accent.[63]

(21)a.	$(\mu_\sigma\mu_\sigma)(\mu_\sigma\mu_\sigma)$
b.	$\mu_\sigma(\mu_\sigma\mu_\sigma)\mu_\sigma$

Shinohara (2000) also aims to capture default antepenultimate accent in terms of foot structure in her analysis of adapations of French words into Japanese. She claims that adaptations from French differ from adaptations from English in that in the former case there is no accent specification. As in Katayama (1998), she assumes the accent is placed as far as possible to the right of the word, while avoiding the final foot or syllable (NONFINALITY-σ/Ft). [64] Shinohara (2000) claims that the default accent should be analyzed as in (22a). However, when it comes to words in which the antepentultimate mora is an epenthesized vowel, which is avoided by the accent and indicated by “< >” in (22b), it is not clear whether we should allow monomoraic feet in order to satisfy the constraint that feet be trochaic, or whether we should rather allow iambic feet (22b).

[63] The two different rankings are as follows: (i) NONFINALITY >> ACCENT-TO-HEAD >> RIGHTMOST >> ALIGN-R AND (ii) NONFINALITY >> ALIGN-R >> RIGHTMOST >> ACCENT-TO-HEAD. RIGHTMOST is a constraint according to which the accentual H tone must be as far to the right of the word as possible. ALIGN-R is defined as a constraint that needs the right edge of every p-word to be aligned with the right edge of some foot.

[64] Note that Shinohara (2000) also needs to assume constraints may be reranked in order to account for variation.

(22)a.	ma(siˀku)ri	‘mâchicoulis (French) = machicolation’
b.	abu(riˀ)ko	‘abricot (French) = apricot’
	a(b<u>riˀ)ko	

There simply is no evidence on the basis of which we can choose between the two options in (22b).

Thus, so far we have looked at three different constraint-based analyses of loanwords that make quite different assumptions about foot structure. The problem is that there is no independent evidence for the assumptions made by the different authors.

Next, consider the analysis of foreign toponyms by Labrune (2012a). In her analysis, in which there is no room for syllables, the domain of NON-FINALITY is the accented (head) foot (NONFINALITY-Ft). Furthermore, a constraint that demands the parsing of moras into feet (PARSE-μ) is ranked above the foot-binarity (FTBIN) constraint. The result of this is that a bimoraic word like *paˀri* ‘Paris’ is analyzed as consisting of two monomoraic feet, i.e. *(paˀ)(ri)*. This analysis is counterintuitive, because if feet have a role to play in Japanese, we would expect a bimoraic word to consist of one bimoraic foot. Labrune (2012a) also argues that, unlike other analyses, her analysis can account for pre-antepenultimate accent in words like *pirenee* ‘the Pyrenees’ and *senegaru* ‘Senegal’, by means of the foot structures *(piˀre)(nee)* and *(seˀne)(garu)*. However, a foot-based analysis of words like this was already proposed by Tanaka (2008). The two analyses are similar in that candidate output forms like **(pireˀ)(nee)* are ruled out by a constraint that requires feet to be trochaic (FOOT/RHYTHMTYPE=TROCHEE). A difference between the two analyses is that in Tanaka’s (2008) the syllable plays an important role, whereas in Labrune’s (2012a) analysis no reference to the syllable is made.

Tanaka’s (2008) approach is restrictive in that only one foot per word is allowed, which is the accented foot. Moreover, he accounts for words that deviate from the antepenultimate pattern in an interesting way. For this reason, we will review his analysis in some detail.

One of the interesting features of Tanaka’s (2008) analysis is that he attributes the occurrence of pre-antepenultimate accent in quadrimoraic words to ALIGN-LEFT(F, ω), a constraint that aligns the left-edge of a foot to the left edge of the p-word. According to Tanaka (2008), among 89 accented loanwords with four light syllables, 50 words (56%) have pre-antepenultimate (-4) accent (e.g. *oˀfisharu* ‘official’), and 38 words (43%) have antepenultimate (-3) accent (e.g. *hareˀruya* ‘Hallelujah’). Tanaka shows that this can be explained if we assume that ALIGN-LEFT(F, ω) is as important as ALIGN-RIGHT(F, ω). The

tableau in (22) shows this;[65] there are two winning candidates for words with four light syllables, which means we would expect both types to be attested in equal numbers. The data collected by Tanaka (2008) indicates that this is indeed the case. For the sake of reference, the constraints that appear in the tableaux in (24), (25) and (26) are summarized in (23).

(23) a. ALIGN-RIGHT(F, ω):
The right edge of a foot is aligned with the right edge of a p-word.[66]
b. ALIGN-LEFT(F, ω)
The left edge of a foot is aligned with the left edge of a p-word.
c. NONFINALITY(σ, F):
The accent does not fall on the final syllable or foot.
d. FT-BIN (Foot-Binarity):
Feet are binary at the level of the mora or syllable.
e. Ranking: NONFINALITY >> FT-BIN >> ALIGN-LEFT, ALIGN-RIGHT >> TROCHEE.

(24) Pre-antepenultimate and antepenultimate accent in σμσμσμσμ loanwords

σμσμσμσμ	NONFIN	ALIGN-L (F, ω)	ALIGN-R (F, ω)	TROCHEE
→a. (σμ˥σμ)σμσμ			**	
b. (σμσμ˥)σμσμ			**	*!
→c. σμ(σμ˥σμ)σμ		*	*	
d. σμσμ(σμ˥σμ)	*!(F)	**		

On the other hand, among 117 accented quadrimoraic words ending in a final heavy syllable, 99 words (85%) are accented on the pre-antepenultimate mora, with only 18 words (15%) showing antepenultimate accent. This tendency is also predicated by the same ranking, although it is not clear from this ranking why antepenultimate accent would occur in the first place.[67] Note that Syllable Integrity is inviolable in Tanaka's (2008) analysis.

[65] The tableaux in (23), (24), and (25) are adapted from Tanaka (2008: 124). Candidates and constraints irrelevant to the discussion are omitted.
[66] Because Tanaka (2008) assumes syllables and not moras are parsed into feet (PARSE-σ), violations of the alignment constraint are counted in terms of the number of syllables intervening between the edge of the p-word and the foot.
[67] The answer to this most probably lies in the fact that the antepenultimate accent used to be the default pattern for these words. In other words, the variation is based on a diachronic shift.

(25) Pre-antepenultimate and antepenultimate accent in $\sigma_\mu\sigma_\mu\sigma_{\mu\mu}$ loanwords

$\sigma_\mu\sigma_\mu\sigma_{\mu\mu}$	NONFIN	ALIGN-L (F, ω)	ALIGN-R (F, ω)	TROCHEE
→a. $(\sigma_\mu\text{ˀ}\sigma_\mu)\sigma_{\mu\mu}$			*	
b. $(\sigma_\mu\sigma_\mu\text{ˀ})\sigma_{\mu\mu}$			*	*!
c. $\sigma_\mu(\sigma_\mu\text{ˀ})\sigma_{\mu\mu}$		*	*!	
d. $\sigma_\mu\sigma_\mu(\sigma_\mu\text{ˀ}_\mu)$	*!*(F,σ)	**		

In accented quadrimoraic words with a light-heavy-light syllable structure, the default pattern is antepenultimate accent. Again this can be accounted for with the same ranking. Notice that in words with this particular syllable structure, the foot-binarity constraint comes into play.

(26) Pre-antepenultimate and antepenultimate accent in $\sigma_\mu\sigma_{\mu\mu}\sigma_\mu$ loanwords

$\sigma_\mu\sigma_{\mu\mu}\sigma_\mu$	NONFIN	FT-BIN	ALIGN-L (F, ω)	ALIGN-R (F, ω)
→a. $\sigma_\mu(\sigma_\mu\text{ˀ}_\mu)\sigma_\mu$			*	*
b. $(\sigma_\mu\text{ˀ})\sigma_{\mu\mu}\sigma_\mu$		*!		**
c. $(\sigma_\mu\sigma_\mu\text{ˀ}\sigma_\mu)\sigma_\mu$		*!		*
d. $\sigma_\mu\sigma_\mu(\sigma_\mu\text{ˀ}_\mu)$	*!*(F,σ)	**	**	

While the pattern in (26a) is the default for words with light-heavy-light syllable structure, this does not hold when the heavy syllable contains the first half of a geminate. In such words, the pre-antepenultimate pattern is observed. In order to account for this, Tanaka (2008) proposes trimoraic feet (27b), as opposed to the bimoraic feet in words with other heavy syllables (27a).

(27) a. mi(saˀi)ru ‘missile’
pa(paˀi)a ‘papaya’
bi(niˀi)ru ‘vinyl’
o(reˀn)zi ‘orange’
b. (soˀneQ)to ‘sonnet’
(buˀtiQ)ku ‘boutique’
(suˀriQ)pu ‘slip’
(maˀjiQ)ku ‘magic’

According to Tanaka (2008), the different behavior of *Q* is caused by its low sonority. That is, syllables ending in *Q* may be treated as light syllables. In the words in (27a) the foot matches a heavy syllable, whereas in the words in (27b) a bimoraic syllable ending in the moraic obstruent is treated as light, the result of which is a trimoraic foot.

Tanaka's (2008) appeal to the low sonority of *Q* makes sense, but it is necessary to formalize how bimoraic syllables can be treated as light. This can be done in terms of prominence alignment constraints (Tanaka 2001; de Lacy 2002, 2007) that penalize certain types of syllables as foot non-heads. In this particular case, a constraint that bans heavy syllables ending in a non-sonorant coda in foot non-heads (28a) can be ranked below ALIGN-L(F, ω), while constraints that ban heavy syllables with a sonorant coda (28b) or a diphthong or long vowel (28c) are ranked above this alignment constraint. Thus, a relatively high ranking of the constraints in (28b) and (28c) as opposed to the constraint in (28a) can explain the difference between heavy syllable with long vowels, diphthongs, and sonorant codas on the one hand, and those with a non-sonorant coda on the other. For our purposes, the two constraints in (28b) and (28c) can be collapsed into the constraint in (28d), which refers to heavy syllables with a sonorant feature linked to the second mora ($X_{[+son]}$).

(28) a. *NONHEAD$_{Ft}$/$\sigma_{CVC[-son]}$: A syllable of the form $CVC_{[-son]}$ is not a foot non-head.
b. *NONHEAD$_{Ft}$/$\sigma_{CVC[+son]}$: A syllable of the form $CVC_{[+son]}$ is not a foot non-head.
c. *NONHEAD$_{Ft}$/σ_{CVV}: A syllable of the form CVV is not a foot non-head.
d. *NONHEAD$_{Ft}$/$\sigma_{CVX[+son]}$: A syllable of the form $CVX_{[+son]}$ is not a foot non-head.

The tableau in (29) shows how the candidate with a trimoraic initial foot is selected when the second syllable contains the first half of a geminate. Only candidates that do not violate NONFINALITY and FT-BIN are considered.[68]

(29) Pre-antepenultimate and antepenultimate accent in $\sigma_{\mu}.CVQ.\sigma_{\mu}$ loanwords

$\sigma_{\mu}\sigma_{CVQ}\sigma_{\mu}$	ALIGN-L (F, ω)	ALIGN-R (F, ω)	*NONHD$_{Ft}$/ $\sigma_{CVC[-son]}$
→a. $(\sigma_{\mu}$ˀ.CVQ)σ_{μ}		*	*
b. σ_{μ}.(CVˀQ).σ_{μ}	*	*!	

The tableau in (30) shows how the higher ranked constraint *NONHEAD$_{Ft}$/$\sigma_{CVX[+son]}$ rules

[68] Note that we need to assume that FT-BIN is defined in such a way that a foot may consist of at least two moras (i.e. "foot minimality") and at most two syllables ("foot maximality").

out trimoraic trochees ending in a heavy syllable with a sonorant second mora. Again only candidates that do not violate NONFINALITY and FT-BIN are considered.

(30) Pre-antepenultimate and antepenultimate accent in σ_μ.CVX$_{[+son]}$.σ_μ loanwords

$\sigma_\mu\sigma_{CVV}\sigma_\mu$	*NONHD$_{Ft}$/ $\sigma_{CVX[+son]}$	ALIGN-L (F, ω)	ALIGN-R (F, ω)
a. (σ_μˈ.CVV)σ_μ	*!		*
→b. σ_μ.(CVˈV).σ_μ		*	*!

Summarizing, in Tanaka's (2008) foot-based approach, several instances of variation in loanword accent can be accounted for in a convincing way. Any proposal for a non-metrical analysis of Tokyo Japanese accent will therefore also need to show that the different types of variation can be accounted for in an equally convincing way.

Tanaka (2008) is not the only researcher who has proposed trimoraic feet. Tanaka (2012, 2013) proposes the quantitative trochee, a foot consisting of either two moras or a heavy and a light syllable, to account for variation in accentuation of words such as those in (31). The examples include the foot structure proposed by Tanaka (2013).

(31) a.	ma(ziˈsya)<n>	vs.	(maˈzi)<syan>	'magician'
b.	sere(moˈni)<i>	vs.	se(reˈmo)<nii>	'ceremony'
c.	are(ruˈgi)<i>	vs.	a(reˈru)<gii>	'allergy'
d.	kaa(diˈga)<n>	vs.	(kaˈadi)<gan>	'cardigan'

Tanaka's (2013) analysis can be summarized as follows. He points out that although a final extrametrical mora with a right-aligned foot is optimal in the default case, such a parsing results in violations of Syllable Integrity in the above examples. In order to avoid this, variant forms in which the whole final syllable is made extrametrical are also produced.[69]

In Tanaka (2008), the variants on the right were analyzed as the default pattern in at least the innovating variety of Tokyo Japanese. However, in his analysis, it is not discussed which constraints can be reranked to generate the variants on the left without making the wrong predictions for other forms. Again, in order to explain a specific phenomenon or type of variation in the accent system, a specific type of foot structure is

[69] In Tanaka (2013), these examples serve as counterevidence to Labrune's (2012a) claim that the syllable does not exist in Tokyo Japanese.

posited without taking into account the system as a whole.[70]

The above discussion shows that a foot-based analysis of default antepenultimate accent is not as straightforward as it has been claimed to be. There is a lack of evidence in favor of a particular approach to foot structure as relevant to Tokyo Japanese accentuation. There is no consensus on the status of degenerate (monomoraic) feet, headedness (left-headed, right-headed, or both), exhaustiveness of parsing, syllable integrity, extrametricality, and the relation of the accent (or H tone) with foot structure. What is more, the question of how to deal with unaccented words (see section 4.2.4) remains untouched in most studies.

An important question is whether the lack of consensus in the literature, and the difficulties in deciding on the right analysis, is caused by the theoretical apparatus, or whether OT rather nicely reflects the ambiguous nature of foot structure in Japanese. In other words, if the linguists cannot decide on the right analysis, maybe this also holds for native speakers, resulting in multiple grammars with identical output forms in most cases, and different output forms in some cases. However, another possibility is that the lack of consensus concerning foot structure in Japanese could simply be related to foot structure playing a relatively unimportant or even no role in the accentual grammar. Before presenting an attempt at a non-metrical, purely tonal account of Tokyo Japanese accent, in the next section it is investigated whether compound accentuation can be used as a more informative source of evidence for metrical representations.

3.2.3. Compound accentuation

Apart from the evidence from prosodic morphology for bimoraic feet discussed in Chapter 3, Poser (1990) also proposes that the foot functions as a counting unit and as a unit of extrametricality or invisibility in nominal compound accentuation. A distinction between "short" and "long" N2s is proposed to be related to foot structure. Thus, short N2s consist of maximally one foot, and long ones of more than one foot. Furthermore, in long compounds, the final foot of N2 is extrametrical, and as a result the original accent cannot be retained.

[70] To capture the difference between the conservative and innovative forms in (31), Kager and Martínez-Paricio (2014) propose yet another type of feet: "internally layered or "minimally recursive" feet (see Martínez-Paricio 2013). According to Kager and Martínez-Paricio (2014), quadrimoraic words with a final heavy syllable such a *mazisyan* would be parsed as *ma.((zi˩.sya)n)* in the conservative variety, violating Syllable Integrity. In the innovative variety, on the other hand, the parsing is *((ma˩.zi.)syan)*, with no violation of Syllable Integrity. Whether Syllable Integrity is allowed or not depends on the ranking of the constraint DEP=μ "Dependent of FtNonMin [i.e. a non-minimal foot involving recursion] must be monomoraic". The analysis is interesting, and it seems worthwhile to investigate whether the analysis can be extended to words of other lengths, both accented and unaccented.

(32)a. Short N2s: consist of maximally one foot

garasu-dama^{0}	'glass bead'	(unaccented)
gaimuˀ-syoo	'Foreign Minstry'	(accent of final σ of N1)
momen-iˀto	'cotton thread'	(accent on first σ of N2)

b. Long N2s: > 1 Ft; rightmost foot extrametrical

yume-monogaˀ<tari>	'dream story'
me-guˀ<suri>	'eyewash'
te-kaˀ<gami>	'hand mirror'
sato-goˀ<koro>	'homesickness'

Poser mentions the three different classes of compounds with short N2s in (32a) (see also (9a) and (9b) above), but he does not give general rules that derive these three patterns. However, he does propose rules for the accentuation of compounds with long members.

(33) Accentuation of compounds with long N2 (adapted from Poser 1990: 99)

a. Mark the final foot of N2 as invisible.
b. If the visible portion of N2 is unaccented, assign an accent to its initial syllable.
c. Otherwise, leave the existing accent in place.

The examples in (32b) are correctly derived by the rules in (33). Note that tones are associated autosegmentally, without the need for a H tone to function as the head of a foot.

According to Poser (1990), foot boundaries need not coincide with syllable boundaries. The examples in (34), which are adapted from Poser (1990: 101), indicate this.

(34)

singata^{0} + zidoˀosya	→	singata-zidoˀ<osya>	'new-style car'
sukiˀn + kuriˀimu	→	sukin-kuriˀ<imu>	'skin cream'
haˀngaa + sutaraˀiki	→	hangaa-sutoraˀ<iki>	'hunger strike'
bi$^{(ˀ)}$zin^{0} + konkuˀuru	→	bizin-konkuˀ<uru>	'beauty contest'

Athough the analysis works for the forms in (30), it was pointed out by Kubozono (1995) that Poser's (1990) analysis cannot account for words in which the accent on the penultimate mora in words with long N2s is preserved, as in (35).

(35)a.	mata^0 + itoˀko	→	mata-iˀtoko, mata-itoˀko	‘a second cousin’
	kamiˀ + omuˀtu	→	kami-oˀmutu, kami-omuˀtu	‘paper diaper’
b.	saki^0 + ototoˀi	→	saki-ototoˀi	‘three days ago’
	kita^0 + deraweˀa	→	kita-deraweˀa	‘northern Delaware’

Invisibility of the final foot does play a role according to Kubozono (1995, 1997), but one that is more subtle. Kubozono’s (1995, 1997, 2011) assumptions on foot structure are that feet may not cross morpheme boundaries (Morpheme Integrity) or syllable boundaries (Syllable Integrity), and the maximum size of feet is two moras (FT-BIN), with monomoraic feet “banned in principle” (Kubozono 2011: 2891). The compound accent generalizations are summarized by Kubozono (2011: 2891) as follows.

(36)a. Preserve the accent of N_2 [the second member] in compounds unless it is on the very final syllable.

b. If the N_2 is unaccented or accented on the final syllable, place a default compound accent on the rightmost foot that is not in compound-final position.

The rules in (36) are illustrated by means of the examples in (37), which can be found in Kubozono (1995, 1997, 2011). According to NONFINALITY, the final foot may bear an accent, as long as it is not aligned with the word edge. For clarity, final accented syllables in N2, which are subject to the condition in (36a), are underlined. The foot structure in N1s is abbreviated, as in Kubozono’s work.

(37)	N1	N2	Compound	
a.	peˀrusya	neˀko	perusya)-(neˀko)	‘Persian cat’
	saki^0	ototoˀi	saki)(oto)(toˀi)	‘three days ago’
	yaˀmato	nadeˀsiko	yamato)-(nadeˀ)(siko)	‘Japanese lady’
	minami^0	amerika^0	minami)-(aˀme)(rika)	‘South America’
b.	abare^0	u<u>maˀ</u>	abareˀ)-(uma)	‘restive horse’
	neˀbada	<u>syuˀu</u>	nebadaˀ)-(syuu)	‘the State of Nevada’
c.	kamiˀ	omuˀtu	kami)-(o)(muˀtu) kami)-(oˀ)(mutu)	‘paper diaper’
	mata^0	itoˀko	mata)(i)(toˀko) mata)(iˀ)(toko)	‘a second cousin’

The forms in (37) are correctly derived if we apply the rules in (36). However, as we have

seen in (35a), for the forms in (37c), variants exist in which the accent falls on the antepenultimate moras. Furthermore, as shown in (38), there is a group of short N2s that lose their accent in compounds and behave as pre-accenting instead.

(38)	N1	N2		
a.	niˀngyo	hiˀme	ningyoˀ)-(hime)	'Little Mermaid'
b.	yoyaku^{0}	seˀki	yoyakuˀ)-(seki)	'reserved seat'

In order to derive these forms correctly, we need to reformulate the generalization in (36) as follows.

(39)a. Preserve the accent of N_2 the second member in compounds unless it is on the very *final foot*.

b. If the N_2 is unaccented or accented on the *final foot*, place a default compound accent on the rightmost foot that is not in compound-final position.

In Kubozono (1995, 1997), the revised rules in (39) are not stated as such in prose, but they do form the basis of his constraint-based analysis. Using the constraints in (40),[71] he explains the observed variation in (37c) in terms of the re-ranking of PARSE(N2), a constraint that dictates that the accent of N2 should be faithfully realized in the output. This is exemplified by the tableaux in (41) and (42), which are based on the discussion in Kubozono (1995).

(40)a. NONFIN(Ft): Accent may not be part of the final foot of the p-word.

b. NONFIN(σ): Accent may not be on the final syllable of the p-word.

c. RIGHTMOSTNESS: Put accent at the right edge of the p-word.

d. PARSE(N2): Preserve the accent of the second member as the CA.

e. ALIGN-CA: Align the accent with the boundary between N1 and N2.

(41) Ranking A: PARSE(N2) >> NONFIN(Ft)

mata#itoˀko	NONFIN(σ)	PARSE(N2)	NONFIN(Ft)	ALIGN-CA	RIGHTMOST
→a. mata)(i)(toˀko)			*	*	*
b. mata)(iˀ)(toko)		*!			***

[71] The (abbreviations of) names of constituents in the constraints are reformulated for the sake of consistency. Compound accent is abbreviated as "CA".

(42) Ranking B: NONFIN(Ft) >> PARSE(N2)

mata#itoˀko	NONFIN(σ)	NONFIN(Ft)	PARSE(N2)	ALIGN-CA	RIGHTMOST
a. mata)(i)(toˀko)		*!		*	*
→b. mata)(iˀ)(toko)			*		***

Importantly, the different rankings also account for the difference between *perusya-neˀko*, in which the accent of N2 is faithfully realized (Ranking A), and *ningyoˀ-hime*, in which the accent is shifted to the left (Ranking B). Thus, the analysis proposed by Kubozono (1995, 1997) implies that there are three types of words: words that select either Ranking A or B (e.g. native words like *itoˀko* 'cousin'), words with Ranking A (e.g. loanwords and native words like *neˀko* 'cat'), and words with ranking B (e.g. Sino-Japanese words and native words like *hiˀme* 'princess'). It seems that native words are moving from Ranking A to Ranking B in the speech of speakers of younger generations (Kubozono 1997; see also Uwano 1997b).

Building on the work of Kubozono (1995, 1997), Tanaka (2001) shows in more detail how the reranking approach can be used to account for variation in compound accent across the three different lexical strata of "Yamato" (native) words, Sino-Japanese, and loanwords, as well as for the emergence of deaccented compounds.

Kubozono, Ito and Mester (1997) present additional evidence for foot structure from compounds with N2s of more than five moras. As shown above, compounds with Ns of more than five moras behave differently from compounds with N2s of four moras or less. However, the situation is actually more complicated than this, because there is also a difference between compounds with N2s consisting of two morphemes or less, and those consisting of three morphemes or more. In the examples in (43), which are adapted from Kubozono, Ito and Mester (1997), morpheme boundaries of bound morphemes are indicated by straight lines "|", and word boundaries before compound formation by "+".

(43)

a.	minami^{0} + amerika^{0}	→	minami-aˀmerika
			‘South America’
	tiˀ|hoo + dan|tai^{0}	→	tihoo-daˀntai
			‘local body’
b.	minami^{0} | kariforunia^{0}	→	minami-kariforunia^{0}
			‘southern California’
	daˀi| + |saku|si|ka^{0}	→	dai-sakusika^{0}
			‘great songwriter’
	tiˀ|hoo + sai|ban|syoˀ	→	tihoo-saibansyoˀ
			‘local courthouse’
c.	tiˀ|hoo +{koo|kyoo| + daˀn|tai}	→	tiˀhoo kookyoo-daˀntai
			‘local public body’

According to Kubozono et al. (1997), the two different conditions of “≥ 5 moras” and “≥ 3 morphemes” can be captured in terms of feet if we assume a morpheme consists of at least one foot. Depending on the number of feet, the parsing into higher constituents differs. In (44a), the output is a single p-word. In (44b), the output is a structure that Kubozono et al. (1997: 158) claim to lie somewhere between a p-word and a p-phrase. Finally, the compound in (44c) consists of two p-phrases. The evidence for this is that two different accents surface, one on the first word and one on the second word.

(44)	Foot structure of N2	Class	Higher Level Parsing
a.	minami-(aˀme)(rika)	N2 ≤ 2 Ft	→ (N1 N2)
	tihoo-(daˀn)(tai)	N2 ≤ 2 Ft	→ (N1 N2)
b.	minami-(kari)(foru)(nia)^{0}	2Ft < N2 ≤3Ft[72]	→ (N1 (N2))
	dai-(saku)(si)(ka)^{0}	2Ft < N2 ≤3Ft	→ (N1 (N2))
	tihoo-(sai)(ban)(syoˀ)	2Ft < N2 ≤3Ft	→ (N1 (N2))
c.	tiˀhoo-(koo)(kyoo)-(daˀn)(tai)	N2 > 3Ft	→ (N1)(N2)

What the words in (44b) and (44c) have in common, is that the accent of N2 is preserved in the output, without assigning a new accent on the boundary of N1 and N2. On the other hand, a compound accent is assigned in (44a).

In Ito and Mester (2007), the difference between the different types of compounds is worked out in more detail. They distinguish between four types of compounds in

[72] Words of three feet or words of two feet and a stray syllable.

Japanese: left-and right-branching "word compounds", and mono-phrasal and bi-phrasal compounds. In their analysis, the words in (44a/b/c) would be analyzed as word compounds (consisting of a maximal p-word) (45a), mono-phrasal compounds (hence the suppression of accent in N1) (45b), and bi-phrasal compounds (45c) respectively.

(45)a. [[minami]$_{\omega}$ [aˈmerika]$_{\omega}$]$_{\omega}$
b. [[minami]$_{\omega}$ [kariforunia^{0}]$_{\omega}$]$_{\Phi}$
c. [[tiˈhoo]$_{\omega}$]$_{\Phi}$ [[kookyoo]$_{\omega}$[daˈntai]$_{\omega}$]$_{\omega}$]$_{\Phi}$

In (45a), the junctural accent is the result of two minimal p-words forming a maximal projection of the p-word (as in Ito and Mester 2007). The two p-words in (45b), however, do not form a new p-word, which needs an explanation. The only possible explanation is that N2 already is a maximal p-word, the result of which is that the compound is phrased as a p-phrase. This is indeed what Ito and Mester (2007) argue for words like that in (45b). The reason they give is that heads of p-words may be maximally binary, and since *(kari)(foru)(nia)* is parsed into three feet, the p-word is projected to the maximal level. The compound in (45c) is bi-phrasal, but here Ito and Mester (2007) do not offer more than the descriptive generalization that in words of this type, N2 consists of more than three feet.

The proposed generalizations are interesting, but do not work for all words (Uwano 1999a). This is also admitted by Kubozono (2004: 130), who gives the following counterexample to the analysis proposed by Kubozono et al. (1997).

(46)a.	kisoˈ + [tai+go]^{0}	→	kiso-taigo^{0}	'basic Thai'
			*?kiso-taˈigo[73]	
b.	kisoˈ + eigo^{0}	→	kiso-eˈigo	'basic English'
			kiso-eigo^{0}	

As pointed out by Kubozono (2004), the difference between (45a) and (45b) is that the former word *taigo^{0}* consists of a free word *taˈi* 'Thailand' followed by a suffix *-go* 'language', whereas the latter consists of two bound morphemes *ei-* 'English' and *-go*. In a foot-based approach, there should be no difference between the two words, because both N2s would consist of two feet.

[73] The asterisk with a question mark is as in Kubozono (2004). The question mark seems to be necessary indeed, as at least one of my informants finds *kiso-taˈigo* acceptable. The reason for this might be that for this person *taigo* is lexicalized and processed as a single word.

Another problem raised by Kubozono (2004) is that there are words with N2s consisting of no more than three feet that show the same behavior as the words in (43c).[74] An example of this is [koohaku [uta gassen]] ('Red-and-White Song Battle'), which is pronounced with two accents even though N2 consists of no more than three feet: *koꜜohaku uta-gaꜜssen* and not **koohaku uta-gaꜜssen.*

Kubozono (2004) argues that an analysis based on morphological structure is more successful. If we ignore the word *kariforunia* for the moment, the difference in morphological structure between the words in (43a), (43b), and (43c) can be accounted for in terms of the morphological structure of N2: those in (43a) consist of a single word, those in (43b) of a word derived from another word, and those in (43c) are compounds of two words.

(47)	Morphological structure of N2	Example
a.	[Word]	[amerika]
b.	[[Word]Suffix]$_{\text{Word}}$	[[sakusi]ka]
c.	[[Word][Word]]$_{\text{Word}}$	[kookyoo][dantai]

Kubozono (2004) says that *kariforunia* can be analyzed as being composed of the quasi-morphemes *kaliforu* and *-nia*. That is, these are not really morphemes in the Japanese language, but they are treated as if they are morphemes by the phonology. Interestingly, there is independent evidence for this idea from the accentuation of words ending in *-Cia* in general: *kanbozia*^{0} 'Cambodia', *ruumania*^{0} 'Rumania', *makedonia*^{0} 'Macedonia', *karedonia*^{0} 'Caledonia', *tanzania*^{0} 'Tanzania' (Kubozono 2011). All of these words take the unaccented pattern, which suggests that we are dealing with a deaccenting quasi-morpheme, just like the morpheme *-ka* above.[75]

According to Kubozono (2004), the idea of "quasi-morpheme' may actually be extended to loanword accentuation in general (see also Sato 2001). For instance, it could be the case that words like those in (40) are processed as phonological compounds, consisting of multiple "words". As Kubozono (2004) shows, theoretically several types of parsing would be possible, such as those in (48).

[74] This word is also given as an exception in Ito and Mester (2007).

[75] Other deaccenting quasi-morphemes are *-Cin* and *-Cingu* (Kubozono 2011; see Giriko 2009 for a comprehensive survey). See Labrune (2012a: 219) for a list of some major deaccenting morphemes, and NHK (1998) for an exhaustive list of such morphemes.

(48) Loanwords reanalyzed as pseudo-compounds

indoneʾsia	←	indo + nesia
		indone + sia
oosutoʾ ria	←	oosuto + ria
		oosu + toria

While the proposal of pseudo-compounds is interesting, the argument seems somewhat circular as there are a lot of counterexamples. In fact, the attentive reader will already have noticed that the words in (48) are counterexamples to the usage of *-Cia* as a deaccenting quasi-morpheme, even though they are also examples of placenames. Still, considering the fact that unaccented words are very rare in words of five moras or longer, it could well be the case that it is exactly the unaccentedness of words like *kaliforunia* that makes speakers reanalyze these words as consisting of two consituents. It is important to realize, however, that Kubozono's analysis only makes sense in a model of the phonology-morphology interface in which phonology has direct access to morphological structure, and not in an approach in which the Indirect Reference Hypothesis is respected. Still, if we allow direct reference, Kubozono's (2004) analysis is an alternative to the foot-based approach proposed by Kubozono et al. (1997).

Thus, although the foot-based approach to phrasal compounds is able to account for the observation that words of five moras or more and words of three morphemes or more behave in the same way, there are some excepions, and an alternative account is available. Still, the number of exceptions seems very small, and the alternative in terms of pseudo-compound structure is not without problems either. Apart from the fact that the pseudo-morphemes do not behave the same in all words, the status of pseudo-morphemes is unclear from a formal point of view.

From the above discussion, we may conclude that the foot-based approach to Tokyo Japanese compound accentuation is attractive in that it accounts for both word-level and phrase-level compounds. For word-level compounds, the constraint-based approach can be said to be an enormous improvement compared to previous approaches, in that the same generalizations can be used for both compounds with short N2s and long N2s. Therefore, it is not surprising that the success of the OT analyses of compound accentuation seem to have contributed to the idea that the foot is indispensable to account for Tokyo Japanese accentuation. However, the success of the constraint-based approach may be at least partly due to its output-based nature and the possibility of re-ranking. The actual role that foot structure plays within compound accentuation can only be evaluated when other alternatives are shown to be less convincing.

3.2.4. Unaccentedness

In many of the metrical approaches to Japanese accent, the status of unaccented words is unclear. In most studies, unaccented words are simply ignored under the assumption that they lack an underlying accent or tone(s). This is also the position taken by Asano (2002), who states that whether or not a word is unaccented is unpredictable. As evidence for this, she shows that unaccented words with different kinds of prosodic structure can be found. Asano (2002: 291) lists the following examples of loanwords.

(49)a.	ree.ru^{0}	‘rail’	e.	bat.teri^{0}	‘battery’
b.	ka.ree^{0}	‘curry’	f.	bai.o.rin^{0}	‘violin’
c.	kya.ra.me.ru^{0}	‘caramel’	g.	me.ki.si.ko^{0}	‘Mexico’
d.	rai.on^{0}	‘lion’	h.	ka.ri.fo.ru.ni.a^{0}	‘California’
e.	pu.re.mi.a.mu^{0}	‘premium’			

Asano (2002) is surely right in that one cannot predict whether a given word will be accented or unaccented with 100% accuracy. However, as discussed in section 4.1, certain tendencies can be observed. Among trimoraic and quadrimoraic words of the native Yamato stratum, unaccented words are in the majority. What is more, among quadrimoraic loanwords, unaccented words are in the majority, despite the fact that loanwords tend to be accented.[76]

Ito and Mester (2012) take up the question of why quadrimoraic loanwords tend to be unaccented. Before we take a look at their analysis, some important statistical data provided in their paper is introduced. The table in (50), which is adapted from Ito and Mester (2012: 30), is a breakdown of the number of unaccented words in quadrimoraic words with different types of syllable structure. The data makes clear that unaccented words are frequent in quadrimoraic words ending in two light syllables. Ito and Mester (2012) also observe that most quadrimoraic accented words are accented on the same syllable as in the source language, which explains the still relatively high number of accented words.

[76] According to a calculation made by Asano (2002) on the basis of the database in Katayama (1998), 68% of words of three moras or more have default antepenultimate accent, 21% have an accent on some other mora, and 11% are unaccented (Asano 2002: 269).

(50)

	Syllable structure	Ratio of unaccented words	
		Kubozono 2006 (963 words)	Ito & Mester 2012 (1289 words)
a.	$\sigma_\mu\sigma_\mu\sigma_\mu\sigma_\mu$	54%	46%
b.	$\sigma_{\mu\mu}\sigma_\mu\sigma_\mu$	45%	69%
c.	$\sigma_\mu\sigma_\mu\sigma_{\mu\mu}$	19%	19%
d.	$\sigma_{\mu\mu}\sigma_{\mu\mu}$	7%	10%
e.	$\sigma_\mu\sigma_{\mu\mu}\sigma_\mu$	24%	22%

The data shows that unaccented words are only frequent in words ending in two light syllables. Why this should be so is one of the questions that Ito and Mester (2012) aim to answer. The central idea of their analysis is that unaccentedness emerges when the constraints RIGHTMOST and NONFINALITY cannot be satisfied at the same time. NONFINALITY refers to both feet and syllables in Ito and Mester (2012); the final foot or the final syllable may not be accented. The definition of RIGHTMOST used by Ito and Mester (2012), which is given in (51), is different from that in Kubozono (1995). It requires the accented foot to be the rightmost foot in the word, and is violated when the accented foot is followed by another foot within the same p-word.

(51) RIGHTMOST: $*F\text{'}\ldots F\ldots]_\omega$

In a word consisting of exactly two feet, it will be impossible for this word to be accented while not violating either of these two constraints (52).

(52)a. $(\sigma_\mu\sigma_\mu)(\sigma_\mu\sigma_\mu)^0$ → RIGHTMOST and NONFINALITY satisfied

b. $(\sigma_\mu\sigma_\mu)(\sigma_\mu\text{'}\sigma_\mu)$ → RIGHTMOST satisfied, NONFINALITY violated

c. $(\sigma_\mu\text{'}\sigma_\mu)(\sigma_\mu\sigma_\mu)$ → RIGHTMOST violated, NONFINALITY satisfied

d. $\sigma_\mu(\sigma_\mu\text{'}\sigma_\mu)\sigma_\mu$ → RIGHTMOST and NONFINALITY satisfied[77]

e. $(\sigma_\mu\text{'}\sigma_\mu)\sigma_\mu\sigma_\mu$ → RIGHTMOST and NONFINALITY satisfied

There are three output forms that satisfy the two constraints. The form in (52a) is the only unaccented one. However, as can be seen in (52d) and (52e), a non-final, rightmost foot is surely possible in quadrimoraic words. The question then is why quadrimoraic words

[77] A non-final foot like this is posited by Ito and Mester (2012) for longer loanwords like *(baru)(se'ro)na* 'Barcelona' and *(asu)pa(ra'ga)su* 'asparagus'.

are parsed into two feet, as in (52a), excluding the forms in (52d) and (52e). According to Ito and Mester (2012), this is due to the following two constraints.

(53)a. INITIAL-FT: A prosodic word begins with a foot.
b. NOLAPSE(TONE): No tonal lapse. Violated by any pair of consecutive low-toned daughters of ω.

The constraint in (53a) speaks for itself, and is crucially violated by the form in (53d).[78] The constraint in (53b) needs more explanation. It is based on the idea that the L tone of the HL bi-tonal pitch accent is associated to all moras following the mora to which the accentual H tone is linked.[79] Whenever two consecutive constituents (foot and foot, syllable and syllable, foot and syllable, etc.) are L throughout, the constraint is violated. Ito and Mester (2012) give the following examples of words in which the constraint is violated (54a) as well as words in which it is not violated (54b).[80]

(54)a. minami-(aˈme)(rika) ‘South America’
H L → only the final foot is wholly L-toned

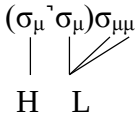
b. *minami-(kaˈri)(foru)(nia) ‘southern California’
H L → two consecutive feet are wholly L-toned

We can now see why the form in (52e) is rejected: it violates NOLAPSE(TONE) because it has two consecutive L-toned syllables which are directly dominated by the p-word.

(55) *(σμˈσμ)σμσμ → two consecutive syllables are wholly L-toned
H L

[78] The ranking of the four constraints discussed here is as follows: INITIAL-FT >> NOLAPSE(TONE), RIGHTMOST, NONFINALITY. See Ito and Mester (2012) for more details and other relevant constraints.
[79] Reference to toneless units makes the wrong predictions for unaccented words. It should be pointed out that according to Pierrehumbert and Beckman (1988) only the H part of the pitch accent is linked even in the surface representation. However, Gussenhoven (2004, to appear) shows evidence that the L tone is associated as well. As shown by Kubozono (1988/1993), the pitch fall from H to L continues throughout the phrase, which could also be interpreted as evidence for L-spreading.
[80] Note that the rejected form is an example of an overlong compound. Ito and Mester (2012) give many of such examples and state that the absence of junctural accent in overlong compounds is caused by NOLAPSE(TONE). However, this constraint does not explain why a word like *kiso-taigo*^{0} ‘basic Thai’ (*?*kiso-taˈigo*) does not get a junctural accent.

If the final two moras form a heavy syllable, on the other hand, the constraint is not violated. Although the L tone is linked to two consecutive moras, they are dominated by the syllable, which forms a single unit.

(56) 

→ only one non-footed syllable is wholly L-toned

Let us now move on to the question of why so many trimoraic words in the native lexicon are unaccented.[81] Ito and Mester (2012) argue that this is due to a constraint that militates against unparsed syllables in roots, PARSE-ROOT-σ.

(57) PARSE-ROOT-σ: Every syllable of a root is parsed [into a foot].

This results in parsings like those in (58a). Note that this constraint must be assumed to be highly ranked in the native and Sino-Japanese stratum only, because otherwise we would expect trisyllabic loanwords to be largely unaccented as well, which they are not (58b). However, loanwords made up of four light syllables often do surface without an accent (58c), because they are parsed into two bimoraic feet. In quadrimoraic words ending in a heavy syllable, on the other hand, the final syllable may be left unfooted, resulting in an accent on the only foot.

(58)a.	(saka)(na)^0	'fish'
	(nezu)(mi)^0	'mouse'
b.	(ba˺na)na	'banana'
c.	(ame)(rika)^0	'America'
d.	(ta˺ku)shii	'taxi'

An interesting consequence of the two different representations in (58a) and (58b) is that they can be used to explain the difference in accentuation between trimoraic simple truncated loanwords and trimoraic compound truncations. According to Ito and Mester (2012), the former are largely accented, whereas the latter are largely unaccented.

[81] Bimoraic words in Tokyo Japanese are mostly accented, violating NONFINALITY. In order to account for this fact, Ito and Mester (2012) propose a constraint MINWORDPROM, according to which minimal p-words must be accented.

(59) a. (seki)(me)^{0} < sekusyon meeto ‘section mate’
b. (teˈre)bi < terebizyon ‘television’

In order to ensure that both morphemes in (59a) project a foot, Ito and Mester (2012) propose the following constraint.

(60) LEXFT: Every lexical morpheme (i.e. full content morpheme, not grammatical formative) minimally projects its own foot.

Although the analysis in Ito and Mester (2012) makes some more interesting claims, the above discussion suffices for the purposes of this chapter. The claim is clear and interesting: unaccentedness is governed by constraints on prosodic structure. Furthermore, the analysis is explicit on questions such as exhaustiveness of parsing, foot heads (left-headed), Syllable Integrity (no violations allowed), and the status of degenerate feet (allowed).

However, Ito and Mester’s (2012) proposal also raises a number of questions. To start with, unaccentedness is not a phenomenon that is only observed in Tokyo Japanese, and it remains to be seen whether correlations between unaccentedness and words of certain lengths and/or certain types of syllable structure can all be explained in terms of the interaction between the above constraints.

There are also some potential empirical problems. More concretely, certain words cannot be generated by the ranking proposed by Ito and Mester (2012). For example, variation is observed among forms like *amaˈzon ~ aˈmazon* ‘Amazon’. Because right-headed feet are not available in Ito and Mester’s (2012) analysis, the potential parsing of *(amaˈ)(zon)* is ruled out. Also, the fact that around half of the quadrimoraic words consisting of all light syllables take the pre-antepenultimate pattern rather than the antepenultimate pattern (Tanaka 2008), as in *oˈfisyaru* ‘official’, cannot be accounted for in their analysis, because *(oˈfi)syaru* violates NOLAPSE(Tone), and *(oˈfi)(syaru)* violates RIGHTMOST.

Also, it is not clear how the accentual tendencies of words in the Sino-Japanese stratum can be accounted for. It has been shown that the accent patterns of Sino-Japanese words are predictable to a considerable extent, with syllable structure playing an important role in the determination of the output forms (Kubozono and Ogawa 2004; Ogawa 2004, 2008). This may not be what we expect though, because due to the constraint LEXFT, every Sino-Japanese content morpheme projects a foot, no matter it consists of a monomoraic syllable, a bimoraic syllable, or two monomoraic syllables.

Another point is that the analysis merely reflects tendencies in parts of the lexicon, something which is of course recognized by Ito and Mester (2012). However, if the preference for unaccentedness is based on stochastic knowledge of the lexicon, a question we may ask is whether this stochastic knowledge could not have a more direct role in the grammar. The relevance of this question may be illustrated by the example of truncation.

Ito and Mester (2012) note that simple trimoraic truncated loanwords are usually accented, whereas trimoraic compound truncations are usually unaccented. An explanation for this may be sought in the canonical patterns of these two types of word formation. The unmarked simple truncated loanword consists of two moras (and two syllables), whereas the unmarked compound truncation is made up of four moras (see Ito 1990). Now, we also know that bimoraic words are mostly accented, and quadrimoraic words mostly unaccented. Therefore, it could be that trimoraic truncations of compounds choose the default accent pattern (that of quadrimoraic truncations) because this is the most frequently occurring pattern for compound truncations in general. Still, this does of course not answer the question of why quadrimoraic loanwords would be unaccented.

There could be functional answers to the question where unaccentedness comes from. For instance, Labrune (2006: 180) proposes that frequency effects may be involved. In a corpus consisting of 52,896 native and Sino-Japanese words, Shibata (1994) found that unaccented words make up as much as 67.7% of all quadrimoraic words, compared to 52.6% for trimoraic words (see Labrune 2006: 180, 232; 2012a: 186-187. Indeed, it may be that 67.7% (2 out of 3) is a robust enough number to rely on, whereas 52.6% (roughly half) is not.

Labrune (2012a) points out that Tanaka (2008) offers a more convincing explanation. Based on data from a database of 56,812 words (Sugito 1995), Tanaka (2008) demonstrates that the higher the type frequency of words of a particular length, the higher the unaccentedness ratio among words of that length is. According to Tanaka (2008), the data in (61) suggests that the absence of the culminativity and boundary clues from accent in unaccented words are, as it were, taken over by it being made up of four moras, which is the canonical length for a Japanese word.

(61) Relation between word length type frequency and unaccentedness ratio

	#μ	Type frequency	Unaccentedness Ratio
a.	4μ	39.7%	74%
b.	3μ	22.1%	51.5%
c.	5μ	15.9%	27.9%
d.	6μ	10.8%	20.9%
e.	2μ	5.3%	16.9%
f.	others	6.1%	9.9%

I agree with Tanaka (2008) and Labrune (2012a) that this is an attractive explanation for the relatively large number of unaccented quadrimoraic words. Even so, the question of why the second most frequent word type, namely that of trimoraic words, does get accented in the default case remains unanswered. A possible reason for this is that in trimoraic words, antepenultimate accent is at the same time a word-initial accent, which makes the accented option relatively attractive compared to unaccented words.[82]

To sum up, Ito and Mester's (2012) proposal that unaccentedness is caused by certain types of metrical parsings structure is an intruiging idea that could have broad implications for the phonological grammar as a whole. Yet, it does not account for all data, and the analysis seems to depend on knowledge of stochastic distributions that could have a more direct role in the grammar. Thus, there remain a number of topics for future studies. First, there is the question to what extent Ito and Mester's (2012) analysis is compatible with claims about foot structure made in studies on default accent in loanwords and compounds, or whether it is able to account for the observed patterns (including the variation) in an alternative way.

Second, a more systematic comparison with more functionalist alternatives would be welcome. Third, evidence for their approach from other dialects would be welcome. Although – due to the lack of sufficient data – the question of unaccentedness in dialects other than Tokyo Japanese will not be taken up in this thesis, there will be extensive discussion of issues related to foot structure in the following chapters.[83]

3.2.5. Accent and high vowel devoicing

Another perennial problem in Japanese accentology is the relation between pitch accent and vowel devoicing (McCawley 1968; Sugito 1969, 2005; Haraguchi 1977; Vance 1987;

[82] As discussed in section 4.2.2, Tanaka (2008) argues that left-edge alignment plays a more important role in Tokyo Japanese accentuation than assumed in most other studies.

[83] In Chapters 4 and 5, two different instances of what resembles Ito and Mester's (2012) INITIAL-FT constraint will be shown to be useful in the analyses of the dialects of Maisaka and Ichihara.

Maekawa 1990; Kitahara 1998; Hasegawa 1999; Hirayama 2009). High vowels in (Tokyo) Japanese tend to get devoiced when preceded by a consonant that is [-voiced], and followed by another [-voiced] consonant or a word boundary (see Vance 1987 for further details). Although most studies of the interaction between accent and vowel devoicing take a descriptive or experimental viewpoint, there has been some discussion in the theoretical literature on how to deal with accent shifts caused by high vowel devoicing (Haraguchi 1991; Tanaka 1992, 2005; Yokotani 1997).

It was already pointed out by Han (1962b) and Sugito (1969/1982) that accents may be realized on devoiced vowels. The accent can still be detected by speakers because of a raised falling pitch in the syllable following the accented syllable (Sugito 1969/1982, Maekawa 1990). Kitahara (1998), based on his own measurements, argues that this fall in pitch is evidence for the reassociation of the accentual H tone to the syllable following the originally accented syllable, but only in cases in which there is no preceding phrasal H tone.

The general claim in the literature seems to be that when an accented vowel is devoiced, it "shifts" to the left, except if the devoiced vowel is in the first or second syllable, in which case it "shifts" to the right (Haraguchi 1977).[84] The reason for the absence of leftward shifts from initial syllables needs no explanation. For words with an accent on the second mora, it has been claimed that the absence of leftward shifts is caused by the extrametricality (Archangeli 1981; Halle 1982), or extratonality (Archangeli & Pulleyblank 1984) of the first mora.[85]

According to Tanaka (1992), an important argument in favor of right-headed feet as unmarked comes from the direction of accent shifts due to high vowel devoicing.[86] In general, accent is assumed to shift to the right in left-headed feet, and to the left in right-headed feet. Because in Tokyo Japanese accent shifts to the left where it can, Tanaka (1992) concludes that Tokyo Japanese feet must be right-headed. Examples given by Tanaka (1992) are shown in (62). The earlier stages of the derivation given by Tanaka (1992) are omitted.

[84] For the sake of convenience, the derivational term "shift" is used. In a non-derivational approach, the "accent shift" can be considered to be a difference in accent location between (i) the input and the output form, or (ii) two paradigmatically related forms.

[85] Archangeli (1981) and Halle (1982) are referred to by Prince (1983); Archangeli and Pulleyblank (1984) is referred to by Haraguchi (1988).

[86] Haraguchi (1991) assumes right-headed unbounded feet in order to explain the direction of accent shifts. See Tanaka (1992) for problems for such an approach.

(62)a. (wari)(bikiˀ)<(ken)> → (wari)(biˀ)kḭ<(ken)> ‘discount coupon’
(kari)(yakuˀ)<(soku)> → (kari)(yaˀ)kṵ<(soku)> ‘tentative appointment’
b. (in)(satu)(kí)<(kai)> → (in)(satu)(kḭˀ)<(kai)> ‘printing machine’
(boo)(siˀ)<(kake)> → (boo)(sḭˀ)(kake) ‘hat rack’

The above forms show that the accent may shift leftward inside a foot (62a), but not across foot boundaries (62b).[87]

Yokotani (1997) presents data on the basis of which he claims accent shifts across foot boundaries. He proposes to adopt the superfoot – a unit that dominates the foot within the Prosodic Hierarchy – as the domain of accent shift. The compounds in (63), which are adapted from Yokotani (1997), show variation due to high vowel devoicing.[88]

(63)

a.i.	zikan^0 + chikaku^0	→	zikan-chḭˀkaku, zikan-chḭkaˀku	‘time perception’
	ko^0 + hituzi^0	→	ko-hḭˀtuzi, ko-hḭtuˀzi	‘lamb’
ii.	kootei^0 + kakaku^0	→	kootei-kaˀkaku, *kootei-kakaˀku	‘official price’[89]
	boosui^0 + kakoo^0	→	boosui-kaˀkoo, *boosui-kakoˀo	‘waterproofing’
b.	asiˀ + kosi^0	→	asḭˀkosi, asḭ-koˀsi	‘legs and waist’
c.	yamaˀguti + keˀn	→	yamagutḭˀ-ken, yamaguˀtḭ-ken	‘Yamaguchi Prefecture’
	nagaˀsaki + keˀn	→	nagasakḭˀ-ken, nagasaˀkḭ -ken	‘Nagasaki Prefecture’

According to Yokotani (1997), the difference between the forms in (63a-i) and (63a-ii), all of which are compounds with N2s that are unaccented in isolation, must be due to the devoicing of the first vowel of N2 in (63a-i). In derivational terms, the accent can be said to shift due to the devoicing of the high vowels. Interestingly, the accent shifts to the right, and not to the left. Why does the accent not shift to the left in e.g. **koˀ-hḭtuzi* instead of *ko-hḭtuˀzi*? Yokotani (1997) points out that this cannot be due to a constraint against accent shifts across morpheme boundaries, as proposed by Yamada (1990). The reason

[87] Vowel devoicing is assumed to involve the loss of a syllable node by Tanaka (1992). Therefore, the devoiced vowel loses its position in the foot completely. However, we could also assume that, apart from the shift of the accent to the left, the foot stays as it is. This is actually the approach taken in Tanaka (2005) in his reanalysis of the phenomenon within OT.

[88] Adapted from Yokotani (1997). The accent forms of the isolation forms of the compound members are from NHK (1998).

[89] For the isolation form *kakaku* ‘price’, NHK (1998) actually lists two accent patterns: *kakaku^0* and *kaˀkaku*.

for this is that in forms like *asi̥-koˀsi* (63b), the accent does shift across a morpheme boundary. What is more, assuming that *asi* and *kosi* both form a foot, the accent shifts across a foot boundary. Yokotani argues that an accent shift is only bound by a superfoot boundary. The assignment of feet and the relevant superfeet in Yokotani's (1997) analysis is as in (64). Superfeet are indicated by curly brackets "{ }".

(64)

a.	zikan^0 + chikaku^0	→	(zi)(kan)-{(chi̥ˀ)(kaku)}
			(zi)(kan)-{(chi̥)(kaˀku)}
	ko^0 + hituzi^0	→	(ko)-{(hi̥ˀ)(tuzi)}
			(ko)-{(hi̥)(tuˀzi)}
b.	asiˀ + kosi^0	→	{(asi̥ˀ)-(kosi)}
			{(asi̥)-(koˀsi)}
c.	yamaˀguchi + keˀn	→	{(yama)(guti̥ˀ)}-(ken)
			{(yama)(guˀti̥)}-(ken)
			*{(yama)(guti̥)}-(keˀn)
	nagaˀsaki + keˀn	→	{(naga)(saki̥ˀ)}-(ken)
			{(naga)(saˀki̥)}-(ken)
			*{(naga)(saki̥)}-(keˀn)

Note that Yokotani (1997) assumes that a compound consists of a single p-word rather than a recursive one. Under this assumption, he is able to claim that the domain of accent shift cannot be the p-word. However, there are at least two alternative analyses. If we assume all members of a compound project p-words, as is assumed in this thesis, the maximal domain in which accent may "shift" can be said to be the p-word. In such an analysis, we only need to assume that the compound in (64b) is lexicalized and consists of a single p-word. A more straightforward approach, however, is to analyze the blocking of an accent shift in the words in (64c) by means of NONFINALITY(σ), the constraint that bans accent in the word-final syllable which was shown to be independently necessary in Section 3.2.3.

It thus appears that whether there is a role for foot structure in accentuation or not, the alleged accent shifts are not bound by foot boundaries. Neither do accent shifts provide us with evidence for the superfoot, athough this conclusion depends on the view that compounds with junctural accents are recursive p-words.

In the above, it was simply taken for granted that accent shifts are phonologically real. However, an important question that has not been satisfactorily answered yet is whether

there are genuine phonological accent shifts due to vowel devoicing in the first place. Or, avoiding the derivational flavor of the term "shift", is there actually evidence that the interaction between accent and high vowel devoicing is phonological rather than phonetic?

As discussed in Prince (1983) and Poser (1984), the so-called leftward shift could be simply the result of the H tone linked to the devoiced vowel not being realized. This could be either the result of the delinking of the H tone or it could be due to phonetic interpretation. In either case, the rightmost realized H-toned mora is the one preceding the accent. This situation is graphically illustrated in (65a), where dotted lines indicate the portion of the pitch curve that is reconstructed by the listener, and the mora shaded in gray the mora that is perceived as being accented. If the pitch on the mora following the originally accented mora is low rather than of a falling nature, the accent may be interpreted to have shifted to the left. On the other hand, in cases in which the drop in pitch is partly realized on the mora following the accented mora (65b), the mora with devoiced vowel will typically be perceived to be accented (Sugito 1969/1982).

(65)a. 

HL

"leftward accent shift"

b. 

H L

"no accent shift"

As for the rightward shift observed from the second to the third syllable, we may refer to the fact that peak delay is more likely when the H peak is directly preceded by a L tone (because of the production of a rise). Furthermore, due to the absence of high pitch before an accent associated to the second syllable, there will simply be no acoustic information on the basis of which the accent can be perceived as having switched to the left.

While these phonetic accounts may work for most cases of the accent shift, they do not explain why accent should shift to the right (or be perceived as having shifted to the right) in the N2 of compounds. One possibility is that genuine accent shifts do exist, and are always rightwards, whereas in the case of putative leftward accent shifts the H tone is only phonetically not realized on the accented syllable.[90]

[90] This possibility is also discussed by Poser (1984), who only refers to rightward shifts of accent on the first or second syllable.

3.2.6. An interim evaluation of the metrical approach

Now that we have reviewed several phenomena in Tokyo Japanese accentuation for which a metrical approach has been argued to be necessary, let us summarize what can be said about the properties and function of foot structure according to these studies.

Starting with the properties of foot structure, Kubozono (1999) mentions that there is no consensus on a number of issues concerning foot structure (66a~d). To these we may add the issues in (66e~h).

(66)a. Directionality of footing (left-to-right or right-to-left)
b. The status of degenerate feet (allowed or not)
c. Exhaustiveness of parsing (Y/N)
d. Syllable Integrity (respected or not)
e. Foot-type (left-headed vs. right-headed)
f. The relation between the accentual H and foot heads (H to foot-head: Y/N)
g. The role of extrametricality/non-finality (mora and/or syllable and/or foot)
h. The status of covert feet (see below)

In the studies reviewed in this section,[91] different positions were taken regarding these issues. In some cases, this is because it does not seem to matter whether all syllables are parsed into feet or not, or whether degenerate feet are allowed or not. This is probably related to the fact that most studies only focus on a limited amount of data. The study by Ito and Mester (2012) shows that as soon as one aims to account for both accented and unaccented words in terms of foot structure, it does become necessary to take a stance on the issues in (66).

Most of the issues listed in (66) were discussed above. However, the status of covert feet (66h), i.e. feet that receive no (perceptible) phonetic realization, has not yet been mentioned. One may wonder whether such feet are not overly abstract. How can such feet and their internal structure (i.e. head vs. non-head) be abstracted from the surface forms if there are no direct phonetic clues for them in the form of secondary stress, vowel reduction or consonantal lenition or fortition?[92] It has been proposed by some researchers that feet without (and maybe even with) a lexical accent do not have an intrinsic head in the first place. Such headless feet have been proposed by Hagberg (1993), Crowhurst and

[91] The issues mentioned by Kubozono (1999) were partly based on the same studies discussed in this chapter.

[92] Interestingly, in a recent study by Kawahara et al. (2014), alternating patterns of jaw displacement are said to possibly reflect metrical foot structure. Pitch accent, on the other hand, did not affect jaw displacement. See also Bennett (2012) for possible experimental evidence for covert feet in Japanese.

Hewitt (1995a), Krämer (2009), and Shimoji 2009).[93] Another possibility is that covert feet are assigned a head by default. In the case of Tokyo Japanese, this would be the left mora, at least according to most studies. In any case, in the absence of direct phonetic evidence for left- or right-headed feet, all that we can do is rely on the position of the accent within the proposed feet, which is assumed to be variable in some of the analyses.

One way in which we might be able to tackle the above issues is to look at microvariation across dialects. Evidence in favor of certain types of structure in one dialect may help us to decide on the analysis of another dialect. Therefore, in Chapters 4, 5, and 6, we will focus on what metrical and tonal approaches have to offer us in analyses of other dialects.

Moving on to the function of foot structure, the five different phenomena that have been claimed to constitute evidence for foot structure are repeated in (67).

(67) a. Surface tonal patterns
b. Default antepenultimate (-3) accent
c. Compound accentuation
d. Unaccentedness
e. Domain of accent shifts

One thing that has become clear is that there does not seem to be any evidence for a crucial role of foot structure (either unbounded or bounded) in determining the surface tonal patterns other than the location of the accentual H. Furthermore, there is no clear evidence for the bimoraic foot playing a role in the interaction of accent and vowel devoicing.

This leaves us with evidence from compound accentuation, default loanword accent, and unaccentedness. The foot-based analyses couched in OT by Kubozono (1995, 1997) and scholars building on his work have indeed proved to account successfully for the behavior of both compounds with long N2s and compounds with short N2s. Yet, it could be that this success is the consequence of the adoption of an output-based framework that allows for variation, rather than the assumption that accent is assigned metrically. Exactly the same holds for the several different metrical approaches to default loanword accent. The fact that so many authors propose different types of metrical parsings at least indicates that a foot-based approach is not as straightforward as it seems.

[93] The possibility of headless feet is enabled by the 'separability hypothesis' (Crowhurst & Hewitt 1995a; cf. Halle & Vergnaud 1987), according to which the construction of feet and foot-internal head assignment are independent processes.

The evidence for foot structure from phrasal compounds looks convincing at first sight; by referring to the foot, it becomes possible to restate "five moras or more" and "three or more morphemes" as "more than two feet". However, there is a small number of exceptions to this generalization, and an alternative in which morpho-syntactic structure plays an important role has been presented by Kubozono (2004).

The study by Ito and Mester (2012) lays out an interesting new view on the emergence of unaccentedness in Tokyo Japanese. Their analysis is insightful and treats both unaccented and accented words of different lexical strata, and both simplex and derived words. Still, the constraint ranking they propose does not make the right predictions for all forms. Also, the fact that they refer to a tonal constraint like NOLAPSE(TONE) raises the question whether tone cannot be assumed to play a more important role in the assignment of pitch accent in general.

3.3. A tonal approach

In this section, an attempt at a non-metrical, purely tonal approach to Tokyo Japanese accentuation is developed.[94] It will be shown that many of the generalizations that have been stated in terms of metrical structure can be reformulated in terms of tone. We will discuss the same five phenomena that were reviewed for the metrical approach. In some cases I will refer back to section 3.2 in order to avoid unnecessary repetitions.

3.3.1. Surface tonal patterns

Obviously, in a non-metrical tonal approach, acounting for the surface tonal patterns is directly related to the tonal representations.

Pierrehumbert and Beckman (1988) argue that the TJ pitch accent is represented on the tonal tier as HL, with only H being associated (see Chapter 2). However, as demonstrated by Gussenhoven (2004, to appear b), there is evidence for the association of the trailing L tone. Therefore, I assume that both the H and the L tones of the pitch accent are associated, as in (68). In the absence of concrete evidence, I take no stance on whether the L tone spreads all the way to the end of the word.

[94] There exists one well-worked-out tonal analysis of Tokyo (and Osaka) Japanese by Clark (1986). While her analysis is interesting, it heavily makes use of rule ordering, a mechanism which is unavailable in the approach taken in this thesis. Also, her analysis is based on the assumption that a default H tone is assigned word-finally in unaccented words, an assumption that has been shown to be untenable from an empirical point of view (Pierrehumbert and Beckman 1988).

(68) Autosegmental-metrical analysis of Tokyo Japanese

Unaccented	*Antepenultimate*	*Penultimate*	*Final*
sakura (-ga)	karasu(-ga)	tamago(-ga)	otoko(-ga)
			
L%H% L%	L%HL L%	L%H L L%	L%H%HL L%
‘cherry-nom’	‘crow-Nom’	‘heart-nom’	‘man-nom’

The phrasal tones are assumed to be inserted by the grammar. Accentual phrases start with a L%H%-sequence, and end with a L% in declarative utterances (Maekawa 2010).

For the lexical representation of the pitch accent, we have the options in (69).

(69) Different analyses of the underlying specification of accent in Tokyo Japanese

Example	*Accent*	*H-only*	*L-only*	*H*L*
karasu(ga)	μ˩μμ(μ)	<u>H</u>ØØ(Ø)	Ø<u>L</u>Ø(Ø)	H*LØ(Ø)
tamago(ga)	μμ˩μ(μ)	Ø<u>H</u>Ø(Ø)	ØØ<u>L</u>(Ø)	ØH*L(Ø)
otoko(ga)	μμμ˩(μ)	ØØ<u>H</u>(Ø)	ØØØ(<u>L</u>)	ØØH*(L)
sakura(ga)	μμμ(μ)	ØØØ(Ø)	ØØØ(Ø)	ØØØ(Ø)

The diacritic accent approach is not adopted, because the goal of this section is to develop a purely tonal analysis. In the absence of evidence against an underlying tonal representation, diacritics are needlessly abstract (see Poser 1984, Pulleyblank 1986).

The L-only approach (Backley and Nasukawa 2013) is not adopted either, as it is unable to capture the boundary marking function of the junctural accent in, for instance, compounds. In compounds, an accentual H tone is aligned with either the right edge of N1, or with the left edge of N2, a situation that cannot be described in terms of L without also referring to the preceding H.

The choice between the H-only approach and the HL-approach is partly related to the position we take on phonological derivations. In the H-only approach, we need a constraint that inserts a L tone following a H tone. This constraint can only be active in the lexical grammar, or else a L tone would be inserted following the H tone of the phrase-initial L%H%.[95] Thus, if we do not refer to accent or foot structure, the H-only approach can only be adopted if we assume the phonological grammar consists of a lexical and a postlexical module. However, if we assume a strictly two-level non-derivational grammar in which phonological computations are only made once, we need to assume HL as the

[95] Another possible analysis would be to assume the relevant constraint can distinguish between accentual H tones and non-accentual H tones.

only possible output representation of the pitch accent.

If the L tone is inserted in the lexical phonology anyway, it does not really make a difference if we simply assume it to be part of the lexical input representation. Therefore, I will assume the lexical accent consists of HL both in the output and the input. Still, I would like to stress that the tonal analysis developed in this chapter does not hinge on the input representation. In an alternative approach, the underlying accent is a diacritic or a H tone with a diacritic (H^* or H᾽), and HL or L is inserted by the grammar.

3.3.2. Default antepenultimate accent

I already mentioned that both the H and the L tones of the pitch accent are assumed to be associated, following Gussenhoven (2004, to appear b). This is a crucial assumption which enables us to arrive at a non-metrical tonal analysis that is strictly local. In a tonal analysis, lexical tones are aligned as far to the right as possible (ALIGN-R), but the rightmost mora (or syllable; see below) is "extratonal" (NONFINALITY). The H and L part of the bitonal pitch accent are kept together by the constraint CONCATENATE, which bans gaps in the form of toneless moras between two associated tones belonging to the same tonal complex (Gussenhoven 2004). The relevant constraints and their ranking are given in (70). The constraint CONCATENATE can be assumed to be undominated in the grammar, because it never happens that an empty tone-bearing unit (TBU) or another tone intervenes between the H and L part of the pitch accent. Therefore, it will be left out of any further discussion or ranking in this chapter. NONFINALITY must outrank ALIGN-R in order to get a penultimate accent.

(70) a. ALIGN-RIGHT(L, ω) = ALIGN-R(L):
The right edge of the L tone is aligned with the right edge of the p-word.
(Gradient: one violation for each TBU intervening between L-toned TBU and the word edge)

b. NONFINALITY-TONE(μ):
A tone is not associated to the final mora of the p-word.

c. CONCATENATE:
Tones in bitonal morphemes are aligned with each other.[96]

d. NONFINALITY-TONE >> ALIGN-R(L)

For an input that consists of a quadrimoraic word and a HL whose location needs to be determined (i.e. a floating HL), the output will be as in (71). The domain of NONFINALITY

[96] Gussenhoven (2004: 150). Gussenhoven (2004) attributes the idea to Riad (1998).

is represented by "< >".

(71) μμμ<μ>
  | |
  HL

The representation in (71) shows that the default antepenultimate pattern can be easily accounted for in terms of tones. However, what about the variant forms that show the pre-antepenultimate pattern? Let us start with words consisting of four light syllables. As discussed above, Tanaka (2008) shows such words surface either with antepenultimate or pre-antepenultimate accent. In Tanaka (2008), a left-edge alignment constraint is the driving force behind the pre-penultimate pattern. Let us start with formulating this constraint in a similar way to ALIGN-RIGHT(L, ω)

(72) ALIGN-LEFT(H, ω) = ALIGN-L(H):
The left edge of the H tone is aligned with the left edge of the p-word.
(Gradient: one violation for each TBU intervening between H-toned TBU and the leftmost TBU in the p-word)

The tableau in (73) shows that if the two alignment constraints have the same ranking, the right predictions are made for words of four light syllables. In words with the antepenultimate pattern (73a), both ALIGN-L(H) and ALIGN-R(L) are violate once, whereas in the pre-antepenultimate pattern, ALIGN-R(L) is violated twice.

(73)

cv.cv.cv.cv.	NONFIN	ALIGN-L (H)	ALIGN-R (L)
→a. cv.cv^{H}.cv^{L}.cv		*	*
→b. cv^{H}.cv^{L}.cv.cv			**
c. cv.cv.cv^{H}.cv^{L}	*!	***	

When we consider forms of five or more syllables, however, the analysis fails to make the right predictions. This is shown in the tableau in (74). The equal ranking of the two alignment constraints causes any form in which the accent falls on the antepenultimate accent or further to the left to be acceptable as an output.

(74)

cv.cv.cv.cv.cv	NONFIN	ALIGN-L (H)	ALIGN-R (L)
→a. cv.cv.cv^{H}.cv^{L}.cv		**	*
→b. cv.cv^{H}.cv^{L}.cv.cv		*	**
→c. cv^{H}.cv^{L}.cv.cv.cv			***
d. cv.cv.cv.cv^{H}.cv^{L}	*!	***	

The problem can be resolved by assuming that ALIGN-L(H) is not a gradient constraint, such as as ALIGN-R(L), but a categorical constraint. In order to distinguish between the two types of constraints, the new constraint is called STRICTALIGN-L(H). The way in which violations are assigned by the constraint is redefined in (75).[97]

(75) STRICTALIGN-LEFT(H, ω) = STRICTALIGN-L(H):
The left edge of the H tone is aligned with the left edge of the p-word.
(Categorical: one violation for the alignment of a H tone with any TBU that is not leftmost in the p-word)

The tableau in (76) shows that for words of four light syllables, the right predictions are made under the ranking NONFINALITY >> STRICTALIGN-L(H), ALIGN-R(L).

(76)

cv.cv.cv.cv	NONFIN	STRICT ALIGN-L (H)	ALIGN-R (L)
→a. cv^{H}.cv^{L}.cv.cv			**
→b. cv.cv^{H}.cv^{L}.cv		*	*
c. cv.cv.cv^{H}.cv^{L}	*!	*	

Furthermore, the same ranking also makes the right predictions for words of five or more light syllables, for which there is only one productive pattern, viz. antepenultimate accent (77). The categorical alignment constraint STRICTALIGN-L(H) is only satisfied in (77c), but this candidate fails to be selected because of its multiple violations of the gradient

[97] In a metrical approach, the distinction between gradient and strict alignment could also be adopted. This would improve Tanaka's (2008) analysis in which it is assumed that ALIGN-L is ranked lower in words of five or more light syllables.

constraint ALIGN-R(L). The candidates in (77a) and (77b) both violate STRICTALIGN-L(H) once, but the former is preferred because it violates ALIGN-R(L) only once, as opposed to the latter, which violates it twice.

(77) Only penultimate accent in words of five or more light syllables

cv.cv.cv.cv.cv	NONFIN	STRICT ALIGN-L (H)	ALIGN-R (L)
→a. cv.cv.cv^{H}.cv^{L}.cv		*	*
b. cv.cv^{H}.cv^{L}.cv.cv		*	**!
c. cv^{H}.cv^{L}.cv.cv.cv			***!
d. cv.cv.cv.cv^{H}.cv^{L}	*!	*	

Next, let us see whether we can account for pre-antepenultimate accent in words with a medial heavy syllable. In the tonal approach such words would have the following lexical tones in their surface representations.

(78)a. misa^{H}i^{L}ru ‘missile’
papa^{H}i^{L}a ‘papaya’
bini^{H}i^{L}ru ‘vinyl’
ore^{H}n^{L}zi ‘orange’
b. so^{H}ne^{L}Qto ‘sonnet’
bu^{H}ti^{L}Qku ‘boutique’
su^{H}ri^{L}Qpu ‘slip’
ma^{H}ji^{L}Qku ‘magic’

Comparing the forms in (78a) and (78b), it seems likely that some markedness constraint is responsible for ruling out forms like **sone^{H}Q^{L}to* with a L tone on the moraic obstruent. What is marked about such forms is of course the association of a tone to a non-sonorous mora (*Q*). Therefore, I assume that the constraint in (79) blocks L tones being linked with moraic obstruents.

(79) *T/$\mu_{[-son]}$:
A tone is not associated to a mora that dominates a [-sonorant] feature.
(A tone-bearing unit is a sonorant mora)

If we take a more careful look at the variation among quadrimoraic words with a medial heavy syllable, it turns out that this cannot be the whole story. According to Tanaka (2008), there are several words with the same structure as those in (78b) that take either the antepenultimate or the pre-antepenultimate pattern. What is more, depending on the word, either the antepenultimate (80a) or the pre-antepenultimate (80b) form is the preferred variant (Tanaka 2008:99-100).

(80)a.	koˀroQke ~ koroˀQke	‘croquette’
	roˀboQto ~ roboˀQto	‘robot’
	zyaˀkeQto~zyakeˀQto	‘jacket’
b.	kuriˀQpu ~ kuˀriQpu	‘clip’
	rakeˀQto ~ raˀkeQto	‘racket’
	suriˀQpu ~ suˀriQpu	‘slip’

Things get even more complicated if we take into consideration the fact that pre-antepenutimate accent in quadrimoraic words with a medial heavy syllable is also observed in cases where the heavy syllable ends in a moraic nasal (Tanaka 2008: 99).

(81)	iˀ.nin.gu	‘inning’
	haˀ.min.gu	‘humming’
	nyuˀan.su	‘nuance’
	puˀ.din.gu	‘pudding’,
	buˀ.ran.ko	‘swing’
	reˀ.gin.su	‘leggings’

The complicated situation is summarized in (82). The patterns in (82a/b/c-i) are based on words which do not seem to show variation. These are given including the frequency with which they are observed (Tanaka 2008: 97). The words in (82c-ii) are the patterns observed for words that do show variation. For these words no information on frequency is available.

(82)	*Syllable structure*		*Attested output forms*	
a.	CV.CVV.CV	→	CV.CV’V.CV	211/225 (94%)
b.	CV.CVN.CV	→	CV.CV’N.CV	
			(CV’.CVN.CV)	14/225 (6%)
c.i.	CV.CVQ.CV	→	CV’.CVQ.CV	
ii.	CV.CVQ.CV	→	CV’.CVQ.CV	(~ CV.CV’Q.CV)
			CV.CV’Q.CV	(~ CV’.CVQ.CV)

What this data first of all shows, is that at least the exceptions must be stored in the lexicon to the extent they are the same for all speakers. If we want to account for the data in terms of constraint ranking, as in Tanaka (2008), we could think of an analysis along the following lines.

The constraint *T/$\mu_{[-son]}$ may be held responsible for the tendency that at least some words take the pre-antepenultimate pattern. However, this pattern is not always chosen, so there must be a constraint that favors the antepenultimate pattern. Looking at a form like *ro^{H}bo^{L}Qto*, two things come to mind that could be wrong with it in terms of tone-syllable relations. First, it is possible that the problem is the association of a L tone with the head of a heavy syllable. L tones are unprominent in comparison to H tones, and heavy syllables are prominent in comparison to light syllables. In the variant form *robo^{H}Q^{L}to*, such an association is avoided. A second possibility is that the preference of H tones to associate to the head of a heavy syllable is what makes the variant form *robo^{H}Q^{L}to* acceptable. Because H tones are prominent, they prefer to associate with heavy syllables (Kim 1997, 2010; Gordon 2006). As pointed out by Gordon (2006), cross-linguistically, syllables of the type CVC$_{[-son]}$ tend to be treated as light. However, even if in a heavy syllable the second mora does not dominate sonorant material, the fact that it consists of two moras may still cause the H tone to be attracted to it.

In Chapter 2 I already introduced a constraint that regulates the alignment between H tones and heavy syllables with sonorant second moras in Tokyo Japanese. Apart from this constraint, which is repeated in (83a), we may also posit a constraint that aligns H tones to heavy syllables with non-sonorant second moras (83b).

(83) a. ALIGN(H, CVX$_{[+son]}$): Align a H tone with a syllable of the form CVC$_{[+son]}$.
b. ALIGN(H, CVC$_{[-son]}$): Align a H tone with a syllable of the form CVC$_{[-son]}$.

The natural ranking for these constraints is as follows: ALIGN(H, CVX$_{[+son]}$) >> ALIGN(H, CVC$_{[-son]}$).

Words of the type *so'neQto* 'sonnet', which have only one variant, can be regarded as the result of the constraint *T/$\mu_{[-son]}$ outranking ALIGN(H, $CVC_{[-son]}$), as in (84).

(84)

so.neQ.to	*T/$\mu_{[-son]}$	Align(H, $CVC_{[-son]}$)
→a. $so^{H}.ne^{L}Q.to$		*
b. $so.ne^{H}Q^{L}.to$	*!	

Words with two variants like *ro'boQto~robo'Qto* 'robot', on the other hand, can be accounted for if the two constraints *T/$\mu_{[-son]}$ and ALIGN(H, $CVC_{[-son]}$) have the same ranking, as in (85).

(85)

ro.boQ.to	*T/$\mu_{[-son]}$	Align (H, $CVC_{[-son]}$)
→a. $ro^{H}.bo^{L}Q.to$		*
→b. $ro.bo^{H}Q^{L}.to$	*	

In (86), the selection of an example of a quadrimoraic word with a medial CVV syllable is illustrated. The absence of variation in such words results from a violation of ALIGN(H, $CVX_{[+son]}$) in the candidate with pre-penultimate accent.

(86)

bi.nii.ru	STRICT ALIGN-L (H)	ALIGN-R (L)
a. $bi^{H}.ni^{L}i.ru$		**
→b. $bi.ni^{H}i^{L}.ru$	*	*

ALIGN (H, $CVX_{[+son]}$)
*!

In (86), the ranking of ALIGN(H, $CVX_{[+son]}$) is left indetermined. However, once we take the word *yooro'Qpa* 'Europe' into account, it turns out that it must have the same ranking as the two tonal alignment constraints. As we can see in (87), because the two tonal alignment constraints and ALIGN(H, $CVX_{[+son]}$) all three have the same ranking, the lower ranked ALIGN(H, $CVC_{[-son]}$) becomes decisive, ruling out the candidate with pre-penultimate (i.e. initial) accent.

(87)

yoo.roQ.pa	STRICT ALIGN-L (H)	ALIGN-R (L)	ALIGN (H, $CVX_{[+SON]}$)	Align(H, $CVC_{[-son]}$)
a. $yo^{H}o^{L}$.roQ.pa		***		*!
→b. yoo.$ro^{H}Q^{L}$.pa	*	*	*	

Quadrimoraic words with a medial heavy syllable ending in the moraic nasal are the final group of words we need to account for. For words of this shape, penultimate accent as in *oreˈnzi* 'orange' is the default pattern, with some exceptions like *puˈdingu* 'pudding'. Except for the word *buˈranko* 'swing', the words in (81) in which the H tone fails to be associated to the heavy syllable CVN could be the result of the preservation of the prominence in the source language. In other words, we may treat these words as exceptional. As for the only word that cannot be treated in this way, *buˈranko*, NHK (1998) also lists the form *buraˈnko*, which means there is at least a "regular form" for this word. The reason why an initially accented variant exists is left unexplained, but this is probably also true in a metrical approach.

Moving on to the variation that can be observed in words ending in a heavy syllable, an obvious solution seems one in which we change the domain of NONFINALITY from the mora to the syllable, as proposed by Tanaka (2013) in his approach based on the quantitative trochee discussed earlier in this chapter. In (88), associated tones are indicated in superscript.

(88) a. mazi^{H}sya^{L}<n> vs. ma^{H}zi^{L}<syan> 'magician'
b. seremo^{H}ni^{L}<i> vs. sere^{H}mo^{L}<nii> 'ceremony'
c. areru^{H}gi^{L}<i> vs. are^{H}ru^{L}<gii> 'allergy'
d. kaadi^{H}ga^{L}<n> vs. ka^{H}a^{L}di<gan> (*kaa^{H}di^{L}<gan>) 'cardigan'

The solution is straightforward for the forms in (88a/b/c). In the example in (88d), the H tone cannot be associated to the syllable non-head, resulting in a pre-pre-antepenultimate pattern. This is due to a constraint that prohibits H tones to be associated to syllable non-heads.

(89) *H/$NONHEAD_{\sigma}$: A H tone is not associated to a syllable non-head.

Obviously, this constraint should outrank ALIGN-R(L). Note that a similar (if not identical) constraint is necessary in foot-based approaches in which the syllable is also assumed to play a role.

From the above, we can conclude that the foot-based generalizations of loanword accent that have been proposed in the literature can be restated in terms of tone, and that a tone-based analysis is surely not inferior to a metrical one.

3.3.3. Compound accentuation

In this section, I will develop a reanalysis of Kubozono's (1995, 1997, 2011) generalizations of compound accent in terms of tone.[98] For the moment, I will assume that Kubozono's generalizations are empirically correct. Some examples of compounds are repeated in (90).

(90)	N1	N2	Compound	
a.	peˀrusya	neˀko	perusya)-(neˀko)	'Persian cat'
	saki^{0}	ototoˀi	saki)-(oto)(toˀi)	'three days ago'
	yaˀmato	nadeˀsiko	yamato)-(nadeˀ)(siko)	'Japanese lady'
	minami^{0}	amerika^{0}	minami)-(aˀme)(rika)	'South America'
b.	abare^{0}	umaˀ	abareˀ)-(uma)	'restive horse'
	neˀbada	syuˀu	nebadaˀ)-(syuu)	'the State of Nevada'
c.	kamiˀ	omuˀtsu	kami)-(o)(muˀtsu)	'paper diaper'
	mata^{0}	itoˀko	mata)-(i)(toˀko)	'a second cousin'

In order to make a fair comparison between the foot-based approach and the tone-based approach possible, the foot-based generalizations of (39) are reformulated as in (91). I have added a constraint that forces the accent to be aligned with the morpheme boundary between N1 and N2, as this constraint is essential to the success of Kubozono's (1995, 1997, 2011) analysis.

[98] This section is partly based on a discussion of Tokyo Japanese compound accentuation in Poppe (2014). In that paper, I proposed a representational solution (floating tones vs. linked tones) to account for the variation in compounds with short N2s. However, this analysis does not seem to be available for the similar type of variation in compounds with long N2s.

(91) Foot-based generalizations (reformulated):

a. Preserve the accent of N2 in compounds unless it is on the final syllable.
b. If the N2 is unaccented or accented on the final syllable, place a default compound accent:
   (i) on the rightmost foot;
   (ii) that is not in compound-final position;
   (iii) while aligning the accent with the boundary between N1 and N2.

The foot-based generalizations can be translated into tone-based ones as follows.

(92) Tone-based generalizations:

a. Preserve the associated tones of N2 unless a H tone is linked to the final syllable.
b. If the N2 contains no tone or a H tone on the final syllable, associate the lexical tones (HL)
   (i) as far to the right as possible;
   (ii) avoiding the final mora;
   (iii) while aligning H with the boundary between N1 and N2 *or* while aligning any tone (H or L) with the right edge of N2.

Either of the two generalizations in (92b-iii) is sufficient for our purposes. What is important is that the foot-based and the tone-based generalizations make use of exactly the same number of conditions.

The generalizations in (92) can be formalized in terms of constraints in the following way. The preservation of tones in N2 is regulated by the faithfulness constraint in (93), which refers to the head p-word, which in the case of compounds is equivalent to N2.[99]

(93) PARSE-ω_{HD}(Tone) The head p-word preserves the input tones of the corresponding input form.

The non-faithfulness to H tones that are associated with the final syllable in the N2 (92a) can be accounted for by the constraint in (94).

(94) NONFINALITY-H(σ) A H tone is not associated to the final syllable.

[99] The constraint in (92) can be seen as the prosodic counterpart of faithfulness to morphological heads (Revithiadou 1999). See Alderete (1995) for instances of faithfulness to lower-level prosodic heads.

This constraint outranks PARSE-ω_{HD}(Tone), and can be assumed to be undominated in the Yamato and Sino-Japanese strata, and will be omitted in the tableaux of words of the Yamato stratum.

Apart from this constraint, there is another non-finality constraint at work that was already introduced in the analysis of loanword accent, NONFINALITY-TONE(μ). The generalizations in (92b-i/ii) follow from the ranking that was also established for the default accent in simplex words: NONFINALITY-T(μ) >> ALIGN-R (L).

The generalization stated in (92b-iii) could be formalized in several different ways. One solution would be to simply mimic Kubozono's (1995) ALIGN-CA and posit the following constraint, where CT stands for Compound Tone.

(95) ALIGN-CT: (i) A H tone is aligned to some edge of N2. *or*
(ii) The HL complex is aligned to some boundary of N2.

While adopting this constraint is certainly possible, a disadvantage is that it refers to "N2", which can be seen as a shorthand for "the head of a compound". It would be preferable if our constraint would refer to a more general category, and if possible a phonological one rather than a morpho-syntactic one. This is possible if we make use of the constraint in (96a), which could be seen as a more abstract version of the constraints in (96b) and (96c).

(96) a. ALIGN(Tone, Left, ω, Right) = ω]T
The left edge of a tone is aligned with the right edge of a p-word.
b. ALIGN(H, Left, ω, Right) = ω]H
The left edge of H tone is aligned with the right edge of a p-word.
c. ALIGN(L, Left , ω, Right) = ω]L
The left edge of a L tone is aligned with the right edge of a p-word.

I assume the constraint is evaluated categorically: either a tone is aligned to the right edge of the p-word, or it is not. The constraint must be ranked below PARSE-ω_{HD}, because it does not cause unfaithful parsings of the input tones of N2. Furthermore, it must be ranked above the alignment constraints ALIGN-R(L) and ALIGN-L(H),[100] because aligning the H and L input tones to the left and right edges of a disyllabic N1 as in (93c) below is not allowed.

The variation described in the foot-based analysis discussed above can be accounted for by switching the ranking of PARSE-ω_{HD} and NONFIN-T(μ), as shown by the example

[100] These two alignment constraints are assumed to have the same ranking as they have in loanwords.

mata-itoko ‘a second cousin’ in (97) and (98) (where the square brackets indicate p-word boundaries). In Ranking A in (97), where NONFIN-T(μ) is outranked by PARSE-ω_{HD}, the tones remain in place. In (98), however, NONFIN-T(μ) causes the tones to move to the left, and ω]T forces the tones to align with the left boundary of N2.

(97) ‘a second cousin’ (Ranking A)

mata#ito^{H}ko^{L}	PARSE -ω_{HD}	NONFIN -T(μ)	ω]T	ALIGN -R(L)	ALIGN -L(H)
→ a. [[ma.ta.][i.to^{H}.ko^{L}]]		*	*		*
b. [[ma.ta.][i^{H}.to^{L}.ko]]	*!			*	
c. [[ma^{H}.ta^{L}.][i.to.ko]]	*!				

(98) ‘a second cousin’ (Ranking B)

mata#ito^{H}ko^{L}	NONFIN -T(μ)	PARSE -ω_{HD}	ω]T	ALIGN -R(L)	ALIGN -L(H)
a. [[ma.ta.][i.to^{H}.ko^{L}]]	*!				*
→ b. [[ma.ta.][i^{H}.to^{L}.ko]]		*		*	
c. [[ma^{H}.ta^{L}.][i.to.ko]]		*	*!		

Evidence for the demotion of PARSE-ω_{HD} to a position below NONFIN-T(μ) comes from the variation of compounds with short N2s. When PARSE-ω_{HD} outranks NONFIN-T(μ), as in (99), the accent is parsed as in the isolation form.

(99) ‘Persian cat’ (Ranking A)

pe^{H}ru^{L}sya#ne^{H}ko^{L}	PARSE-ω_{HD}	NONFIN-T(μ)
→a. [[pe.ru.sya][ne^{H}.ko^{L}]]		*
b. [[pe.ru.sya^{H}.][ne^{L}.ko]]	*!	

However, in (100), it is NONFIN-T(μ) that rules out candidate (100a) in favor of (100b).

(100) ‘Little Mermaid’ (Ranking B)

ni^{H}n^{L}gyo#hi^{H}me^{L}	NONFIN-T(μ)	PARSE-ω_{HD}
a. [[nin.gyo][hi^{H}.me^{L}]]	*!	
→b. [[nin.gyo^{H}.][hi^{L}.me]]		*

Both in Kubozono's (1995, 1997) approach, as well as in the tonal approach developed here, the reranking of constraints is only minimal. I would like to point out that what Kubozono calls the "Little Mermaid pattern" of (100) is becoming more dominant among speakers of younger generations. While I know of speakers that pronounce both forms with the Little Mermaid pattern, I do not know of speakers that pronounce both forms with the "Persian cat pattern". Thus, the grammar in (100) represents the innovative variety.

One may wonder whether we really need reranking of constraints, a question that is relevant for both the foot-based and the tonal analyses. The co-phonology approach (Orgun 1996; Antilla 1997; Inkelas 1998, 2007, among others) to variation has been criticized for its excessive power (Scheer 2011). To constrain the power of the co-phonology approach, it has been proposed that different rankings are dominated by a "master ranking" (Antilla 1997). In the case at hand, this would be a situation in which the constraints PARSE-ω_{HD} and NONFIN-T(μ) have the same ranking. Different words may "choose" one of the sub-rankings of this master ranking.

Another approach would be to assume the constraint PARSE-ω_{HD} is an OO constraint, as in Nishimura (2013). Nishimura (2013) proposes that there are two classes of words: those indexed for a high-ranked OO constraint, which is assigned to words by default, and those for which the learner has concluded that they are subject to a lower-ranked OO constraint. This approach is attractive in that it can account for the behavior of loanwords. As pointed out by Kubozono (1997) and Tanaka (2001), the accent of the isolation form of loanwords is parsed faithfully in compounds even if it falls on the final syllable (e.g. *suna'kku* 'snack' + *ba'a* 'bar'→ *sunaku-ba'a* 'snack bar'). In the OO-correspondence approach, this is the result of the assignment to the default high-ranked OO class when the word enters the language. In the co-phonology approach, PARSE-ω_{HD} can be assumed to outrank NONFINALITY-H(σ) in the loanword stratum.[101]

As for the peculiar behavior of compounds with overlong N2s, I refer to the discussion in 3.2.3. In a non-metrical analysis, the only choice we seem to have is to adopt Kubozono's (2004) morphology-based analysis. The consequence of this is that in a tone-only approach to Tokyo Japanese accent, the Indirect Reference Hypothesis cannot be maintained, unless the status of the pseudo-morphemes is clarified in terms of prosodic constituents.

[101] As Nishimura (2013: 175) points out, the assignment of an index to a highly ranked OO-correspondence constraint makes sense when we realize that relatively unfamiliar words are more prone to misunderstanding than more familiar words. The same holds for the assignment of a default co-phonology in which faithfulness is highly ranked.

3.3.4. Unaccentedness

In 3.2.4, Ito and Mester's (2012) proposal that unaccentedness is a consequence of a particular type of metrical parsing was reviewed. Following this, some functionalist explanations were also discussed. In a purely tonal approach, it seems that such functionalist explanations are the only possible explanation. Therefore, if the analysis proposed by Ito and Mester (2012) can be modified in such a way that it also provides us with a convincing explanation for the variation observed among accented loanwords, unaccentedness would speak in favor of the metrical approach.

3.3.5. Accent and high vowel devoicing

In 3.2.5 I demonstrated that there is no evidence for a role of foot structure in accounting for the shift of accent as the result of vowel devoicing. The only thing that might be worth mentioning about a tonal approach in relation with this phenomenon is that a constraint against H tones associated with non-sonorant moras would be a particular instance of the constraint *T/$\mu_{[-son]}$, which was introduced above in (75).

3.4. Summary: foot structure, tone, or both?

In this chapter, the evidence for metrical structure from the accent patterns of Tokyo Japanese was critically evaluated, and where existing alternatives were available, these were discussed. Furthermore, a tone-only analysis of default loanword accent and accent in non-phrasal compounds was developed. I showed that at least for these two phenomena, the tonal approach is not necessarily inferior to the metrical approach. The two tonal analyses that were developed are perfectly in harmony with each other, which shows that at least for these two phenomena, a tonal approach can be considered seriously as an alternative. Still, the tonal analysis is based on two specific assumptions about L tone association that require further investigation: (i) L tones may associate to moraic obstruents, and (ii) the trailing L tone of the HL pitch accent associates to a single TBU. If the first assumption turns out to be invalid, a tone-only analysis can only be maintained if the location of tones is computed in terms of alignment without association, which may be problematic.

By making reference to the foot we are able to explain not only accent in loanwords and word-level compounds, but also phrase-level compound accent and unaccentedness. For these two phenomenon, a non-metrical tonal approach is only possible if we seek for explanations outside the phonological grammar (the lexicon and morphology), which speaks in favor of the metrical approach. Even so, it is not clear to what extent the analyses of the different phenomena presented by different researchers are compatible with each

other. Therefore, based on the Tokyo data alone, the metrical approach cannot, at this point, be conclusively shown to be superior to the tonal approach.

A third possible line of analysis in which both metrical structure and tone are relevant was not considered, because this would involve the comparison of a lot of different ways in which the two can be combined. However, in theory such an approach is a perfectly valid option. In fact, we did see that in the metrical analysis proposed by Ito and Mester (2012), a constraint directly referring to tone actually plays a crucial role. Combining insights of metrical approach and the tonal approach developed in this chapter may yield interesting results.

To sum up, this chapter has shown that although a metrical approach to Tokyo Japanese accentuation seems to do better than a purely tonal approach in terms of empirical scope, a number or issues remain as to what type of feet are actually allowed and what the relative roles of feet and tone are.

Still, the different proposals that have been made can function as a guide in our search for further evidence for metrical structure and/or tone being active in the lexical phonology. In the following two chapters, a number of dialects will serve as case studies in which the relative roles of metrical structure and tone in Japanese accentuation are investigated further.

4. The Maisaka variety of Shizuoka Japanese: a counting system

In this chapter, the nature of Japanese accentuation is investigated based on data from a dialect that is closely related to Tokyo Japanese, spoken in Shizuoka Prefecture in the former town of Maisaka, now part of Hamamatsu City. All Maisaka data presented in this chapter are taken from Yamaguchi (1984). The generalizations stated by Yamaguchi (1984) are in line with the data in Maekawa (1970) and Terada (1970), although these studies are less comprehensive.

The Maisaka dialect is relevant to this thesis in that its accent system is what we may call an "odd-even counting system", i.e. an accent system in which the location of the accent is (partly) determined by whether a word has an even or odd number of moras and syllables. For many languages, such counting has been shown to be the result of metrical foot structure (see Hayes 1995). As far as I know, in this sense Maisaka Japanese is unique among Japanese dialects.[102]

4.1. Data

4.1.1. Nouns

In (1), the accent patterns of nouns consisting of one to five moras are displayed.

(1) Maisaka Japanese nominal accent (based on Yamaguchi 1984)

	1 mora	2 moras	3 moras	4 moras	5 moras
0	ha^{0} 'tooth'	kawa^{0} 'river'	suzume^{0} 'sparrow'	itazura^{0} 'mischief'	atarimae^{0} 'ordinary'
1	teˀ 'hand'	miˀmi 'ear'	kaˀiko 'silkworm'	oˀokami 'wolf'	kiˀnkakuji 'Kinkakuji (temple)'
2		(-)	sugaˀta 'figure'	muraˀsaki 'purple'	migiˀhidari 'right-left'
3			atamaˀ 'head'	asagaˀo 'morning glory'	(-)
4				otootoˀ 'younger brother'	daidokoˀro 'kitchen'
5					arigatamiˀ 'value'

[102] See Lawrence (1990) for a counting system in the Nakijin Okinawan variety of Ryukyuan.

In Maisaka Japanese, starting with words from two to four moras, there are at least three important generalizations that need an explanation. First of all, all accented bimoraic nouns have an initial accent (as in (2a)) in their isolation forms. Secondly, word-initial accent in words with more than two syllables is only allowed if the syllable in question is bimoraic ((2b-i) and (2c-i)). Thirdly, and most importantly, words that have (i) an even number of syllables and (ii) penultimate accent in isolation, undergo an accent shift one syllable to the right when they are followed by a particle. Examples of this accent shift are given in (2a) and (2c-iii).

(2)

a.	**a**ˈ.me	‘rain’	a.m**e**ˈ.-ga	‘rain-NOM’
b.i.	kaˈi.ko	‘silkworm’	kaˈi.ko.-ga	‘silkworm-NOM’
ii.	su.gaˈ.ta	‘figure’	su.gaˈ.ta.-ga	‘figure-NOM’
iii.	a.ta.maˈ	‘head’	a.ta.maˈ.-ga	‘head-NOM’
c.i.	oˈo.ka.mi	‘wolf’	oˈo.ka.mi.-ga	‘wolf-NOM’
ii.	mu.raˈ.sa.ki	‘purple’	mu.raˈ.sa.ki.-ga	‘purple-NOM’
iii.	as.a.**ga**ˈ.o	‘morning glory’	a.sa.ga.**o**ˈ.-ga	‘morning glory-NOM’
iv.	o.too.toˈ.	‘younger brother’	o.too.toˈ.-ga	‘younger brother-NOM’

According to Yamaguchi (1984), *kaˈiko* and *oˈokami* are underlyingly /kaiˈ.ko/ and /ooˈkami/, with a constraint against accents on non-head moras causing the accent to shift to the left. Under this analysis, there are no initially accented nouns longer than two syllables at all. While this is an interesting proposal, it has two problems. First of all, there is the abstractness problem: should we allow underlying representations that differ from surface representations in the absence of alternations? A second and more serious problem is revealed when we take words of four moras into account. Yamaguchi argues that since *asagaˈo* and *otootoˈ* are in complementary distribution; the underlying form of *otootoˈ* can be analyzed as *otooˈto*. However, if this is true, the question is why in this case the accent would shift to the right rather than to the left (**otoˈoto*), as in *kaˈiko* and *oˈokami*. Therefore, it seems better to simply assume that *kaˈiko, oˈokami,* and *otootoˈ* are specified in the lexicon as they appear on the surface. Still, we do need an explanation for the fact that there are (i) no words consisting of an even number of syllables, ending in two light syllables, with final accent ($\mu\mu\mu_\sigma\mu_\sigma$ˈ) in isolation, and (ii) no words consisting of more than two syllables, starting with two light syllables, with initial accent (*μ_σˈ$\mu_\sigma\mu$..).

The non-accidental gaps in the accent system that become apparent when we ignore accented heavy syllables are highlighted in (3). Numbers denoting the accent location (in the leftmost column) refer to the n^{th} mora counted from the left.

(3)

	1 mora	2 moras	3 moras	4 moras	5 moras
0	ha^0	kawa^0	suzume^0	itazura^0	atarimae^0
	‘tooth’	‘river’	‘sparrow’	‘mischief’	‘ordinary’
1	teˀ	miˀmi	*μσˀμσμ	*μσˀμσμμ	*μσˀμσμμμ
	‘hand’	‘ear’			
2		*μμˀ	sugaˀta	muraˀsaki	migiˀhidari
			‘figure’	‘purple’	‘right-left’
3			atamaˀ	asagaˀo	*μμμσˀμσμ
			‘head’	‘morning glory’	
4				otootoˀ	daidokoˀro
				*μμμσμσˀ	‘kitchen’
				‘younger brother’	
5					arigatamiˀ
					‘value’

Words of five or more moras generally consist of multiple morphemes. It is interesting to see that depending on the position of the morpheme boundary, accent shifts do or do not occur. In (4a), where the second morpheme consists of three moras, no accent shift is observed. In (4b), where the second morpheme is made up of two moras, the accent does shift to the right when a particle is attached.

(4) a. dai-dokoˀro ‘kitchen’
dai-dokoˀro-ga ‘kitchen-NOM’
b. abura-m**u**ˀsi ‘plant louse’
abura-mu**si**ˀ -ga ‘plant louse-NOM’

Finally, monomoraic words are special in that they exhibit accent shifts even though they have an odd number of moras. As the examples in (5) makes clear, the accent shifts when a bimoraic particle is attached, but not when a monomoraic particle follows.

(5) a. teˀ ‘hand’
b. teˀ-ga ‘hand-NOM’
c. te-kaˀra ‘hand-ABL’

Thus, another question that must be answered is why monomoraic forms pattern with bimoraic and quadrimoraic words rather than trimoraic ones.

4.1.2. Verbs

The same generalizations that hold for nouns also apply to verbs. In (6), forms of verbs with a non-past form of two moras are given (henceforth referred to as bimoraic verbs).[103] The example contains two verbs: one verb with a root ending in a vowel (V-verb), and one verb with a root ending in a consonant (C-verb).

(6)

	mi- ‘see’	kak- ‘write’
non-past	miˀru	kaˀku
transitional	mir**u**ˀto	kak**u**ˀto
conditional I	mirya**a**ˀa	kakya**a**ˀa
imperative	my**o**ˀo	kaˀke
past	miˀta	kaˀita
conditional II	mit**a**ˀra	kaˀitara
gerund-foc	mit**e**ˀmo	kaˀitemo
negative	miˀn	kakaˀn

Notice that in (6), the accent shifts to the right in all forms in which the shift does not result in a word-final accent (**miruˀ*) or an accent on a dependent mora (**kaiˀta*).

The forms in (7) are examples of “trimoraic verbs”. As with trimoraic nouns, in general no accent shifts can be observed. The negative forms are exceptional in that they show a shift of the accent in all cases. However, this can be regarded as an idiosyncratic property of the negative suffix, which also shows peculiar behavior in terms of accent in Tokyo Japanese (see e.g. McCawley 1968).

[103] The glosses of the realized grammatical exponents are adapted from Otoguro’s (2009) study on Tokyo Japanese.

(7)

	tabe- ‘eat’	aruk- ‘walk’
non-past	tabeˀru	aruˀku
transitional	tabeˀruto	aruˀkuto
conditional I	tabeˀryaa	aruˀkyaa
imperative	ta.byoˀo	aruˀke
past	tabeˀta	aruˀita
conditional II	tabeˀtara	aruˀitara
gerund-foc	tabeˀtemo	aruˀitemo
negative	tabeˀn	arukaˀn

However, when we turn to the quadrimoraic verbs, just like in nouns of the same length, accent shifts can again be observed. The examples in (8) are transitive (tr.) and intransitive (intr.) forms sharing the same root.[104]

(8)

	atume- ‘gather (tr)’		atumar- ‘gather (intr.)
present	atumeˀru	=	atumaˀru
transitional	atumer**u**ˀto	=	atumar**u**ˀto
conditional I	atumery**a**ˀa	=	atumary**a**ˀa
past	atumeˀta	=	atumaˀtta
conditional II	atumet**a**ˀra	≠	atumaˀttara
gerund-foc	atumet**e**ˀmo	≠	atumaˀttemo
negative	atumeˀn	=	atumeˀn

The same generalizations hold for adjectives, which are not treated in this thesis.

To summarize, the following questions require an answer.

(9) a. How can we explain the gaps in the accent system in (3)?

b. Why does the accent only shift in (i) forms with penultimate accent and an even number of moras/syllables, and (ii) monomoraic words?

[104] The imperative form of these verbs is not given by Yamaguchi (1984).

4.2. A metrical analysis of the accent patterns

4.2.1. Nouns

Based on the observation that a distinction is made between even and odd numbers of moras/syllables, we may expect that foot structure is involved. The question then is what this foot structure looks like and what constraints it is subject to.

Let us start with words with an even numbers of moras and syllables, as the foot structure of such words should be relatively straightforward because we do not have the problem of possible leftover syllables. Based on the claims made for Tokyo Japanese, let us assume that feet are bimoraic. The foot structure of *aˈme*, *asagaˈo*, and *muraˈsaki* would then be as in (10). The prosodification of particles is ignored for the moment.

(10)

a.	(**a**ˈ.me)	'rain'	(a.m**e**ˈ).-ga	'rain-NOM'
b.	(a.sa.)(**ga**ˈ.o)	'morning glory'	(a.sa.)(ga.**o**ˈ.)-ga	'morning glory-NOM'
c.	(mu.raˈ.)(sa.ki)	'purple'	(mu.raˈ.)(sa.ki.)-ga	'purple-NOM'

If these are the metrical parsings of words with an even number of moras and syllables, we now can say that, from a derivational point of view, accents only shift within a foot (10a/b), but not across foot boundaries (10c).

There are two potential reasons why accents shift from penultimate to ultimate position only when a particle follows. First, this could be due to a high-ranked constraint against a final accentual H tone on the final mora NONFINALITY(μ) within the domain of the (minor) phrase.[105] Second, perhaps a shift to the domain-final mora is avoided to save the L part of the HL pitch accent from being deleted or being left unassociated. Because it is not clear from the literature what happens with the trailing L tone in words with word-final accent, I will simply assume the NONFINALITY constraint.

(11) NONFINALITY(μ)$_{\Phi}$

The accentual H tone is not associated with the last mora of a p-phrase.

Next, we move on to the important question of why an accent would shift to the right in the first place. Again, there are at least two possible explanations. In one analysis, rightward accent shifts are triggered by a general pressure for the accentual H tone to be

[105] The reason why the domain would be the p-phrase rather than the p-word is that in compounds like *abura-muˈshi* 'plant louse' (see 4b), which according to the discussion in Chapter 2 would be a maximal p-word, the accent shifts when a particle is attached, i.e. *[[abura][musiˈ]]-ga*.

aligned as far to the right of the p-word as possible, i.e. the (gradient) constraint ALIGN-R(H, ω).

A second possible reason is that tones do not prefer to be aligned to the right edge of the word, but to the right edge of a foot. To make things more complicated, this could be either due to a preference for iambic feet in favor of trochaic feet, or due to a preference for a H tone to be aligned with the rightmost syllable of a foot, without caring about left- or right-headedness.[106] Because it does seem that the Maisaka dialect has unbalanced light-heavy iambic feet (e.g. *(mirya˺a)* 'see-$COND_I$') as well as monosyllabic feet (e.g. *(ka˺i)(tara)* 'write-$COND_{II}$'), adopting a constraint IAMB (12) does not seem to be inappropriate.

(12) IAMB
Feet are right-headed under syllabic analysis.
Allowed: $(\sigma_\mu\sigma_{\mu\mu})$, $(\sigma_\mu\sigma_\mu)$, (σ_μ), $(\sigma_{\mu\mu})$
Not allowed: $(\sigma_{\mu\mu}\sigma_\mu)$

However, we don't know whether a word like *(a˺me)* 'rain' has a left-headed foot with H associated to the foot-head or a right-headed foot with the H tone associated to the non-head of the foot. In other words, is headedness partially determined by the H tone or is it a separate kind of status that is determined by constraints unrelated to tone? The two different options are illustrated in (13) by the word *ame* 'rain', in which foot heads are underlined.

(13) a. $(\underline{a}^H.me^L)$ b. $(a^H.\underline{me}^L)$

Unfortunately, it is not possible to answer this question. However, I submit that in cases in which the only possible evidence for headedness is the pitch accent, as in a case like (13), it is more natural to assume that the head is the syllable to which the H tone is associated. Therefore, I assume the representation in (13a).

This means that the constraint IAMB is dominated by the constraint ACCENT-TO-HEAD (Katayama 1998: 189), which is responsible for the association of the accentual H tone with a foot-head.

[106] As discussed in Chapter 3, in Katayama (1998) feet are assumed to be left-headed in all cases, with H tones being able to associate to foot-non-heads and even unfooted syllables.

(14) ACCENT-TO-HEAD = A-TO-HEAD
The accentual H tone is associated to the head of a foot.

The constraint A-TO-HEAD can be assumed to be undominated in the grammar. Therefore, it must outrank NONFINALITY(μ), which is violated in words like *atamaˀ* 'head'. The interaction of the three constraints is exemplified in the tableaux in (15), where foot-heads are again underlined. As it is not necessary to distinguish between H and L tones in the tableaux, the HL pitch accent is indicated by the accent mark "ˀ", which of course stands for the accentual H tone.[107]

(15) a.

aˀme	A-TO-HEAD	NONFIN(μ)	IAMB
a. (aˀ.me)	*!		
b. (a.meˀ)		*!	
→ c. (aˀ.me)			*

b.

aˀme-ga	A-TO-HEAD	NONFIN(μ)	IAMB
a. (aˀme)ga	*!		
→b. (ameˀ)ga			
c. (aˀme)ga			*!

In (15a), the candidate with an iambic foot is rejected because of the constraints A-TO-HEAD and NONFINALITY(μ). When a particle is attached, as in (15b), however, the accentual H can be associated to a foot with a right-aligned head without violating NONFINALITY(μ).

Of course the accent shift involves a violation of a faithfulness constraint. This constraint is NOFLOP-PROM (Alderete 1999: 18, 2001: 16), which is defined in (16).

(16) NOFLOP-PROM = NOFLOP
Corresponding prominences have corresponding sponsors and links.[108]

[107] Although accented words of two light syllables show only a single pattern, I assume the accent is specified in the input. The reason for this is that this is necessary anyway for longer words, as there is no data from loanwords (or other sources) to determine what the default pattern is.
[108] Adapted from (Alderete 2001: 216). See Alderete (1999) for formal definitions of NOFLOP constraints. The NOFLOP constraint could also be defined in terms of tone. However, because there is

NOFLOP must of course be ranked below IAMB, as in (17), since otherwise no accent would ever shift in order to create an iamb.

(17)

aˈme-ga	A-TO-HEAD	NONFIN(μ)	IAMB	NOFLOP
a. (aˈme)ga	*!			
→b. (ameˈ)ga				*
c. (aˈme)ga			*!	

For words that do not exhibit accent shifts, the accentual H must be located rightmost within the foot in both isolation forms and suffixed forms. This is the case not only for quadrimoraic/quadrisyllabic words like *muraˈsaki* (18a), but also for trimoraic/trisyllabic words like *sugaˈta* 'shape' in which the accent does not shift when particles follow (18b).

(18) a.i. (muraˈ)(saki)
ii. *mu(rasaˈ)ki
b.i (sugaˈ)ta vs. (sugaˈ)ta-ga
ii. (*)su(gaˈta) vs. *su(gataˈ)-ga

It is of crucial importance not only that the initial syllable is footed, but also that it is binary. As we can see in (18b), if we were to assume the foot structure in (18b-ii), we would expect the accent to shift when followed by the nominative particle. Therefore, the parsing must be as in (18b-i).

The next challenge is to explain why the final accent of atama does not shift to the left. The tableau in (19) shows that this should be possible under the present ranking in which NONFINALITY(μ) outranks NOFLOP.[109]

no need to distinguish between H tone shifts and L tone shifts NOFLOP-PROM is adopted.

[109] The undominated constraint A-TO-HEAD and candidates violating this constraint are omitted in this and all following tableaux in this chapter where the constraint is not relevant.

(19)

atamaˀ	NONFIN(μ)	IAMB	NOFLOP
←a. (ata)(maˀ)	*!		
→b. *(ataˀ)ma			*
c. a(taˀma)		*!	*
d. a(tamaˀ)	*!		

The actual output form in (19a) is rejected in favor of the unattested form in (19b). What could be the reason for this? In a derivational analysis, feet can be said to built first, after which the accent may only shift within a foot. To develop a non-derivational account, let us focus on the output that should have won (19a), spell out its crucial properties, and translate these into constraints. Looking at the form in (19a), we can see that (i) the final syllable must be a monomoraic foot, and (ii) the accentual H tone must be located in this same foot.

To ensure that the final mora is a monomoraic foot, we need a constraint on foot-alignment. A rather standard approach would be to assume that the direction of footing is determined by ALL-FEET-RIGHT, which in languages with degenerate feet favors word-initial binary feet, and word-final degenerate feet (see Crowhurst and Hewitt 1995b). Here I adopt the definition proposed by Hyde (2012: 803), to which I added the informal definition that each foot be aligned with the right edge of a p-word.

(20) ALL-FEET-RIGHT: *ω, F, σ/[...F...σ...]$_{ω}$
Align each foot with the right edge of a p-word.
("Assess a violation mark for every ω, F, σ such that F precedes σ within ω.")

As demonstrated by Crowhurst and Hewitt (1995b), it takes exhaustive parsing (PARSE-σ >> FT-BIN) for this constraint to yield initial binary feet. In the absence of evidence against exhaustive foot parsing, we may simply assume that PARSE-σ indeed outranks FT-BIN. However, I prefer to take a conservative approach to metrical parsing in which degenerate feet are kept to a minimum (FT-BIN >> PARSE-σ). Therefore, I adopt an additional constraint that penalizes unfooted syllables before feet. The constraint in (21) is one of the "adjacent alignment" constraints proposed by Houghton (2013).[110]

[110] Houghton (2013) distinguishes between "adjacent alignment" and "between alignment". Both constraints are defined in terms of interveners (here: an unparsed syllable) that precede or follow the aligned category (here: a foot). As the name suggests, adjacent alignment is concerned with interveners

(21) *-o-/F

A foot is not preceded by an unparsed syllable ("-o-") within a p-word.

(Assign one violation for every F that is directly preceded by -o-).

Obviously, *-o-/F must outrank ALL-FEET-RIGHT. Under this ranking, feet will try to be aligned with the right edge of the p-word as much as possible, while avoiding an unparsed syllable. The result of this is an initial binary foot. The interaction of the two alignment constraints with FT-BIN and PARSE-σ is illustrated by the tableau in (22).[111] The fate of the (unattested) candidate **(ataˀ)ma*, which is not considered in (22), is discussed later on in this section.

(22)

atamaˀ	A-TO-HD	*-o-/F	FT-BIN	PARSE-σ	ALL-FT-RT
→a. (ata)(maˀ)			*		*
b. a(tamaˀ)		*!		*	
c. (ata)maˀ	*!			*	
d. (a)(tamaˀ)			*		**!

Note that the initial unaccented covert foot is crucially important in this analysis. Therefore, if the metrical analysis is on the right track, Maisaka Japanese provides us with important evidence for the existence of covert feet.

In (22), the constraint A-TO-HEAD forces the final syllable to be parsed into a monomoraic foot. However, in words with non-final accent such as sugata 'shape', the final syllable is left unparsed, as can be seen in (23).

(23)

sugaˀta	*-o-/F	FT-BIN	PARSE-σ	ALL-FT-RT
→a. (sugaˀ)ta			*	*
b. (sugaˀ)(ta)		*!		*

that are directly adjacent to the aligned category. Between alignment, on the other hand, is about an intervening element that is anywhere between the aligned category and the edge. The edge to which the category in question is aligned is the p-word. As in this particular case it does not matter whether we choose a constraint defined in terms of adjacent alignment or between alignment, the former is adopted as it is of a more local nature.

[111] PARSE-σ is assumed to be ranked above ALL-FEET-RIGHT under the assumption that binary feet will be built as long as there are enough leftover syllables. However, nothing hinges on this assumption, and we might as well assume that PARSE-σ has a lower ranking than that in (20), which would be possible due to the work that A-TO-HEAD and *-o-/F do in this analysis.

Now that we have dealt with the assignment of foot structure in words with an odd number of syllables, let us continue with the problem of why the accent does not shift to the left in order to avoid a violation of NONFINALITY(μ).

Assuming the foot structure for final-accented words in (22a), what seems to be needed is a constraint that penalizes accent shifts across foot boundaries. There are two types of constraints that we could use here. First, we could assume that in the learning process, foot structure is built as in (23), with the accent staying in place because there is no evidence for alternations. This can be formalized in terms of an OO-correspondence constraint that forces the accented mora/syllable in the input to be part of the accented foot in the output.

(24) FAITH-OO(F᾽):

If a pair of outputs stands in an OO-correspondence relation, the accentual-H-bearing foot dominates the same syllables in both members of the pair.[112]

A second option is to make use of a constraint that penalizes forms in which the accented syllable in the input is not part of the accented foot in the output. Cassimjee & Kisseberth (1998) propose the constraint INCORPORATE(F-Sponsor) in Optimal Domains Theory, where "F" stands for "feature". For Maisaka Japanese, we may posit the constraint in (25), where the word "incorporate" is replaced by "parse". The "sponsor" is the unit to which the H tone is linked in the input. In this case this is the accented mora and/or syllable.

(25) PARSE(H-Sponsor):

The sponsor of a H tone in the input is incorporated in the foot that dominates the H tone in the output.[113]

The only difference between the two constraints is that whereas the one in (24) compares two output forms, the one in (25) compares the input with the output form. Here I adopt PARSE(H-Sponsor), for ease of exposition.[114] The role this constraint plays in the

[112] See Davis (2005) for a similar constraint. Also note that it may be possible to refer to the "accentual-H-bearing foot" as the "head foot", i.e. the head of a prosodic word. However, if the head foot of a word is the accent-bearing foot, the question is how we determine the head foot of unaccented words.

[113] See also Bickmore (1996) for tone-sponsor alignment and Kim (2002) for foot-sponsor alignment constraints to account for tone displacement in restricted tone languages of the Bantu family. In an approach that is based on underlying diacritic accents, the constraint could be redefined so that it refers to the sponsor of the diacritic accent.

[114] In Chapter 6 it will be shown, however, that by adopting FAITH-OO(F᾽) we can account for the accent patterns of verbs, which do not have a specific sponsor syllable.

grammar is exemplified in the tableau in (26).

(26)

atamaˀ	PARSE (H-Sponsor)	*-o-/F	FT-BIN	PARSE-σ	ALL-FT-RT
→a. (ata)(maˀ)			*		*
b. a(tamaˀ)		*!		*	
c. (ataˀ)ma	*!			*	
d. (a)(tamaˀ)			*		**!

When nouns are followed by particles of more than two moras, the accent does not shift to these particles. This indicates that particles in general fall outside the domain of foot assignment for nouns.

(27) a. atamaˀ-ga ‘head-NOM’
b. atamaˀ-demo ‘even a head’

A straightforward explanation for this is that particles fall outside the p-word as the result of an alignment constraint like in (28).

(28) ALIGN-R(p-word, m-word) = ALIGN-R(ω, W)
The right edge of a p-word is aligned with the right edge of an m-word.

The tableau in (29) shows how the form in (27b) is generated by the grammar.[115] Candidates that violate Proper Bracketing (Nespor & Vogel 1986) are not considered, nor are candidates that place the accent on a foot that does not conain the sponsor (violating violate PARSE(H-Sponsor)).

[115] I assume ALIGN-R(ω, W) outranks FT-BIN, because Yamaguchi (1984) does not report accent shifts for trimoraic words when followed by a particle consisting of a single heavy syllable, such as *-nyaa* (which is a contraction of the dative marker *-ni* and the topicalizer *-wa*). If they had the same ranking, the accent would shift in order to satisfy PARSE-σ. In any case, even if such shifts would exist, PARSE-σ would still rule out the forms in (29b/c/d/), which are all unattested.

(29)

atamaʾ-demo	ALIGN-R(ω, W)	FT-BIN	PARSE-σ	ALL-FT-RT
→a. [(ata)(maʾ)](demo)		*		*
b. [(ata)(maʾ)]demo		*	*!*	*
c. [(ata)(madeʾ)]mo	*!		*	**
d. [(ata)(madeʾ)mo]	*!		*	*, ***

Next, consider the forms in (30).

(30)	Accent	Foot structure	Gloss
a.	teʾ	(teʾ)	'hand'
b.	teʾ-ga	(teʾ-ga)	'hand-NOM'
c.	te-kaʾra	(te-kaʾ)ra	'hand-ABL'
d.	te-nyaʾa	(te-nyaʾa)	'hand-DAT.TOP'

When a monomoraic word is followed by a bimoraic syllable, the constraint ALIGN-R(ω, W) is violated, and the particle is incorporated into the p-word. There are two possible explanations for this. The first possibility is that p-words need to branch (Ito 1990, Ito and Mester 1992; Ussishkin 2000). The constraint in (31) is adapted from Ussishkin (2000: 44).

(31) P-WD-BRANCHING = ω-BRANCH
A p-word branches at the foot level or the syllable level.

Because ω-BRANCH is more important than the alignment of the boundaries of a p-word and an m-word, the accent shifts in (32b) so that IAMB can be satisfied. However, when the particle is only monomoraic, the accent shift is blocked by NONFINALITY(μ), as in (32a).

(32)a. 'hand-NOM'

teʾ-ga	ω-BRANCH	ALIGN-R (ω, W)	NONFIN(μ)	IAMB
→a. [(teʾga)]		*		*
b. [(tegaʾ)]		*	*!	
b. [(teʾ)]ga	*!			

b. ‘hand-ABL’

teˀ-kara	ω-BRANCH	ALIGN-R (ω, W)	NONFIN(μ)	IAMB
→a. [(tekaˀ)ra]		*		
b. [(teˀka)ra]		*		*!
c. [(te)](kara)	*!			

Another possible explanation, which in the light of cross-dialectal typology is more interesting, is that the p-word-initial foot needs to branch. We could even assume that this is because the initial foot is the head foot, which is not such a strange idea when we realize that unaccented words must also have a head.

(33) INITIAL-FOOT-BINARITY: A p-word has an initial binary foot.

An extra argument in favor of an INITIAL-FOOT-BINARITY constraint is that it also explains why all words have initial feet in the first place. If we adopt INITIAL-FOOT-BINARITY, there is no need to refer to both ALL-FEET-RIGHT and *-o-/F to get p-word-initial bimoraic feet. What is more, it may even be the case that there is independent motivation for this constraint from the prosodic morphology of (Tokyo) Japanese, where, as discussed in Chapter 2, [(σσ)σ] parsings are preferred to [σ(σσ)] parsings. Ito and Mester (2012: footnote 34) point out that their INITIALFT requires “a proper bimoraic foot”. The Maisaka data can be seen as evidence for such an interpretation of the constraint.

The tableaux in (34) show how INITIAL-FOOT-BINARITY forces the parsing of the (first syllable of the) particle inside the p-word.

(34)a. ‘hand-NOM’

teˀ-ga	INIT-FT-BIN	ALIGN-R (ω, W)	NONFIN(μ)	IAMB
→a. [(teˀga)]		*		*
b. [(tegaˀ)]		*	*!	
b. [(teˀ)]ga	*!			

b. ‘hand-ABL’

teˀ-kara	INIT-FT-BIN	ALIGN-R (ω, W)	NONFIN(μ)	IAMB
→a. [(tekaˀ)ra]		*		
b. [(teˀka)ra]		*		*!
c. [(te)](kara)	*!			

Returning to the forms in (30), note that the final mora of *te-nyaˀa* ‘hand-DAT.TOP’ in (30d) is assumed to be part of an iambic ($\sigma_{\mu}\sigma_{\mu\mu}$) foot. However, in theory metrical parsing might ignore syllable boundaries, resulting in *(te-nyaˀ)a*. The reason for choosing the parsing *(te-nyaˀa)* is that it nicely fits with the idea that Maisaka Japanese is an iambic language, and that it obviates the need to assume violations of Syllable Integrity. While there is no conclusive evidence for such trimoraic iambs, there is evidence against trimoraic feet consisting of an initial heavy and a following light syllable ($\sigma_{\mu\mu}\sigma_{\mu}$). The absence of an accent shift from the first to the second syllable in forms like *kaˀiko-ga* (35) indicates that a disyllabic trochee with an initial heavy syllable is not allowed. The fact that *kaˀiko* is not realized as **kaiˀko* is of course that the accent may not be dominated by a dependent mora, i.e. a syllable non-head (*H/NONHEAD$_{\sigma}$ or A-TO-HEAD$_{\sigma}$).

(35)

a.	(**a**ˀ.me)	‘rain’	(a.m**e**ˀ.)-(ga)	‘rain-NOM’
b.	(kaˀi.)(ko)	‘silkworm’	(kaˀi.)(ko.)-(ga)	‘silkworm-NOM’

Finally, the existence of both shifting and non-shifting words among 5μ-compounds can be explained by the different location of the p-word boundary within the compound, as in (36). Thus, the even/odd distinction is also relevant for the second members of compounds, so that compounds of the same length but with different internal structures have the accent on different locations.

(36) a.i.	[[(dai)][(dokoˀ)ro]]	‘kitchen’
ii.	[[(dai)][(dokoˀ)ro]]ga	‘kitchen-NOM’
b.i.	[[(abu)ra][(m**u**ˀsi)]]	‘plant louse’
ii.	[[(abu)ra][(mus**i**ˀ)]]-ga	‘plant louse-NOM’

In (37) the ungrammatical forms of accented words from one to four moras and the forms into which such potential input forms are repaired are summarized.

(37)

	1 mora	2 moras	3 moras	4 moras	5 moras
1	(teˀ)	(miˀmi)	*(μσˀμσ)(μ) →(μσμˀσ)(μ)	*(μσˀμσ)(μμ) →(μσμˀσ)(μμ)	*(μσˀμσ)(μμ)μ (μσμˀσ)(μμ)μ
2		*(μμˀ) →(μˀσμσ)	(sugaˀ)ta	(muraˀ)(saki)	(migiˀ)(hida)ri
3			(ata)(maˀ)	(asa)(gaˀo)	*(μσˀμσ)(μμ) →(μσμˀσ)(μμ)
4				(otoo)(toˀ) *(μμ)(μσμσˀ) →(μμ)(μˀσμσ)	(dai)(dokoˀ)ro

The constraint ranking that is crucial for the "repairs" shown in (35) is NONFINALITY(μ) >> IAMB >> NOFLOP. Furthermore, NO repairs are made in forms like *(ata)(maˀ)*, (*otoo)(toˀ)*, and *(asa)(gaˀo)* because PARSE(H-Sponsor) outranks NONFINALITY(μ) and IAMB.

The ranking of the constraints used in the proposed analysis can be summarized as in (38) or (39), where the former gives the ranking of the analysis based on ω-BRANCH, and the latter the ranking of the analysis based on INITIAL-FOOT-BINARITY. Because for some constraints it is not clear how they are ranked relative to each other, the ranking is divided into two groups of constraints for which the relative rankings can be deduced from the tableaux introduced above.

(38) a. A-TO-HEAD, ω-BRANCH, PARSE(H-Sponsor) >> NONFIN(μ) >> IAMB >> NOFLOP
b. *-o-/F >> FT-BIN >> PARSE-σ >> ALL-FEET-RIGHT

(39) a. A-TO-HEAD, INIT-FT-BIN, PARSE(H-Sponsor)>> NONFIN(μ) >> IAMB >> NOFLOP
b. INIT-BIN-FT >> FT-BIN >> PARSE-σ

If the proposed metrical analysis is on the right track, Maisaka Japanese provides us with evidence for (i) the coexistence of right- and left-headed feet in a single system, (ii) monomoraic feet (at least accented ones), and (iii) covert feet, i.e. feet that have no (clearly audible) phonetic correlate.

Before we move on to see whether a purely tonal analysis is available, however, I will first show that the analysis proposed for nouns naturally extends to verbs.

4.2.2. Verbs

In 4.1.2, the accent patterns of accented verbs with non-past forms of two, three, and four moras were presented. In this section, it is shown that the metrical analysis can be naturally extended to verbs.

The prosodic structure proposed for the different forms are presented in (40) and (41). The verbs in (37) are V-verbs, those in (38) are C-verbs. As above, square brackets indicate p-word boundaries. The lexical accent is assumed to be floating, as in the more recent literature on verbal accent (Kitagawa 1986; Haraguchi 1991; Kubozono 2008; Nishiyama 2010; Yamaguchi 2010; Poppe 2012). The floating nature of the accent is expressed by the subscripted [acc] feature.

(40)	*mi*-[acc] 'see'	*tabe*-[acc] 'eat'	*atume*-[acc] 'gather(tr.)'
non-past	[(miˈru)]	[(tabeˈ)ru]	[(atu)(meˈru)]
transitional	[(miruˈ)to]	[(tabeˈ)ru]to	[(atu)(meruˈ)]to
conditional I	[(miryaˈa)]	[(tabeˈ)(ryaa)]	[(atu)(meryaˈa)]
past	[(miˈta)]	[(tabeˈ)ta)]	[(atu)(meˈta)]
conditional II	[(mitaˈ)ra]	[(tabeˈ)(tara)]	[(atu)(metaˈ)ra]
gerund-foc	[(miteˈ)mo]	[(tabeˈ)(temo)]	[(atu)(meteˈ)mo]

(41)	*kak*-[acc] 'write'	*aruk*-[acc] 'walk'	*atumar*-[acc] 'gather(intr.)'
non-past	[(kaˈku)]	[(aruˈ)ku]	[(atu)(maˈru)]
transitional	[(kakuˈ)]to	[(aruˈ)ku]to	[(atu)(maruˈ)]to
conditional I	[(kakyaˈa)]	[(aruˈ)(kyaa)]	[(atu)(maryaˈa)]
past	[(kaˈi)ta]	[(aruˈi)ta]	[(atu)(maˈt)ta]
conditional II	[(kaˈi)(tara)]	[(aruˈi)(tara)]	[(atu)(maˈt)(tara)]
gerund-foc	[(kaˈi)(temo)]	[(aruˈi)(temo)]	[(atu)(maˈt)(temo)]

The metrical structure in these forms is assigned according to the same principles as in nouns. The data shows that in paradigmatically related forms, a switch from a trochaic to an iambic foot is allowed as long as the head of the trochaic foot is parsed into the iambic foot. As a general "rule", the accentual H falls on the rightmost foot in the shorter (more basic) forms of the paradigm. Or, alternatively, we could say that the accentual H tone falls on the foot that contains the rightmost vowel of the stem. We will come back to the formal details of the analysis in Chapter 6, where accentual alternations in the verbal paradigms of different dialects are analyzed.

In conclusion, a metrical analysis of the accent system in the Maisaka dialect is able to account for the observed gaps in the accent system and synchronic accent shifts. The analysis is based on a default iambic foot that in some environments fails to surface due to other highly ranked constraints.

4.3. A tonal analysis of the accent patterns

In theory one could think of a number of tonal analyses. For instance, we could posit listed allomorphs: words that exhibit accent shifts have two allomorphs, one with a H tone on the penultimate syllable, and one with a H tone on the final syllable. The form in which the H tone is as far to the right as possible (ALIGN-R) without being phrase-final (NONFINALITY) would be selected. An alternative to this allomorphy-based approach is one in which the words that show accent shifts have a floating HL complex, while in words that do not show accent shifts, at least the H tone is linked in the lexicon. In this analysis too, the default accent pattern would be defined as linking the HL complex to the right edge of the p-word, docking on the final two moras.

Although analyses like this are theoretically possible, they do not explain anything. For instance, in the floating tone analysis, there is no natural connection between the number of moras and syllables on the one hand, and the existence of words with floating tones on the other. Floating tones would only exist for words with one, two, or four syllables, but not for words with three syllables. Of course, this could be an accidental gap. Still, the gaps in words of three and five moras would be left unexplained. On top of this, an even more serious problem is that an analysis based on a distinction between linked tones and floating tones does not make sense for verbs, for which accent is predictable for words of any length. The predictable nature of verbal accent, in combination with the alternations that can be observed in the verbal (and adjectival) paradigms, is the main reason why verbal accent is often analyzed as “floating” in Tokyo Japanese.

If we have another look at the gaps in the accent system (39), it is striking that four of the six gaps involve the word-initial syllable.

(42)

	1 mora	2 moras	3 moras	4 moras	5 moras
1	teˀ	miˀmi	*μσˀμσμ	*μσˀμσμμ	*μσˀμσμμμ
			→μσμˀσμ	→μσμˀσμμ	μσμˀσμμμ
2		*μμˀ	sugaˀta	muraˀsaki	migiˀhidari
		→μˀσμσ			
3			atamaˀ	asagaˀo	*μσˀμσμμ
					→μσμˀσμμ
4				otootoˀ	daidokoˀro
				μμμσμσˀ	
				→μμμˀσμσ	

Based on this observation, we may hypothesize that the HL complex is not aligned to the right edge of the word (i.e. penultimate accent), but the H tone is aligned as far to the left of the p-word, avoiding the final syllable (NONFINALITY), and if possible, the first syllable as well (NONINITIALITY; see Buckley 2013). This would also explain the lack of shifts of the accent from the second to the third syllable. What is more, the examples *asagaˀo* and *daidokoˀro* are (at least historically) compounds. While the meaning of *asa-gaˀo* 'morning glory' cannot be predicted from the morphemes of which it exists, one could still argue that the morphemes are recognized as structural units, the result of which is that the form is analyzed as a compound. If so, accent falls on the second syllable of the second p-word of the compound, the favored position. The same analysis is possible for *dai-dokoˀro* 'kitchen', so the only two gaps left would be the absence of final accent in words of two and four moras (ending in two light syllables). These cannot be explained in terms of NONFINALITY, because words of three moras and words of four moras which do not end in two light syllables, may have final accent. As there does not seem to be an alternative explanation available, we cannot but conclude that metrical foot structure is necessary to convincingly account for the accent shifts and the gaps in the nominal accent system.

Another clear advantage of the metrical analysis lies in the fact that it naturally extends to verbs. The alternative analysis sketched for nouns in which accent on the second syllable from the left is the default is not available for verbs. In verbs of four moras the accent is preferably placed on the syllable including the fourth mora, and if this is mora is phrase-final, on the third mora, e.g. non-past *atumeˀru* vs. transitional non-past *atumeruˀ* ('to collect'). In the tone-only approach, this fact cannot be accounted for.

4.4. Summary

In this chapter I have demonstrated that a metrical analysis of the Maisaka data convincingly accounts for the local accent shifts and the gaps in the accent system. Furthermore, a convincing tone-only analysis is not available. This may not come as a surprise. After all, the fact that a distinction is made between words with even and odd numbers of syllables is characteristic of metrical structure.

The proposed metrical approach to the accentual system of Maisaka Japanese only works if (i) both trochaic and iambic feet are allowed, with the latter being the default foot form; (ii) degenerate feet are allowed, and (iii) covert feet are allowed. These are important findings for both Japanese prosodic typology and prosodic typology in general. The co-existence of both trochaic and iambic feet in Tokyo Japanese is discussed by Tanaka (2001), and has been tacitly assumed in analyses of Japanese compound accent. However, alternative non-metrical analyses were not considered, and as we have seen in Chapter 3, the evidence for foot structure in Tokyo Japanese cannot be solved by looking at Tokyo Japanese in isolation. However, if we look at the Tokyo data from the point of view of Maisaka, the idea that both trochaic and iambic feet may exist in Tokyo Japanese seems more natural. Interestingly, Houghton (2013) demonstrates that "switch languages" in which trochaic and iambic feet both occur are attested in the Australian languages Yidiny and Wargamay.

Another interesting finding was that case particles may be incorporated into the p-word so that a word- and/or foot-minimality constraint can be satisfied. Thus, Maisaka Japanese resembles Tokyo Japanese in that particles are preferably parsed outside the p-word projected by the noun, but if this entails the violation of a word- or foot-minimality constraint, the particle is incorporated into the p-word.

Furthermore, the evidence for degenerate feet may help us in further investigations of Tokyo Japanese and other dialects. As pointed out in the previous chapter, researchers have not reached consensus on the status of monomoraic feet in Tokyo Japanese. The evidence for such feet in Maisaka Japanese is thus very welcome.

Finally, the evidence for covert feet is important evidence for abstract phonological representations. While covert feet have been proposed in the cross-linguistic literature, due to their abstract nature they have not been accepted by all researchers. Maisaka Japanese shows clear evidence that such abstract feet can play a role in the accent system, which is a significant finding.

The iambic nature of feet in the Maisaka dialect raises questions about metrical coherence (Lahiri 2001), i.e. the relation between the iambic feet needed for accentuation, and the trochaic feet that have been proposed for prosodic morphology. One problem here

is that not much is known about prosodic morphology in dialects. Even if Maisaka speakers make use of the trochaic feet in prosodic-morphological word formation, this may be due to the influence of the standard dialect. On the other hand, if the p-stem rather than the foot is the bimoraic unit that is observed in the templates of prosodic morphology, there is no problem in the first place.

In sum, the discovery of foot structure in Maisaka Japanese is important, but in order to elucidate the nature of metrical structure in Japanese, it is important to study a wider range of dialects. In the next chapter, a start is made by examining three additional dialects in which tone and vowel quality interact.

5. Tone-vowel interaction across Japanese dialects

In a number of Japanese dialects, tone is reported to interact with vowel height (Haraguchi 1984; Nitta 2001). In these dialects, high tones seem to avoid high vowels. According to Odden (2001), such interaction of tone with vowel height is only attested in Japanese. Interaction between stress and vowel height, on the other hand, is a relatively well-known phenomena (see e.g. de Lacy 2007). An important question is whether the interaction is really between tone and vowel height (or sonority)is direct, or rather indirect, with both tone and vowel height or sonority are sensitive to metrical structure. According to de Lacy (2007) – who does not mention the attested patterns in Japanese dialects – the interaction between tone and sonority can only be indirect, i.e. mediated by syllable or foot structure. In other words, foot structure constraints referring to sonority and foot structure constraints referring to tone are argued to be necessary, but constraints in which tone and sonority directly refer have not been proved to be necessary (de Lacy 2007).[116]

If de Lacy (2007) is right, we may expect to find evidence for metrical structure in dialects that exhibit tone-vowel interaction. The goal of this chapter is to investigate whether metrical structure is indeed involved, and to compare metrical and tonal analyses in cases in which both are available.

The three dialects that form the subject of this chapter are Matsue Japanese (spoken in Shimane Prefecture), Ichihara Japanese (spoken in Chiba Prefecture), and Kanazawa Japanese (spoken in Ishikawa Prefecture). For the first two dialects, original fieldwork data was collected in order to confirm some of the claims made in the literature.

There are a number of other Japanese dialects in which tone and vowel height are reported to interact: the Shizukuishi (Iwate) dialect (Uwano 1984) and the Tsuruoka dialect (Nitta 1994, Haraguchi 2001) spoken in the Tohoku region, the Kaga and Noto dialects, spoken in Ishikawa Prefecture (Hirayama 1952), the Toyama dialect (Sanada 1998), the Sanagishima dialect (Akinaga 1966) and the Takamatsu dialect (Wada 1958) spoken in Kagawa Prefecture, the Tanohana dialect (Shimizu 2006) spoken in Ehime Prefecture, and the Oki dialects (Hiroto & Ōhara 1952), spoken in Shimane Prefecture. In order to gain a full understanding of tone-vowel interaction in Japanese, a contrastive

[116] In more recent years, other evidence for tone-vowel quality interaction has emerged. Odden (2007) argues that in Tupuri the feature [+high] triggers 'superhigh' tones. This pattern can be seen as the phonologization of the relatively high intrinsic F_0 of high vowels. As noted by Hombert 1977, 1978 and Hombert et al. 1979, this intrinsic F_0 is in general not enough to trigger interaction between tone and vowel height, which means the Tupuri case is exceptional. What is interesting here is that Tupuri shows the opposite pattern of the one attested in Japanese dialects, in which H tones prefer non-high vowels. Mortenson (2013) presents evidence for tone-vowel quality interaction in Shuijinping Mang. Becker & Jurgec (to appear) argue that there is interaction of tone with another vowel feature, namely ATR, in Slovenian.

study of the phonetics and phonology of these dialects is necessary. However, as such a large-scale study lies beyond the scope of this thesis, in this chapter we focus on the three dialects of Matsue, Ichihara, and Kanazawa.

The Matsue and Ichihara dialects were selected because they are both known as Tokyo-type dialects, and therefore may shed light on the issue of metrical structure in the Tokyo dialect. As we will see below, Matsue and Ichihara Japanese show some striking resemblances in the way in which vowel quality influences the tonal patterns.

The Kanazawa dialect was selected because in this dialect not only vowel quality, but also consonant voicing exerts influence on the tonal patterns. Furthermore, at least on the surface this interaction seems to be of a non-local nature (Odden 2001, 2007), which makes the Kanazawa dialect an important testcase for any phonological theory.

The structure of this chapter is as follows. In 5.2, relevant data from the three dialects is introduced. In 5.3, 5.4, and 5.5, analyses of tone-vowel height interaction in the tree dialects are proposed. Following this, in section 5.6, vowel devoicing in the Tohoku dialects is argued to be a potentially interesting topic for further research in the light of the topic of this chapter. In 5.7, the major findings and remaning issues of this chapter are discussed.

Before starting the discussion of the data from the three dialects disussed in this chapter, it is useful to summarize what we know about H-tone-alignment in Japanese in general, as this will facilitate the discussion of phonetic alignment or "anchoring" (Ladd 2006) as opposed to phonological association of tones in the three dialects.

5.1. Phonetic anchoring of H tones in Japanese

It is well known that peak delay occurs in Tokyo Japanese (Neustupný 1966, Sugito 1969/1982, Ishihara 2006, among others). Sugito (1969/1982) demonstrated that in words with initial accent, even if the F_0 peak is phonetically aligned with the syllable following the accented syllable, if a fall in pitch can be detected in the immediate post-accentual syllable, the accent is perceived as belonging to the initial syllable. The most important clue to determine the location of accent is thus the falling pitch in the post-accentual syllable.

Ishihara (2006) investigated the phonetic alignment of the H tone in words with different accent patterns, and found the following tendencies.

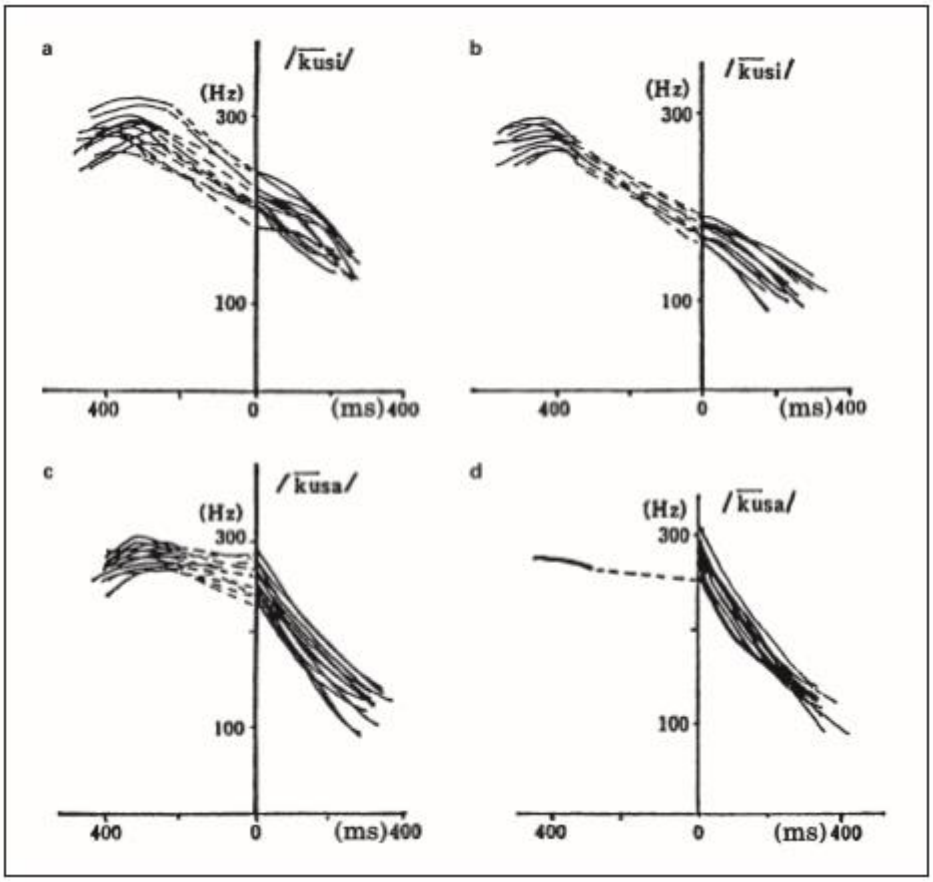
(1)

Accent	*Phonetic alignment*
Initial:	H is aligned with the beginning of the first vowel of the syllable following the accented syllable (i.e. the second syllable).
Second:	H is aligned around the end of the accented syllable.
Third/Fourth:	H is aligned around the middle of the vowel of the accented syllable.

According to Ishihara (2006), differences in vowel height have no significant influence on the anchoring of H tones with segmental material.

This conclusion stands in contrast with Sugito's (1981/1982, 2003) findings for Osaka Japanese. She demonstrates that even in the Osaka dialect, in which tone does not interact with vowel height phonologically, there is a relation between the onset of pitch change and the onset of segmental transitions. More concretely, Sugito (2003) shows that the onset of pitch transitions in initially accented words like *kuˈsi* 'comb' and *kuˈsa* 'grass' are different. The graphs in (2), taken from Sugito (2003:9), illustrate the different F_0 patterns of the two words as pronounced twelve times by two different speakers. The dotted lines are interpolations between the pitch of the first vowel and second vowels. The 0-point on the x-axis indicates the start of the vowel of the second syllable. The differences between *kuˈsi* in (2a) and (2b) on the one hand, and *kuˈsa* in (2c) and (2d) on the other hand are clear. Sugito (2003) explains that in *kuˈsi*, the pitch starts to drop "near the end of the first vowel", whereas in *kuˈsa*, it starts to drop "in the vicinity of the onset of the second vowel" (Sugito 2003: 8).

(2) F_0 contours of *ku˺si* and *ku˺sa* in Osaka Japanese (Sugito 2003: 9)



Based on complementary evidence from muscle activity, Sugito (2003) concludes that there is a preference to synchronize the onset of the pitch fall with the onset of the formant transition towards the non-high vowel.

While Sugito's findings are interesting, they are based on a limited data set, as pointed out by Ishihara (2006). Nevertheless, the results do at least suggest that the tone-vowel height interaction that can be observed in some Japanese dialects may have a phonetic basis. To what extent the phonetically based interaction between tone and vowel height has been phonologized across dialects, and the consequences of this phonologization for the organization of the phonological grammar, is an important research topic. In this chapter, a start is made by focusing on the three dialects of Matsue, Ichihara, and Kanazawa.

5.2. Data from three dialects

In this section, data of the three dialects discussed in this chapter is introduced. The first dialect to be discussed is Matsue Japanese, which has the least complicated accent system

of the three dialects.

5.2.1. Matsue Japanese

Matsue Japanese is spoken in the city of Matsue, Shimane Prefecture. Matsue Japanese is part of the larger dialect group of Izumo Japanese. The analysis and the discussion in this paper are based on two types of data. First of all, reference is made to important descriptive studies by Hiroto & Ōhara (1952) and Uwano (1981). The data in Okuda (1975) was also consulted. Data on the contemporary dialect was obtained during a fieldwork session conducted in November 2012 in Matsue City, Yokohama Town.

5.2.1.1. Data from previous studies

Let us start with the pitch patterns of nouns. The data in (3a) shows that, as in Tokyo Japanese, we are dealing with what Uwano (1981) calls an *n+1* system for words of *n* moras: a word has a 'pitch accent' on any of its moras, or is 'unaccented'. The data in (3b/c/d) involves deviations from the more regular patterns due to syllable structure (3b) or the feature content of particular segments (3c/d). The tonal patterns given are nothing else than approximations of the phonetic patterns. As we will see below, at least the unaccented forms show tonal underspecification even on the surface. To distinguish between accentual H tones and phrasal H tones, accentual H tones are accompanied by the accent mark, i.e. Hˀ. As in Tokyo Japanese, pitch accent can be analysed as a HL tonal complex, which in this chapter is represented as HˀL.[117]

(3) Accent/tone patterns of Matsue Japanese nouns (based on Hiroto & Ōhara 1952)

	Noun		*Noun-NOM*		*Noun-ABL*	
a.	kaˀbuto	HˀLL	kaˀbuto -ga	HˀLLL	kaˀbuto-kara	HˀLLLL
	inoˀti	LHˀL	inoˀti-ga	LHˀLL	inoˀti-kara	LHˀLLL
	atamaˀ	LHHˀ	atamaˀ-ga	LHHˀL	atamaˀ-kara	LHHˀLL
	sakana	LHH	sakana-ga	LHHH	sakana-kara	LHHHH
	(*kabuto* = 'beetle'; *inoti = 'life'; atama* 'head' ; *sakana* ='fish')					
b.	koori	HHH	koori-ga	HHHH	koori-kara	HHHHH
c.	kagamiˀ	LHHˀ	kagami-gaˀ	LHHHˀ	kagami-kaˀra	LHHHˀL
d.	kuziraˀ	LLHˀ	kuziraˀ-ga	LLHˀL	kuziraˀ-kara	LLHˀLL
	(*koori* ='ice'; *kagami* = 'mirror'; *kuzira* ='whale')					

[117] Uwano (1981) posits a "lowering kernel" for Matsue Japanese, which is the same type of accent kernel that he proposes for Tokyo Japanese (Uwano 1999b, 2012b).

The data in (3b) shows tone is weight-sensitive. This is attested in many other dialects, among which Tokyo Japanese, and can be accounted for in terms of the tone-weight alignment constraints introduced in Chapters 2 and 3. Although the problem of weight-sensitive tone in Matsue Japanese deserves more attention that it is given here,[118] weight-sensitive tone is not further analysed, because the focus in this chapter is on the interaction between tone and vowel height.

As can be seen in (3c), the lexical HˇL shifts to the following nucleus and thereby avoids the association of the Hˇ part with the noun-final high vowel: *kagami**i***ˇ ~ *kagami-g**a***ˇ. When Hˇ is associated to the final syllable, the L part is left unassociated. The crucial point here is that in the isolation form *kagami*ˇ, the lexical Hˇ tone does not shift to the left (**kagaˇmi*), which means the Hˇ tone is associated to a high vowel even though there is an adjacent non-high vowel to the left. This suggests that Hˇ tone shifts are restricted in terms of direction, something we will need to account for.

In (3d), another instance of the interaction between tone and vowel height can be observed. The form *kuzira*ˇ is realized with a L tone not only on the initial syllable, but also on the second one.

Judging from the forms in (3), a possible generalization seems to be that H tones avoid a high vowel that is followed by a non-high vowel in the syllable to its immediate right. This generalization is informally stated in (4a), and schematically in (4b) (see Uwano 1981 and Nitta 2001 for similar generalizations stated in terms of pitch movement).

(4) a. H tones avoid associations with syllables headed by high vowels that are directly followed by a syllable with a non-high vowel.

b. *CI^{H}CA (where C = consonant, I = high vowel, and A = non-high vowel)

However, things get more complicated if we also take into account constructions in which nouns are followed by particles with a high vowel in the initial mora, such as the dative marker *-ni*. According to Okuda (1975: 42), in forms like *kagamiˇ-ni* 'mirror-DAT', the accent does not shift. This would make sense according to the generalization in (4a). However, Hiroto & Ōhara's (1952) data, on the other hand, seem to suggest vowel shifts are also obligatory when particles with high vowels are involved.[119] It may be the case

[118] Matsue Japanese is particularly interesting in terms of tone-syllable weight interaction, because a distinction is made between derived and non-derived heavy syllables. Words starting with non-derived heavy syllables of the shape CVV or CVN receive an initial H tone (e.g. *koori*^{0} HHH 'ice'), but words starting with a "derived" heavy syllable, i.e. *kaada*^{0} derived from *karada*^{0} 'body' by the deletion of the onset of the second syllable, are assigned the tonal pattern LLH.

[119] Hirotu & Ōhara (1952) mention that the particle *-ni* is often deleted in everyday speech. Therefore, their examples include few cases of nouns with an accent on a final high vowel followed by the particle

that the system described by Okuda (1975) existed at some stage in Matsue Japanese. However, because Okuda's (1975) data is based on secondary sources, here I will follow Hiroto & Ōhara (1952), who collected the data themselves.[120]

It is important to distinguish between accentual tones and non-accentual boundary tones. As pointed out by Uwano (1981), the function of the initial LH-sequence that also surfaces in unaccented forms like *karada*^{0} [LHH] is to mark the left boundary of a phrase. Thus, the initial LH tones can be assumed to be default boundary tones rather than lexical tones, just like in Tokyo Japanese. Evidence for this comes from the realization of phrases consisting of multiple phrases, such as those in (5) (see Uwano 1981).

(5)	karada^{0}	LHH	kono^{0} karada^{0}	LH HHH	'(this) body'
	kagami˺	LHH	kono^{0} kagami˺	LH HHH	'(this) mirror'
	kuzira˺	LLH	kono^{0} kuzira˺	LH HHH	'(this) whale'

As can be seen in (6), the interaction between tone and vowel height can be observed in words of all lengths, except for monomoraic words.[121] Numbers denoting accent location (in the "Accent" column) refer to the *n*th mora counted from the left. The different phrasal tone patterns that are assigned when the words are followed by the nominal particle *-ga* are indicated by "a" (LH..) and "b" (LLH..). Words with final accent that are subject to accent shift are indicated with an apostrophe. Thus, the class *2b'* in (6) stands for a word with an accent on the second syllable from the left, with a LLH phrasal tone pattern, and shifting accent in the nominative construction.

Notice that among words of two moras or more, there are no words in which the accent falls on a high vowel that is followed by a non-high vowel. Thus, the *CI^{H}CA condition is satisfied in nouns across the lexicon.

-ni. One example that I found is *hikari˺* 'light' vs. *hikari-ni˺* 'light-NOM', the latter of which is described as being accented on the particle.

[120] The data I collected myself does not show consistent accent shifts, and therefore does not help us with this issue. It should be pointed out that not all particles pull the accent to their initial syllables. For instance, the particle *-bakari* 'only' is pre-accenting, and thus consistently places a H tone on the final TBU of the noun, even when the final vowel is a high vowel.

[121] Thus, the monomoraic word *yu˺* 'hot water' in (4) is realized with a high tone even when followed by the particle *-ga*. Assuming that this is an instance of positional faithfulness (i.e. faithfulness to strong positions, in this case the word-initial mora/syllable; see Beckman 1998), monomoraic words will not be discussed in this chapter.

(6) Table of tone patterns in Matsue Japanese (based on Hiroto & Ōhara 1952)

a = Φ[LH..]; a'= Φ[LH..] + accent shift; b= Φ[LL..]; b'= Φ[LL..] + accent shift;

Accent	1 moras	2 moras	3 moras	4 moras
0	e^0-ga	$hana^0$-ga	$karada^0$-ga	$niwatori^0$-ga
	LH	LHH	LHHH	LHHHH
	'handle'	'nose'	'body'	'chicken'
0b		usi^0-ga	$hazime^0$-ga	$karu\text{-}waza^0$-ga
		LLH	LLHH	LLHHH
		'cow'	'beginning'	'acrobatics'
1	yuˀ-ga	aˀki-ga	kaˀbuto-ga	aˀisatu-ga
	HL	HLL	HLLL	HLLLL
	'hot water'	'autumn'	'helmet'	'greeting'
2a		udeˀ-ga	inoˀti-ga	uguˀisu-ga
		LHL	LHLL	LHLLL
		'arm'		'bush warbler'
2b'		asi-gaˀ		
		LLH(L)		
		'leg'		
3a			atamaˀ-ga	honeguˀmi-ga
			LHHL	LHHLL
			'head'	'carcass'
3a'			kagamiˀ-ga	
			LHHL	
			'mirror'	
3b			kuziraˀ-ga	mugimaˀki-ga
			LLHL	LLHLL
			'whale'	'wheat planting'
4a			kagamiˀ-ga	simoyakeˀ-ga
			LHHL	LHHHL
				'chillblains'
4a'				usotukiˀ-ga
				LHHHH
				'liar(-NOM)'
4b				usotukiˀ-ga
				LHHHH
				'liar(-NOM)'
4b'				yokubariˀ-ga
				LLHHH
				'greed'

Next we turn to the behavior of tones in verbs and adjectives. Although Matsue Japanese verbal accent is treated in more detail in Chapter 6, it is useful to introduce the basic patterns here. Verbs and adjectives in all cases are morphologically complex, and therefore provide the opportunity to check whether the sensitivity of tone to vowel height is synchronically active. The examples in (7) are taken from Hiroto & Ōhara 1952. The

data shows that, like in nouns, the association of lexical tones is sensitive to vowel height: the Hˀ part of the HˀL pitch accent avoids association with a high vowel whenever possible: *nobiˀ-ru* [LHˀL] ~ *nobi-taˀ* [LLHˀ] (cf. *sodateˀ-ru* [LHHˀL] ~ *sodateˀ-ta* [LHHˀL]). The different tone patterns of *kari-ru* [LHH] and *kari-ta* [LLH] show that non-accentual H tones also avoid high vowels.

(7) Pitch patterns of Matsue Japanese verbs

	Non-past		*Past*		*Gloss*
a.	kari-ru^{0}	LHH	kari-ta^{0}	LLH	‘loan’
	nigi-ru^{0}	LHH	nigiQ-ta^{0}	LL(L)H	‘clasp’
b.	nobiˀ-ru	LHˀL	nobi-taˀ	LLHˀ	‘stretch’
	kabuˀ-ru	LHˀL	kabuQ-taˀ	LL(L)Hˀ	‘put on (a hat)’
d.	kazureˀ-ru	LLHˀL	kazureˀ-ta	LLHˀL	‘count’
e.	sodateˀ-ru	LHHˀL	sodateˀ-ta	LHHˀL	‘raise’

While pairs like *nobiˀru ~ nobitaˀ* (‘stretch-NONPAST~PASTˀ) indicate that tone and vowel height interact synchronically, the matter is complicated by what in derivational analyses is known as cyclicity (Cole 1995). Some derived forms preserve or copy the accent of their bases, which results in opaque structures in which a H tone surfaces linked to a high vowel even though a non-high vowel is available in the following vowel. This phenomenon and other alternations in the verbal paradigm are dealt with in more detail in Chapter 6.

Finally, the data in (8) shows the behavior of adjectives, which also display the interaction between tone and sonority.

(8) Pitch patterns of Matsue Japanese adjectives

	Non-past		*Continuative*		*Gloss*
a.	akaru-i^{0}	LHHH	akaru-ku^{0}	LHHH	‘bright’
b.	surudoˀ-i	LLHˀL	surudoˀ-ku	LLHˀL	‘sharp’
c.	uresiˀ-i	LHHˀL	uresiˀ-ku	LHHˀL	‘happy’
	sabisiˀ-i	LHHˀL	sabisiˀ-ku	LHHˀL	‘lonely’
d.	tiisaˀ-i	HHHˀL	tiisaˀ-ku	HHHˀL	‘small’

Summarizing, data from nouns, verbs, and adjectives shows that H tones avoid syllables headed by high vowels that are followed by syllables headed by non-high vowels, and, in the case of nouns, word-final high vowels followed by any other vowel.

Before we move on, some remarks on the phonetic realization of the high vowels are in order. As noted by Hiroto (1986), the weakness of Matsue Japanese vowels is probably related to their centralized realization. However, this cannot be the whole story, because H tones can be freely associated to high vowels that are followed by another high vowel. Because there are no non-centralized high vowels in Matsue Japanese, [ɨ] and [ɯ̈] can be analyzed as the phonetic realizations of /i/ and /u/. Therefore, from now on I will refer to the interaction of tone and vowel quality as involving tone and vowel height ([+high] and [-high]).

5.2.1.2. Fieldwork data

During fieldwork conducted in November 2012, data from Matsue Japanese was recorded. The primary goal of the fieldwork was to verify to what extent the patterns described in the literature can still be observed, and to enable examination of the actual phonetic data.[122] The two informants from Yokohama Town, Matsue City were both males in their seventies.

The data obtained from the fieldwork conducted in 2012 differs in minor yet significant ways from the data in Hiroto & Ōhara (1952) and Uwano (1981). As we will see below, a number of important changes have occurred in the behavior of the pitch accents. We will start with the boundary tones, however, the behavior of which does not seem to have changed since 1952.[123]

The pitch track in (9a) suggests that the [LLH…] pattern is the result of spreading of the L boundary tone rather than interpolation from L in the first syllable to H in the third syllable. Because the H tone does appear in the second syllable when both the second and third syllables contain a high vowel, as in *kusuri*^{0} ‘drug’ (9b; note that the vowel in the first syllable is devoiced or not realized at all), we are not simply dealing with peak delay.[124] If we were dealing with peak delay, we would expect the form in (9b) to be

[122] The data was recorded with a Sony Linear PCM Recorder (PCM-M10) and a Sony ECM-MS 957 microphone in a silent room. The recordings were digitized at a 44.1 Khz sampling rate and stored in WAV format.

[123] One of the two informants of Matsue Japanese did show different patterns in terms of the domain of the boundary tones, as he consistently separated every word(-particle combination) as if it were a separate phrase. This means that the tonal domain for this informant may be the p-word rather than the next higher level of the p-phrase. Still, the different phrasing may also have had to do with the fact that the speaker was aware of being recorded, and thereby pronouncing the words in a phrase carefully as p-words.

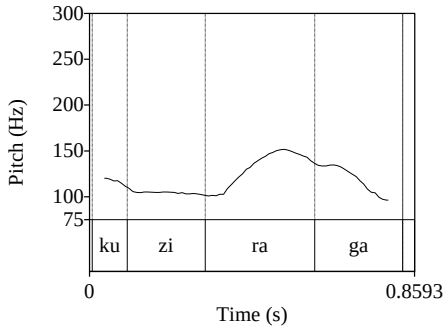
[124] As shown by Myers (2000), some accounts of tonal displacement are the result of peak delay, while others are not. The shift in Matsue Japanese must be phonological because high vowels followed by other high vowels do bear H tones, and in certain verbal forms even a syllable with a high vowel followed by one with a non-high vowel may bear a H tone (see Chapter 6).

realized with the [LLH..] pattern similar to that in (9a) rather than the [LH..] pattern of (9c). Note that the fact that *kuzira*ˀ is a word with final-accent is not relevant here, as words without high vowels in their second syllables such as *atama*ˀ are pronounced as [LHH].

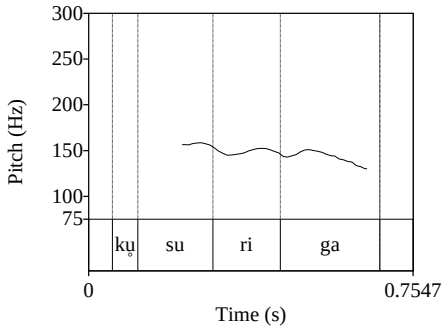
A comparison of *sakana* (9c) and *sakana-ga* (9d) suggests that the phrasal H tone does not spread, but interpolates with a final L% boundary tone. In this sense, the behavior of the phrasal H in Matsue Japanese is similar to that of Tokyo Japanese as reported in Pierrehumbert & Beckman (1988). Whether spreading of the boundary H tone is completely absent in this dialect, is not clear, but an analysis based on interpolation in which some TBUs are underspecified for tone seems to be in line with the data.

(9) Pitch tracks (obtained through Praat; Boersma & Weenink 2011)

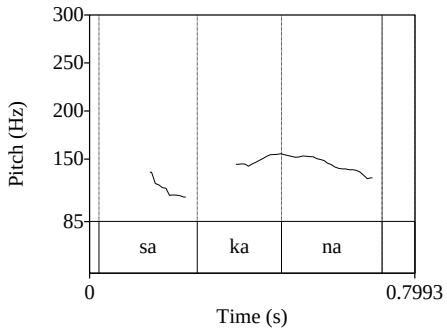
a. *kuzira*ˀ*-ga* LLHˀL (‘whale-NOM’)



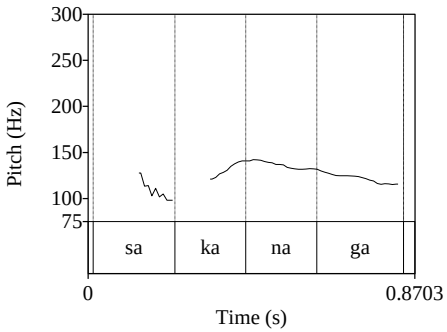
b. *kusuri*^{0}*-ga* LHHH (‘drug-NOM’)



c. *sakana*^{0} LHH (‘fish’)



d. *sakana*^{0}*-ga* LHHH (‘fish-NOM’)



In terms of underlying pitch accent specifications, the data in Hiroto & Ōhara (1952) and Uwano (1981) generally corresponds with the patterns found in the 2012 data, apart from a class of words that are important in the light of this study: final-accented words. In my fieldwork data of Matsue Japanese, the pitch accent shifts in noun+particle constructions

could not be observed. The words are either pronounced with the pitch accent staying in place in all constructions, or seem to have merged with the unaccented class. In other words, the shift of lexical pitch accents from nouns to particles in contemporary Matsue Japanese seems to be disappearing.[125]

5.2.2. Ichihara Japanese

Ichihara Japanese is part of the larger dialect group known as Boso Japanese, named after the Boso Peninsula that lies to the east of Tokyo Bay. The accent system of the Ichihara dialect is very similar to that of its Tokyo counterpart. A striking difference, however, is that in the Ichihara dialect, tone interacts with vowel height.

5.2.2.1. Data from previous studies

Boso Japanese is a relatively understudied variety of Japanese. Apart from a large-scale study on diachronic change by Hirayama et al. (1974), few detailed studies exist. The data discussed below is mainly taken from Fujiwara (1972, 1979), a native speaker of the dialect, and original fieldwork conducted in Ichihara City in November 2013. Some data can also be found in Kinda'ichi (1942), who reports patterns that are similar to the ones described in Fujiwara (1972, 1979).

The data from Ichihara Japanese in Fujiwara (1972, 1979) suggests that the dialect basically is a more complicated version of Matsue Japanese. As in Matsue Japanese, there is a difference in phrasal tone assignment ([LH..] vs. [LLH..]), depending on the vowel height of the nuclei of the second and third syllables.

(10) Phrasal tone assignment

a.	$CV_1.CA_2(.CV_3)$	→	LH..	e.g.	hatake^0	LHH	'field'
b.	$CV_1.CI_2.CI_3$	→	LH..	e.g.	nezumi^0	LHH	'mouse'
c.	$CV_1.CI_2.CA_3$	→	LLH..	e.g.	tobira^0	LLH	'door'

What is more, like in Matsue Japanese, noun-final accentual H tone shifts to the first syllable of the following particle when the syllable bearing the accent in the isolation form is headed by a high vowel.

[125] The same observation was made in Uwano (1981) for the newer variety of Matsue Japanese. Also, the same holds for the other two varieties of Izumo Japanese (Hikawa and Hirata) for which I collected data.

(11) Shifts of noun-final accentual H tones

a.	inuˀ	LH	~	inu-gaˀ	LLH(L)	‘dog(-NOM)’
b.	nioiˀ	LHH	~	nioi-gaˀ	LHHH(L)	‘smell(-NOM)’

The Ichihara dialect differs from the Matsue dialect in that the accent does not shift to a particle starting with an initial high vowel such as *-ni*, i.e. *inuˀ-ni* ‘dog-DAT’.

A characteristic that seems to be unique to the Ichihara dialect (and other varieties of Boso Japanese) is that accentual H tones in general, including those lexically associated with non-high vowels, are also realized in the syllable following the accented syllable. This is especially clear in words with initial-accent, because the initial two syllables are pronounced on a high pitch (Kinda’ichi 1942; Fujiwara 1972, 1979). The condition for the late realization of H tone is that the syllable following the accented syllable (i) contains a non-high vowel, or (ii) contains a high vowel and is followed by another high vowel (10c-ii). According to Fujiwara (1972), the realization of H on the second syllable is optional in the latter case.

(12)

	Example	Accent/Tone Pattern	
a.	aˀme-ga	HˀHL	‘rain-NOM’
b.	aˀsahi-ga	HˀHLL	‘morning sun-NOM’
c.i.	aˀki-ga	HˀLL	‘autumn-NOM’
ii.	aˀki-ni	HˀLL~HˀHL	‘autumn-DAT’

Now, Fujiwara (1972) actually analyzes words with a H on the second syllable (e.g. *ame-ga* HHˀL and *asahi* HHˀL) as accented on the second syllable. However, based on the fact that *aˀki-ni* can be realized as either HLL or HHL, Kawakami (1984) correctly suggests that analyzing these words as having initial accent is more likely.[126] Therefore, in this thesis, I assume that the underlying accent locations of Ichihara Japanese are as in (12).[127]

The data in Fujiwara (1972, 1979) suggests that H tones also spread one mora/syllable to the right when they are word-medial. The examples in (13) can be found in Fujiwara (1972, 1979). Thus, I do not follow Fujiwara (1972, 1979) in analyzing the forms in (13) as being accented on the rightmost H-toned syllable.

[126] Although Kawakami (1984) did not make any phonetic measurements, he proposes that the realization of the H tone on the second syllable is a phonetic effect rather than a phonological process. See 5.2.2.2 for discussion.

[127] My Ichihara informant told me that in forms like *aˀme*, the first syllable is “stronger” than the second, which is in line with an analysis in which the firt syllable is accented.

(13)	akaˀ-gai	LHˀHL	‘blood clam’
	temaˀ-tin	LHˀHL	‘wages for labour’
	amaˀ-mizu	LHˀHL	‘rain water’
	ao-zoˀra	LHHˀH	‘blue sky’

The accent and tonal patterns of Ichihara Japanese are summarized in (14). All words are followed by the nominative particle *-ga*. The tonal patterns on based on Fujiwara (1972, 1979), but the analysis of some of accent locations are based on the above discussion, as well as considerations discussed below.

(14) Table of accent/tone patterns in Ichihara Japanese

a = Φ[LH..]; b= Φ[LL..]; b’= Φ[LL..] + accent shift; c=Φ[Hˀ..] + H spreading

	1 mora	2 moras	3 moras	4 moras
0a	ha^0-ga	sake^0-ga	nezumi^0-ga	nokogiri^0-ga
	LH	LHH	LHHH	LHHHH
	‘leaf’	‘alcohol’	‘mouse’	‘saw’
0b		azi^0-ga	tobira^0-ga	yakusoku^0-ga
		LLH	LLHH	LLHHH
		‘taste’	‘door’	‘promise’
1a	haˀ-ga	aˀkᵢ̥-ga	kaˀbuto	aˀisatu-ga
	HˀL	HˀLL	HˀLL	HˀLLLL
	‘tooth’	‘autumn-NOM’	‘helmet-NOM’	‘greeting’
1c	hiˀ-ga	aˀme-ga	aˀsahi-ga	akacan-ga
	HHˀ(L)	HHˀL	HHˀLL	HHˀLLL
	‘fire’	‘rain’	‘morning sun’	‘baby’
2a		kawaˀ-ga	komuˀgi-ga	agohige-ga
		LHˀL	LHˀLL	LHˀLLL
		‘river’	‘wheat-’	‘goatee’
2b		inuˀ-ga		
		LLHˀ(L)		
		‘dog’		
3a			nioiˀ-ga	tamaneˀgi-ga
			LHHHˀ(L)	LHHˀLL
			‘smell-’	‘onion’
3b			musumeˀ-ga	sibugaˀki-ga
			LLHL	LLHLL
			‘daughter’	‘sour persimmon’
4a				yamaimoˀ-ga
				LHHHˀL
				‘yam’
4b				orimono-ga
				LLHHL
				‘cloth’

The behavior of monomoraic words with a high vowel is interesting. The H tone does not shift, but is delayed until or spreads to the following particle. According to Fujiwara (1972, 1979), this also happens when a particle with a high vowel like the dative marker *-ni* follows. In monomoraic words with non-high vowels, on the other hand, no such delay or spreading is observed.

(15)	a.	hiˀ-ga	HHˀ(L)
		hiˀ-ni	HHˀ(L)
	b.	haˀ-ga	HˀL

Although words starting with a heavy syllable are not given in (14), it is interesting to note that heavy syllables in Ichihara Japanese behave differently from those in Tokyo and Matsue Japanese. In Ichihara Japanese, CVV syllables behave as heavy for phrasal tone assignment, resulting in the initial-H pattern [HH..]. CVN syllables, on the other hand, do not count as heavy, and are assigned the [LLH..] pattern.[128] This forms extra evidence for tone-weight alignment constraints that refer to different types of heavy syllables.

5.2.2.2. Fieldwork data

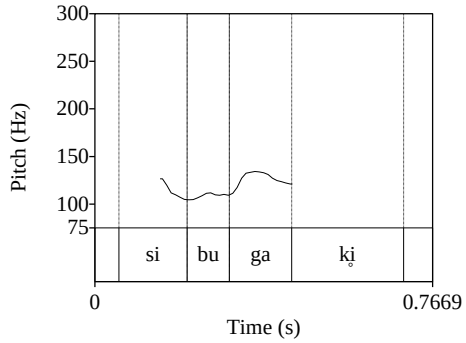
In November 2013, data from the Ichihara dialect was collected in the Yōrō-keikoku Valley. The informant was a male speaker who was 80 years old. The main goal of the fieldwork was to verify the claims made in the scarce literature on the dialect.[129] As I will show now, the collected data was fairly consistent with the claims made in Fujiwara (1972, 1979).

The difference between the phrase-initial [LH.. and [LL.. patterns were as described above. The [LL pattern can be observed in the pitch track in (16). The sharp rise from the second to the third syllable suggests the initial phrasal L tone is associated not only to the first, but also the second syllable.

[128] Heavy syllables with a diphthong do not always show the same behavior. In some cases, they are treated as two separate syllables, as in *aida* LLH 'interval', whereas in other cases, they are treated as a single heavy syllable, as in *tuitati* HHHH 'the first day (of a month)'.

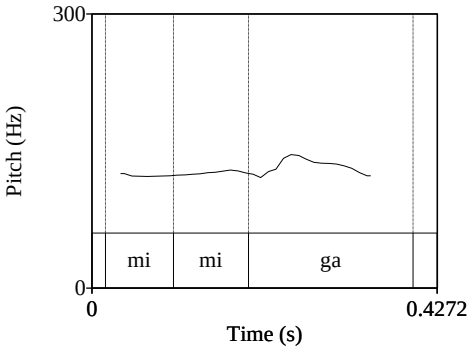
[129] The data was recorded with a Sony Linear PCM Recorder (PCM-M10) and a Sony ECM-MS 957 microphone in a silent room. The recordings were digitized at a 44.1 Khz sampling rate and stored in WAV format.

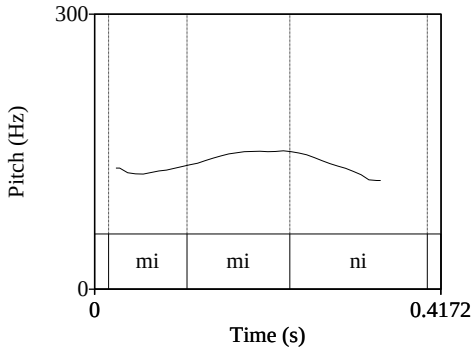
(16) *sibuga˺ki* LLH˥L (‘sour persimmon’)



Furthermore, at least for bimoraic words, the shift of the accent from a word-final high vowel to a following particle starting with a non-high vowel was also confimed. The pitch track in (17a) exemplifies the tone shift and again the phrase-initial double L. The pitch track in (17b) shows the accent remains in place when the noun if followed by the particle *-ni*.

(17) a. *mimi-ga˺* LLH˥ (‘ear-NOM’) b. *mimi˺-ni* LH˥L (‘ear-DAT’)

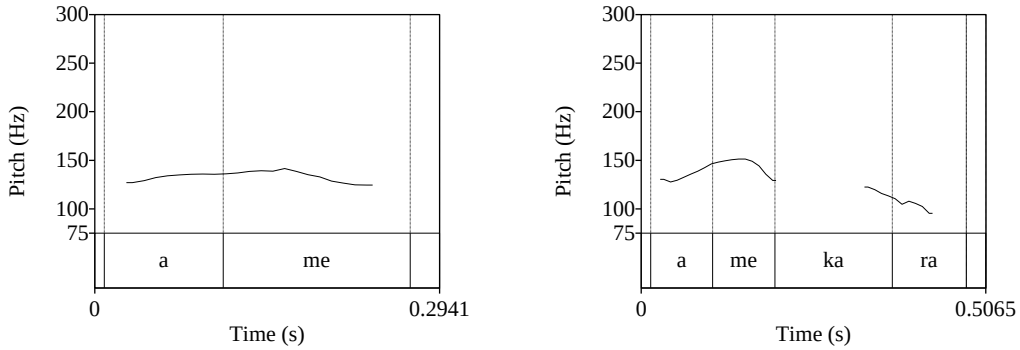




Based on these informal observations, we may conclude that the phrase-initial [LL.. pattern and the accent shift have a phonological status.

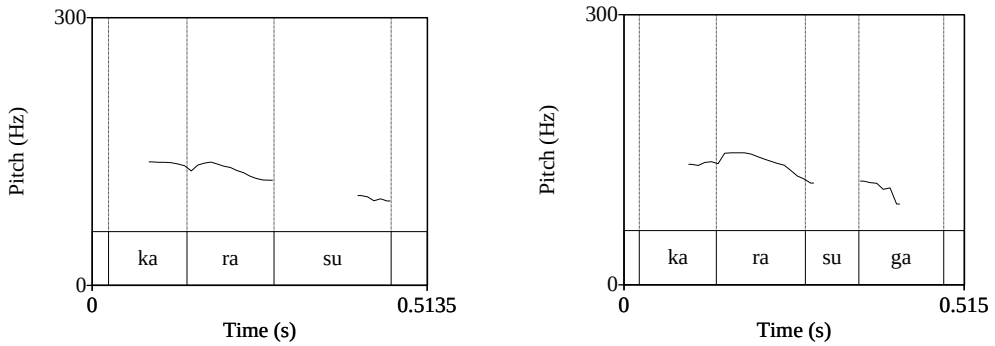
The status of peak delay of H tones that are not immediately preceded by L tones (i.e. the “H-spreading pattern”) is less clear. As can be seen from (18), the H tone is phonetically aligned to the second syllable. The example on the right even shows a pitch boost on the second syllable before the following L. This pitch boost is absent in the isolation form on the left. One possible explanation for this is that in the isolation form the H tone is really spread to the second mora, and there is no mora left for the trailing L tone to associate with.

(18) a. *aˈme* ‘rain’ b. *aˈme-kara* ‘crow-ABL’



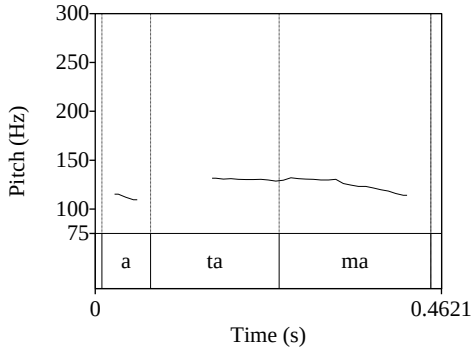
In longer words, the delay of the accentual H tone can also be observed. The example in (19) shows that the H tone is delayed until at least the start of the second syllable. However, the pitch begins to fall within the second syllable, which means that we cannot rule out the possibility that the peak delay is phonetic.

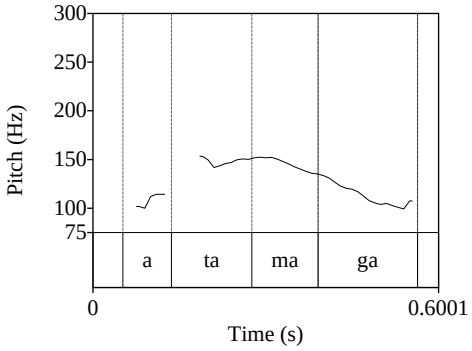
(19) a. *kaˈrasu* ‘crow’ b. *kaˈrasu-ga* ‘crow-NOM’



On the other hand, the pitch track in (20) shows that in words with a final pitch accent, for which no peak delay is reported in Fujiwara (1972, 1979), the peak can also be found relatively early in the accented syllable. This means that H tone association need not to imply phonetic alignment with the right edge of the accented syllable. However, the difference in alignment may be (at least partly) caused by the fact that the accent is associated with the initial syllable in (19) and with the third syllable in (20). That is, the Ichihara Japanese patterns in (19) and (20) resemble those of Tokyo Japanese as reported by Ishihara (2006) (see (1) above). Therefore, judging from the pitch tracks, the late realization of the accentual H in (19) cannot be taken as evidence for the phonological association of this tone to not only the first, but also the second syllable.

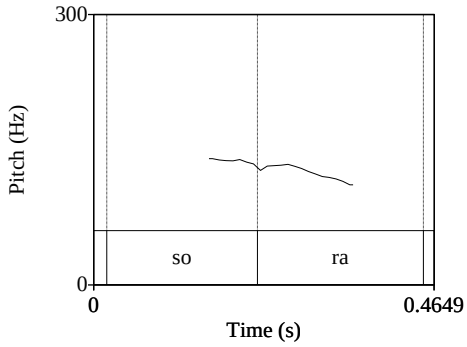
(20) a. *atamaˀ* ‘head’ b. *atamaˀ-ga* ‘head-NOM’

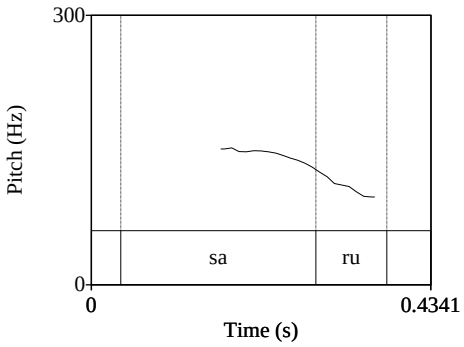




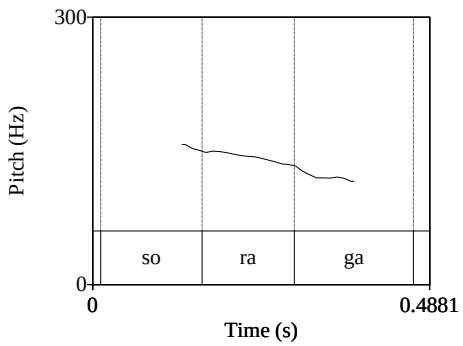
As for vowel height as a condition on tone spreading, there was no clear difference in the “spreading” of the H tone to the second syllable in initially accented words with non-high vowels in the second syllable as opposed to initially accented words with high vowels in the second syllable. The forms in (21) provide a good example. Although the forms *soˀra* ‘sky’and *saˀru* ‘monkey’ in (21) look different (HH/HF vs. HL), the difference between *soˀra-ga* and *saˀru-ga* in (22) is less clear.

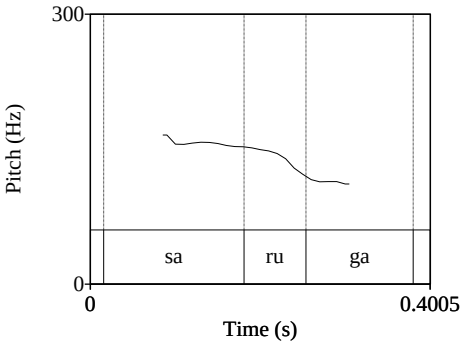
(21) a. *so’ra* ‘sky’ b. *sa’ru* ‘monkey’





(22) a. *soˀra-ga* ‘sky-NOM’ b. *saˀru-ga* ‘monkey-NOM’





Without a more detailed study including measurements based on a sufficiently large set of data, it is hard to decide whether the peak delay of the H tone in initially accented words is phonetic or phonological in nature. The peak delay might be phonetic and only slightly more advanced than in Tokyo Japanese, without changes in the association of H tones. On the other hand, if we are dealing with spreading, spreading of the initial accentual H tone to the second syllable might now (optionally) be allowed in environments in which this was not allowed before. Another possibility is that the Ichihara dialect is becoming more similar to the Tokyo dialect, and that H tone spreading used to tbe phonological but now simply is phonetic peak delay.

Because the environments in which peak delay of "spreading H tones" are claimed to occur by Fujiwara (1972, 1979) are similar to the environments that are relevant to the assignment of phrasal tones and to accent shifts, one is tempted to analyze the H-spread as phonological. Such an approach also makes sense if we take into account that Fujiwara (1972, 1979), a native speaker of the dialect, explicitly argues in favor of the accentual H tone being realized one syllable to the right compared to Tokyo Japanese (which he calls the accented syllable, an assumption that is not adopted in this thesis). Therefore, I tentatively assume that all instances of tone-vowel height interaction in Ichihara Japanese are phonological.

5.2.3. Kanazawa Japanese

Kanazawa Japanese is a dialect spoken in Kanazawa City, Ishikawa Prefecture. In this dialect, not only vowel height but also consonantal voicing plays a role in the tonal grammar. As I have not been able to collect fieldwork data on the Kanazawa dialect myself, for Kanazawa Japanese I completely rely on the data in the work by Uwano and Nitta (1982, 1983), Nitta (1989, 2001), and Uwano (1996, 1997a, 2009). I would like to stress that the proposed analyses will be based on my own interpretations of the data that they present.

Let us start again with the phrasal tones. In Kanazawa Japanese, there is a difference in phrasal tone assignment [LH..] or [H..], depending on the vowel height of the nuclei of the second and third syllables, and the voicing of the onset of the second syllable. This is shown in (23), where "K" stands for a voiceless consonant, "G" for a phonetically voiced consonant (or rather a non-voiceless consonant), and "C" for any consonant. The subscript numbers indicate the *n*th syllable of the word counted from the left edge. The words in (23a~f) can be found in Nitta (2001); the words in (23g/h) can be found in Uwano (2009).

(23)		Condition	Example	Pitch pattern	Gloss
	a.	..KV_2(..)	kata	LH	‘shoulder’
			aki	LH	‘autumn’
	b.	..GI_2 #	umi	HH	‘sea’’
			mizu-gi	HHH	‘swimsuit’
	c.	..$GI_2.CA_3$	suzume	HHH	‘mouse’
	d.	..$GI_2.CI_3$	nezumi	LHH	‘sparrow’
	e.	CV_1V_1	koori	HHH	‘ice’
	f.	CV_1N_1	oNna	HHH	‘woman’
	g.	CV_1Q_1	koQpu	LLH	‘cup’
	h.	$CV_1C\underset{\circ}{V}_2$	asi̥ta	LLH	‘tomorrow’

Like in Tokyo Japanese and Matsue Japanese, words starting with CVV or CVN (23e/f) are assigned an initial H tone phrase-initially, and in words starting with CVQ or CVC$\underset{\circ}{V}$ (23g/h) the H tone only shows up on the third syllable. What is essential about the data in (23) is the difference between (23a) and (23d) as opposed to (23b) and (23c). When the second syllable contains a voiced onset and a high vowel, and the third syllable contains a non-high vowel (23c), the H tone associates to the first syllable. The same happens when a second syllable with a voiced onset and a high vowel is followed by a morpheme boundary (23b). In all other cases, i.e. (23a) and (23d), the default LH pattern is associated to the first two syllables. Thus, these examples show that the tone that associates to the first syllable is related to the segmental make up of the second and third syllables. This means that, counting in either moras or syllables, the tonal assignment is non-local (Odden 2001). That is, a feature is associated based on the structure of a non-adjacent element. Importantly, the different patterns are in a relation of complementary distribution (Nitta 2001).

The phrasal nature of the tones is betrayed by the examples in (24), which are adapted from Nitta (2001: 165).

(24)				
a.	hune	LH	‘ship’
	dare-no hune	LHH HH	‘who-GEN ship’
b.	hari	HH	‘needle’
	dare-no hari	LHH HH	‘who-GEN needle’
c.	nezumi	LHH	‘mouse’
	soko-no nezumi	LHH HHH	‘there-GEN mouse’

d.	suzume	HHH	‘sparrow’
	soko-no suzume	LHH HHH	‘there-GEN sparrow’

The accentual system of Kanazawa Japanese is such that the conditions in (23) are satisfied by the lexically specified accentual H tones, except in words ending in GI with final accent, such as *tori*ˀ ‘bird’, which violates the condition in (23b). Notice that in words with final accent, the accent is realized as a falling tone (F).[130]

(25) Table of accent/tone patterns in Kanazawa Japanese (adapted from Nitta 2001).

	1 mora	2 moras	3 moras	4 moras
0a		hune	nezumi	itazura
		LH	LHH	LHHH
		‘ship’	‘mouse’	‘mischief’
0b	te	umi	suzume	harigane
	H	HH	HHH	HHHH
	‘hand’	‘sea’	‘sparrow’	‘wire’
1	keˀ	kuˀgi	kˀuruma	aˀisatu
	F	HL	HLL	HLLL
	‘hair’	‘nail’	‘car’	‘greeting’
2a		sakeˀ	ataˀma	asaˀgao
		LF	LHL	LHLL
		‘alcohol’	‘head’	‘morning glory’
2b		toriˀ	x	x
		HF		
		‘bird’		
3a			sakuraˀ	tamaneˀgi
			LHF	LHHL
			‘cherry blossom’	‘onion’
3b			kuziraˀ	azisaˀi
			HHF	HHHL
			‘whale’	‘hydrangea’
4a				murasakiˀ
				LHHF
				‘purple’
4b				aburamiˀ
				HHHF
				‘fat’

[130] It should be pointed out that accentual H tones different from phrasal tones in that they also avoid syllables with high vowels with voiceless onsets that are followed by a syllable with a non-high vowel (see Nitta 2001). However, there are no shifts of accents from such syllables in noun-final position to following particles, as in Matsue and Ichihara Japanese. In this chapter, only phrasal tones are treated.

According to Uwano and Nitta (1983), words like *toriˀ* are realized not as [HF] but rather as [HM] or [H-ML]. They also point out that this causes near-neutralization with initially accented words such as *kaˀgi* 'keyˀ, but that in the latter word there is a larger difference between the pitch levels of the first and second moras, the first mora in *kaˀgi* being higher than the high tone in the initial mora in *toriˀ*, and the low tone in the second mora in *kaˀgi* lower than that of the second mora in *toriˀ*.

One could still argue that apart from the lexical tones of the pitch accent, the so-called phrasal tones too are lexical, and subject to a process of deletion in non-phrase-initial position. If so, there is no need to specify the conditions in (23) in the grammar. While this is certainly a possible analysis, there is evidence from verbal forms (Nitta 1989) and compound accentuation (Uwano 1996, 1997; Nitta 2001) that the conditions in (23) do play a role in the synchronic grammar. The verbal forms in (26) are taken from Nitta (1989).

(26) a.	yabuˀru	LHˀL	'tear.NONPAST'
	yabuˀri-ni	LHˀLL	'tear.CONT-DAT'
	yaˀburya	HˀLL	'tear.COND'
	yaˀburema	HˀLLL	'tear.IMP'
	yaburan^{0}	HHHH	'tear-NEG'
b.	naraˀbu	LHˀL	'line up-NONPAST'
	naraˀbi-ni	LHˀLL	'line up.CONT-DAT'
	naraˀbya	LHˀL	'line up.COND'
	naraˀbema	LHˀLL	'line up.IMP'
	naraban^{0}	LHHH	'line up-NEG'

Nitta (1989) shows that the verbs in (26) belong to the same class in terms of which forms are assigned an accent.[131] However, the actual location of the accent within the stem differs depending on the the same conditions that were introduced above. As in nouns, the H tone avoids being associated to *GI* in a sequence *GI.CA* (26a).

Something similar can be observed in compounds (as in (27); data from Uwano 1997a). In (27a), in which the accent falls on a syllable with a non-high vowel, the accent is

[131] Even though the accent location differs in the forms in (26), they must be treated as belonging to a single class. See Nitta (1989) for a detailed discussion of verbal accent, including discussion of how not only the first syllable in sequence like *GI.CA*, but also the first syllable in sequences like *KI.CA* is avoided by accentual H tones, the result of which is that there are three different tonal patterns for words of accented verbs. In Nitta (2001: 176) the difference in accent location between the forms *haˀzimaru* 'begin' vs. *kasaˀnaru* 'overlap' vs. *atumaˀru* 'gather' is also mentioned.

faithfully parsed in the compound. However, when an accent is associated with a syllable of the shape *GI* in the same word-final position (27b), the accented is shifted one syllable to the left in the compound.

(27) a.	sakeˀ	LFˀ	‘liquor’
	furumai-zakeˀ	HHHH-HFˀ	‘offered liquor’
b.	ebiˀ	HFˀ	‘shrimp’
	kuruma-eˀbi	HHH-HˀL	‘scampi’

It goes without saying that the Kanazawa patterns can only be phonological; there is no way in which the realization of a H tone on the first syllable of a word could be phonetically dependent on the segmental make-up of the second and third syllables.

5.2.4. The issues

The data presented above raises questions about the nature and the domain of tone-vowel height interaction: is the interaction direct, with constraints or rules in which tone and vowel height refer to each other directly, or indirect, i.e. mediated by a prosodic domain such as the foot, the head of which interacts with both tone and vowel height?

There are two important facts that need to be accounted for. First, there is the fact that in all three dialects, both lexical and phrasal H tones interact with vowel height. This suggests that in both cases H tone is involved, and not accent. Second, in all three dialects, it seems that not only the vowel of the syllable to which the H tone is potentially associated, but also the following vowel plays a role. In other words, the “weakness” of high vowels as TBUs is context-dependent. This is especially true for Kanazawa Japanese, in which not only the vowel height of the vowel in the following syllable, but also the voicing of the preceding onset matters. In this sense, the three dialects are quite similar in the way vowel height interacts with tone. An interesting question therefore is whether it is possible to analyze the different dialects in similar ways.

The type of analysis is of course also related to possible phonetic or functional motivations behind the tone-vowel height interactions. Sugito’s (1981/1982, 2003) work on Osaka Japanese might serve as a guide in looking for answers in this respect. It seems likely that similar physiological factors set the accent shifts in the three dialects in motion at an earlier stage. An important difference between Osaka Japanese on the one hand and Matsue, Ichihara, and Kanazawa Japanese on the other, is that in the former the physiological patterns have not been phonologized, while in the latter they arguably have. In Osaka Japanese, both *kuˀsi* ‘comb’ and *kuˀsa* ‘grass’ are still perceived as having initial

accent, and there are no reports that in words with non-initial accent, the phonological location of the H tone differs depending on vowel height. The same cannot be said about the three dialects discussed above. There is broad consensus in the literature that in Matsue and Kanazawa Japanese, the accent system has been reorganized in such a way that H tones in certain environments simply do not occur. Similar claims are made for Ichihara Japanese by Kinda'ichi (1942) and Fujiwara (1972, 1979), and although not all claims could be confirmed by the fieldwork I conducted, the dialect is clearly different from dialects like Tokyo and Osaka Japanese in the degree to which H tones are sensitive to vowel height.

From the point of view of phonology, there are a number of different possible ways in which we may interpret the interaction between tone and vowel height. In his discussion of Matsue Japanese, Haraguchi (1984) suggests that the low sonority of high vowels causes the accent and tone shifts.[132]

Nitta (2001) proposes that high vowels (followed by non-high vowels) are "too weak" to bear pitch changes (see also Uwano 2012a). In a discussion of Nitta's (2001) paper, Odden (2001) proposes that it is the shorter length of high vowels that triggered the alternation in Matsue Japanese. However, this is unlikely to be the whole story, because this does not explain why H tones do not shift to the next vowel when the vowel of the immediately following syllable is also high. If the shift is blocked because the following high vowel is also too short for the H to be realized, why would the H tone not shift to the leftmost non-high vowel following a lexically accented vowel?

An explanation in terms of length may be possible though if we take mora isochrony (see Warner and Arai 2001) into account. The longer intrinsic length of non-high vowels might cause the onset of the same syllable to be produced relatively early as opposed to the onset of syllables with high vowels. As a result of this, the preceding vowel is shortened, and therefore not able to bear a H tone. While such an explanation sounds attractive, it is not clear how it can deal with the failure of a H tone to be associated to the second syllable in Matsue forms like *kabuQta*ˀ {LLØHˀ} 'put on-PAST' and *nigiQta*^{0} {LLØH} 'clasp-PAST', in which H avoids high vowels in bimoraic second syllable. Also, it seems that the location of the H tone is decided before the phonetic implementation. The Matsue form in (9a) and the Ichihara forms in (16) and (17a) suggest that this location is determined in the phonology, because rather than showing a gradual rise from the first to the second syllable, the first two syllables are pronounced on a L tone, with the pitch only rising from the second to the third syllable. Furthermore, in the case of Kanazawa Japanese, the location of the phrasal H tone must be determined in the phonology, as the

[132] As in his 1977 analysis, accent and tone are treated separately by Haraguchi (1984).

interaction between H tone and vowel height/consonant voicing involves a non-local output condition. The following sections investigate what kind of prosodic structure, and what kind of constraints on tonal associations are necessary to account for the patterns in the different dialects.

5.3. Matsue Japanese: analysis

Based on the observations made above in section 5.2.1, I assume the surface specifications in (28). In the absence of evidence for the spreading of the trailing L tone to the right in accented words, the moras following the L-toned one are left underspecified. However, assuming they do spread – like the phrasal L tones do when they can – would not affect the analysis.

(28)	Example	Tone pattern	Surface representation
a.	kaˀbuto	HˀLL	HˀLØ
	inoˀti	LHˀL	LHˀL
	atamaˀ	LHHˀ	LHHˀ
	karada	LHH	LHØ
b.	koori	HHH	HØØ
c.	kagamiˀ	LHHˀ	LHHˀ
d.	kuziraˀ	LLHˀ	LLHˀ

The lexical and phrasal pitch patterns that need to be explained can be summarized as in (29). The L tones between brackets in (29b) are the tones that would be assigned to a following word in the same phrase.

(29) a. Phrasal tone assignment

i.	$CV_1.CA_2(.CV_3)$ →	LH..	e.g.	karada^0	LHØ	‘body’
ii.	$CV_1.CI_2(.CI_3)$ →	LH..	e.g.	sizimi^0	LHØ	‘mussel’
iii.	$CV_1.CI_2.CA_3$ →	LLH..	e.g.	hazime^0	LLH	‘beginning’

b. Shifts of noun-final accentual H tones

i.	asiˀ	LH	vs.	asi-gaˀ	LLHˀ(L)	‘leg(-NOM)’
ii.	kagamiˀ	LHH	vs.	kagami-gaˀ	LHØHˀ(L)[133]	‘mirror(-NOM)’

[133] Here I follow the AM-analysis of Tokyo Japanese and assume that the H tone of the phrase-initial LH complex is not the same tone as the lexical H. Although this analysis seems obvious because of the need for such a tone in unaccented words, below we will actually see that it may be better to assume that in accented words the accentual Hˀ spreads to the left (or that the two tones fuse into a single H tone).

5.3.1. Problems for a foot-based analysis of tone-vowel interaction

As was mentioned above, if de Lacy's (2007) claim that the interaction between tone and sonority must be mediated by foot structure is correct, we may expect evidence for foot structure in a dialect like Matsue Japanese. Also, the fact that the notion of "weak vowels" reminds us of languages with metrical stress (see Hayes 1995) may be a reason to be optimistic about the chances of finding evidence for foot structure in this dialect.

A metrical analysis would have to be based on the idea that high vowels are preferred in metrically weak positions, and H tones preferred in metrically strong positions. Constraints that reflect these cross-linguistic preferences are proposed by de Lacy (2007).

(30) Two potentially relevant constraints (adapted from De Lacy 2007)

a.	*NONHEAD$_{Ft}$/H:	High tones do not appear in foot non-heads.
b.	*HEAD$_{Ft}$/*i•u*:	High vowels do not appear in foot heads.

By adapting the two constraints in (30), an analysis in which H tones and high vowels are sensitive to foot structure, but not directly to each other, becomes a theoretical option.

The next step is to find out whether these two constraints help to account for the observed patterns. Consider the adjectival forms in forms in (31).

(31) a.	surudoˀ-i	LLHˀL	'sharp-NONPAST'
b.	sabisiˀ-i	LHHˀL	'lonely-NONPAST'
c.	kurusiˀ-i	LHHˀL	'painful-NONPAST'

It seems natural to assume that the quadrimoraic words in (31) would consist of two feet each: *(suru)(doˀi)* and *(kuru)(siˀi)*. If these are the metrical parsings, the evaluation of candidates with different possible tonal associations for the two words would proceed as in (32). Forms preceded by an asterisk are those with tone patterns that are not actually attested. The leftward arrow marks a form that should be selected is not selected. As we can see, the two constraints fail to select the right candidate for *sabisiˀi*. Thus, under the assumption that the words in (31) are parsed into two bimoraic feet, the two constraints do not account for the Matsue Japanese patterns. The problem is that candidates without a H tone on a high vowel in the second syllable of the initial foot are always more optimal than those with a H tone on that vowel, regardless of the ranking of *Hd/*i•u*.

(32) Why *NONHD/H and *HD/*i•u* are not sufficient (foot heads underlined)

	*NonHd$_{Ft}$/H	*Hd$_{Ft}$/*i•u*
← a.i. (sa^{L}<u>bi</u>^{H})(<u>si</u>$^{H'}$i^{L})		*!*
ii. (<u>sa</u>^{L}bi^{H})(si$^{H'}$i^{L})	*!	*
→ iii.*(<u>sa</u>^{L}bi^{L})(<u>si</u>$^{H'}$i^{L})		*
b.i. *(su^{L}<u>ru</u>^{H})(<u>do</u>$^{H'}$i^{L})		*
ii.*(<u>su</u>^{L}ru^{H})(do$^{H'}$i^{L})	*!	*
iii.(<u>su</u>^{L}ru^{L})(do$^{H'}$i^{L})		*

No matter what other constraints that refer to foot-internal structure we posit, the fact that *sabisi ̀i* [LHH ̀L] and *surudo ̀i* [LLH ̀L] show different tonal patterns cannot be accounted for by means of the two constraints introduced above, as the expected output form *(sa^{L}<u>bi</u>^{H})(<u>si</u>$^{H'}$i^{L})* is harmonically bounded by **(<u>sa</u>^{L}bi^{L})(<u>si</u>$^{H'}$i^{L})*.

Changing the formulation and definitions of the constraints does not help either. For instance, we could try reformulating the constraints positively rather than negatively (see Yip 2001) as in (33).

(33) Two potentially relevant positively formulated foot structure constraints

a. Hd$_{Ft}$/H: Foot heads have H tones associated to them.

b. Hd$_{Ft}$/*a•e•o* Foot heads have non-high vowels.

The tableau in (34) shows that using these positive markedness constraints makes no difference. This time, the form with the attested tonal patterns in (34b-iii) is harmonically bound by the more harmonic form in (34b-i).

(34) Evaluation using Hd$_{Ft}$/*a•e•o* and Hd$_{Ft}$/H (foot heads underlined)

	Hd$_{Ft}$/*a•e•o*	Hd$_{Ft}$/H
→ a. i. (ku^{L}<u>ru</u>^{H})(<u>si</u>$^{H'}$i^{L})	**	
ii. (<u>ku</u>^{L}ru^{H})(<u>si</u>$^{H'}$i^{L})	**	*
iii.*(<u>ku</u>^{L}ru^{L})(<u>si</u>$^{H'}$i^{L})	**	*
→ b. i. *(su^{L}<u>ru</u>^{H})(<u>do</u>$^{H'}$i^{L})	*	
ii. *(<u>su</u>^{L}ru^{H})(<u>do</u>$^{H'}$i^{L})	*	*
← iii. (<u>su</u>^{L}ru^{L})(<u>do</u>$^{H'}$i^{L})	*	*

It should be clear that constraints that mediate between tone and vowel height in a metrical

analysis are not useful. In the absence of further enhancements to this approach, we may conclude that a foot-based approach to tone-vowel height interaction in Matsue Japanese is not available. Therefore, it seems necessary to allow for constraints that regulate tone-vowel height interaction directly.

5.3.2. Tone and vowel height interacting directly

5.3.2.1. Prominence alignment

The first step is to formalize the interaction between tone and vowel height in terms of a "prominence alignment" (Prince & Smolensky 1993) constraint. Tanaka (2003) points out that Prince and Smolensky's (1993) theory of harmonic alignment actually predicts that tone and sonority might interact, because both tone and sonority involve prominence scales. The prominence scales for tone and sonority in Matsue Japanese are given in (35a) and (35b), and the constraint rankings that can be derived from these scales are given in (35c). Note that we do not need to divide the group of non-high vowels into low and non-low vowels, because they behave in the same way in terms of their interaction with H tones.

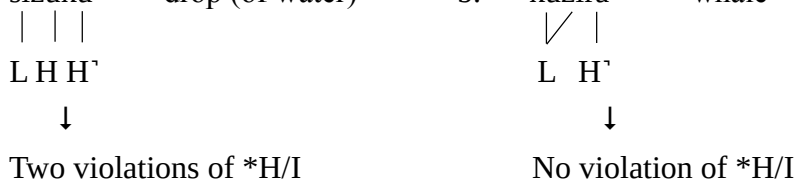
(35) a. H > L
b. a, o, e >> i, u
c. *H/*i•u* >> *H/*a•e•o*
*L/*a•e•o* >> *L/*i•u*

As only the constraint *H/*i•u* seems to be relevant, it is defined as in (36).

(36) *H/$V_{[high]}$ = *H/I
A H tone is not associated to a syllable headed by a high vowel.
(One violation is assigned for every H tone that is linked to one or more high vowels).

To see how this constraint works, consider the forms in (37). If we adopt the AM-approach to phrasal tones in Tokyo Japanese for Matsue Japanese, according to which the word-initial rise in pitch is a property of the phrasal level, a form like (37a) would contain two H tones, an accentual one and a non-accentual one. However, if this is the case, there is no reason why the form in (37b) is not allowed to assign a phrasal H tone to the second syllable and violate the same constraint once.

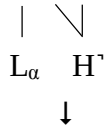
(37) a. sizuku꜀ ‘drop (of water)’ b. kuzira꜀ ‘whale’



Two violations of *H/I No violation of *H/I

This means that if we adopt the constraint *H/I, we cannot assume the representation in (37a). Instead, it is necessary to assume that the lexical H tone is linked to both the third syllable (its sponsor) and the second syllable, as in (38).

(38) sizuku꜀ ‘drop (of water)’



One violation of *H/I

The representation in (38) also involves a violation of *H/I, but one that cannot be avoided. Because spreading the H tone does not add any violations of the constraint, it spreads to the second syllable.

A potential problem for the formulation of the tone-vowel height constraint is that it makes the wrong predictions for a possible variety in which H tones spread to the right in unaccented words.[134] In such a variety, a form of the shape CV.CI.CA.CI would be realized as *$CV^{L}.CI^{H}.CA^{H}.CI^{H}$ rather than $CV^{L}.CI^{L}.CA^{H}.CI^{H}$, because *H/I is violated anyway by the association of the H tone to the final syllable. In order to rule out such potential forms, the *H/I can be replaced by the alignment constraint in (39c), which is a Boolean conjunction (Crowhurst and Hewitt 1997) of the two constraints in (39a/b).[135]

(39) a. ALIGN(H, Left, [-high], Left) = ALIGN-L(H, A)
The left edge of a H tone is aligned with the left edge of [-high].
b. ALIGN(H, Right, [-high], Right) = ALIGN-R(H, A)
The right edge of a H tone is aligned with the right edge of [-high].

[134] As discussed in Chapter 2, in Tokyo Japanese both H-spreading and interpolation of the phrasal H with the utterance-final L can be observed (Sugahara 2003). It is quite likely that the same situation holds for Matsue Japanese.

[135] Crowhurst and Hewitts's (1997) Boolean conjunction is based on the idea that a conjunction "A ∧ B is true if and only if each conjoined expression, or proposition, is true" (Crowhurst and Hewitt 1997: 7). This type of Boolean conjunction thus differs from Local Constraint Conjunction in Smolensky (1993, 2006), which Corresponds to Boolean disjunction.

c. ALIGN-L(H, A) ∧ ALIGN-R(H, A) = ALIGN(H, A)
The left and right edges of a H tone are aligned with the left and right edges of [-high] (One violation is assigned for each high vowel to which a H tone is left- or right-aligned).

Stated informally, ALIGN(H, A) only penalizes the association of H tones to syllables or moras dominating high vowels that are on the left or right edge of a stretch of TBUs that share the same H tone. Therefore, in cases in which a H-toned high vowel is not the leftmost or rightmost vowel with this H tone, the constraint is not violated. Apart from this, the violations of ALIGN(H, A) will be the same as those of *H/I. Therefore, the alignment constraint is adopted instead of *H/I. Notice that the constraint can be considered a formalization of Sugito's (1981/1982, 2003) empirical observation that the target of H tones is preferably anchored to non-high vowels. Furthermore, it makes sense from the point of view of harmonic alignment, even if the constraint is not formed by means of the abstraction of a markedness constraint derived from an independently motivated hierarchy.

Next, we may consider why a H tone is linked to the second TBU in the first place. What comes to mind is a constraint against an initial sequence of two L-toned TBUs, i.e. an initial tonal lapse.[136] Let us call this constraint *LAPSE-LEFT (adapted from Ito 2013; see Gordon 2002 for a stress-based version of the same constraint; see Zoll 2003 for more on tonal lapses).

(40) *LAPSE-LEFT = *LAPSE-L
A p-phrase does not start with a sequence of two non-H-toned (sonorant) moras.

As phrase-initial tonal lapses (i.e. word/phrase-initial [LL..]) are allowed in order to satisfy the constraint that regulates tone-sonority interaction (e.g. *kuzira*ˀ [LLHˀ]), the ranking must be ALIGN(H, A) >> *LAPSE-L. In (41) it is shown that even though ALIGN(H, A) is violated, *LAPSE-L is still relevant in deciding on the output form of *sizuku*ˀ [LHHˀ].

[136] The same is suggested by Odden (2001) for Kanazawa Japanese, in which the ban against initial tonal lapses is never violated (if we only count sonorant moras as TBUs), as we will see in Section 5.5.

(41) Evaluation of *sizuku* ‘drop (of water)’

sizuku | H’L	ALIGN (H, A)	*LAPSE-L
(a) sizuku | | | LH H’(L)	**!	
→ (b) sizuku | \ L H’(L)	*	
(c) sizuku / | L H’(L)	*	*!

It is not unimportant to consider what happens when the accent of *shizuku*᾽ shifts to a following particle. In such a case, ALIGN(H, A) is not violated, which means spreading of H would involve an additional violation of ALIGN(H, A). However, forms in which the first three moras are all sonorant and have L tones are not reported by Hiroto and Ōhara (1952). This means that there must be a constraint against a tonal “extended lapse” or “long lapse” (Nespor and Vogel 1989; Elenbaas and Kager 1999; Gordon 2002; Kager 2005), *LONG-LAPSE-LEFT.

(42) *LONG-LAPSE-LEFT:
A p-phrase does not start with a stretch of three non-H-toned (sonorant) moras.

Obviously, this constraint would refer to three TBUs in a row, which means it is not strictly local (Buckley 2009). One possible solution for this is to posit a word initial foot, and reformulate the constraint in terms of foot structure.

(43) *LONG-LAPSE-LEFT (foot-based):
A p-phrase does not start with a stretch of L-toned TBUs that is longer than a foot.

The absence of initial L-toned lapses is the only potential piece evidence for metrical structure that can be found in the tonal system of Matsue Japanese nouns. If this word-initial foot really exists, it is all the more striking that it does not play a role in the interaction between tone and vowel height.[137]

[137] See Chapter 6 for evidence for foot structure in Matsue Japanese from verbal accent.

The next issue that must be accounted for is where the phrase-initial L tone comes from. The most straightforward way to enforce the insertion and association of a L tone to the left edge of an accentual phrase is to assume an alignment constraint like the one in (44).

(44) ALIGN(Φ, Left, L, Left) = ALIGN-L(Φ, L)
The left edge of a p-phrase is aligned with the left edge of a L-tone.

This L tone needs to be inserted, and thus violates a constraint against the insertion of L tones, DEP-L, which is ranked below ALIGN-L(Φ, L). The constraint ALIGN-L(Φ, L) must also outrank *LAPSE-L, because violations ALIGN(H, A) are not avoided by associating the H tone to the initial syllable at the expense of the initial L tone, which means ALIGN-L(Φ, L) outranks ALIGN(H, A). If ALIGN(H, A) >> *LAPSE-L, by transitivity ALIGN-L(Φ, L) >> *LAPSE-L

(45) Evaluation of *sizuku*ˀ 'drop (of water)'

sizuku | HˀL	ALIGN-L (Φ, L)	*LAPSE-L	DEP(L)
→ (a) sizuku | \ L Hˀ			*
(a) sizuku / | L Hˀ		*!	*
(b) sizuku \ | Hˀ	*!		

Next, we need to explain why the boundary L tone associates to the first two TBUs in words like *kuzira*ˀ [LLHˀ] rather than leaving the second TBU unspecified for tone. For this I posit a constraint that forces a L tone to be adjacent to a H tone.

(46) ALIGN(L, Right, H, Left):
The right edge of a L tone is aligned with the left edge of a H tone.[138]

[138] Because it is assumed that the lexical H tone may spread to the left, we cannot adopt the constraint CONCATENATE which forces tones belonging to bitonal morphemes to align with one another (Riad 1998; Gussenhoven 2004). The alignment constraint could be the result of tones on the tonal tier being organized in a tonal complex (Akinlabi & Liberman 2001; Lee 2006).

As long as this constraint is ranked above a constraint that prohibits spreading (*SPREAD), the right predictions are made.[139]

In unaccented words, a H tone is inserted to avoid a violation of *LAPSE-L. The violation of DEP(H) ("don't insert H tones") that this entails is omitted from the tableau in (47). The tableau in (47) also shows the role of ALIGN(H, A) in words with an initial tonal lapse.

(47) Evaluation of *hazime*^{0} 'beginning'

hazime	ALIGN-L (Φ, L)	ALIGN (H, A)	*LAPSE-L	DEP(L)
(a) hazime | | LH		*!		*
→(b) hazime \ | L H			*	*
(c) hazime \ \ | H	*!			

The assumption that lexical H tones spread leftward is not uncontroversial. Therefore, independent evidence for this spreading operation would be welcome. The compounds in (48), which are reported in Hiroto and Ōhara (1952), may be viewed as evidence for this. The L tones between brackets correspond to non-sonorant moras that can be assumed to be toneless.

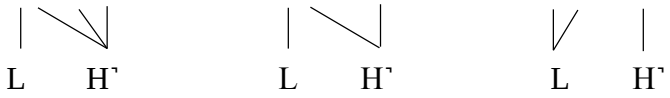
(48) a.	wari-kɯ̥tiˀ	LL(L)Hˀ	'slander'
	wari-kɯ̥tɨ̥-gaˀ	LL(LL)Hˀ	'slander-NOM'
b.	hari-sɨ̥goˀto	LL(L)HˀL	'needlework'
	hari-sɨ̥goˀto-ga	LL(L)HˀLL	'needlework-NOM'

The forms in (48a) both violate *LAPSE-L, even though in the isolation form ALIGN(H, A) is necessarily violated. Spreading the H tone to the second syllable would satisfy *LAPSE-L, yet spreading does not take place. The reason for this seems to be that spreading is blocked due to the devoiced vowel in the third syllable. The markedness constraint *T/$\mu_{[-son]}$ can be hold responsible for this. Gapped representations in which a

[139] In accented words, a constraint against multiple peaks in a p-phrase will block the insertion of a tone to the right of the trailing L tone.

TBU is skipped (49b) are not well-formed (Archangeli and Pulleyblank 1994), so the only solution is to leave the second syllable without a H tone. The initial L tone is spread to the second syllable instead. I assume the L does not spread to the syllable with the devoiced syllable, again because of *T/$\mu_{[-son]}$.

(49) a. * wari-kṵci b. * wari-kṵci c. wari-kṵci



Crucially, no H tone is inserted, which should be possible if the H tone on the second syllable were independent of that of the accentual H tone and there were no constraint like ALIGN(H, A). In other words, the absence of a non-high vowel in the third syllable is not enough for a high vowel in the second syllable to have a H tone associated with it.

The data in (48) may also be taken as evidence for the phonological nature of devoicing, at least in this particular instance. If the devoicing were phonetic, we would expect a H tone on the second syllable.[140]

We could assume that the spreading of the accentual H is responsible for the H tone on the second syllable in all accented words with an accent on the third syllable or further to the right. Since there is no conclusive evidence for this, I will take a conservative approach and allow spreading of H tones only in situations where this results from the avoidance of violations of both ALIGN(H, A) and *LAPSE-LEFT.

The analysis in which a lexical H tone spreads to the left in order to avoid the violation of a phrase-initial tonal lapse is also proposed by Lee and Zhang (2014) for South Kyungsang Korean, a dialect with a pitch accent system that shares a lot of properties with the Tokyo Japanese accent system.

Finally, it should be mentioned that it could be the case that instead the accentual H tone does not spread to the left, but that the non-accentual H spreads to the right and fuses

[140] The conclusions about H spreading and devoicing being phonological are based on the assumption that the second syllable in the words in (48) is pronounced as a full syllable. However, Hiroto and Ōhara (1952) explain that in syllables with /r/ as an onset, the /r/ and the following vowel tend to get deleted, with concominant lengthening of the vowel of the preceding syllable. If this process applies to the cases in (27), the surface forms would be *waa-kṵci* and *haasi̥goto*. Indeed, the form *waa-kuci* is attested in data of the "Central Izumo dialect" described by Katō (1937). If these are the actual surface forms, the fact that the H tone does not link to the second mora can be regarded as an effect of the constraint *H/NONHEAD$_{\sigma}$. While it cannot be completely ruled out that the failure of H tone spreading results from *r*-deletion, Hiroto and Ōhara (1952) do explicitly list more casual forms with *r*-deletion, so there is no strong reason to doubt the forms were not actually pronounced the way they are given in (48).

(see e.g. Myers 1997) with the accentual H. The important thing is that it seems that the stretch of high pitch must be analysed as phonologically represented by a single H tone.

5.3.2.2. Constraints on the displacement of accentual H tones

In this section, we will look at how the displacement of accentual H tones from noun-final high vowels to a following particle can be modelled. Some relevant data is repeated in (50).

(50) Pitch patterns of Matsue Japanese nouns

a.	kaˀbuto	HˀLØ	kaˀbuto-ga	HˀLØØ	kaˀbuto-kara	HˀLØØØ
b.	inoˀti	LHˀL	inoˀti-ga	LHˀLØ	inoˀci-kara	LHˀLØØ
c.	ituˀtu	LHˀL	ituˀtu-ga	LHˀLØ	ituˀtu-kara	LHˀLØØ
d.	atamaˀ	LHHˀ	atamaˀ-ga	LHHˀL	atamaˀ-kara	LHHˀLØ
e.	kagamiˀ	LHHˀ	kagami-gaˀ	LHØHˀ	kagami-kaˀra	LHØHˀL

The fact that H tones are allowed to shift at all shows that some faithfulness constraint that prohibits such shifts can be violated. Let us assume the relevant constraint is NOFLOP(PROM) (see Chapter 4).

(51) NOFLOP-PROM = NOFLOP
Corresponding accents have corresponding sponsors and links.[141]

There are two important facts about the data that needs to be accounted for: (i) a H tone never shifts to the left, and (ii) a H tone may only shift to the immediately following syllable. To give concrete examples of both conditions, there are no forms like **kagaˀmi* instead of *kagamiˀ*, and a word like *ituˀtu* does not show an accent shift when followed by the nominative particle (**itutu-gaˀ*). Although we could posit a constraint against leftward tone shifts, this is not really insightful. Although it is true that (local) rightward movement is natural in tone languages (Hyman 2007), it would be preferable if the absence of leftward shifts can be shown to be the result of a more general constraint.

An possible analysis is one in which tone shifts within a particular domain are not allowed. The idea behind this would be that adding association lines to material that could have been linked in the input is not allowed. This idea is formalized in van Oostendorp's (2007) Coloured Containment model in terms of the constraint ALTERNATION. According to ALTERNATION, material with the same colour (i.e. in the same morpheme) must be

[141] See Alderete (1999) for a formal definition.

linked by association lines of the same colour (i.e. association lines that are specified in the input). However, this constraint cannot be adopted, because the spreading of H tones to the left in order to avoid phrase-initial lapses is allowed.

The basic idea behind the constraint can be adopted, however: NOFLOP can be relativized to the domain of the p-word, which makes it an OO-constraint.

(52) NOFLOP-PROM-ω
Within a p-word, corresponding accents have corresponding sponsors and links ("No accent shift to a TBU that belongs to the same p-word as the sponsor").

When the H tone is shifted to a monomoraic particle, the trailing L tone is deleted or left unassociated, violating PARSE(L), which is formulated in such way that it bans both the deletion and the floating of lexical L tones.[142]

(53) PARSE(L)
A L tone in the input is associated to a TBU in the output.

Avoiding a H tone linked to a high vowel is more important than parsing the trailing L tone: ALIGN(H, A) >> PARSE(L). However, a tone may not flop to a syllable that in the output is in the same p-word as the sponsor, hence the ranking is NOFLOP-ω >> ALIGN(H, A) >> PARSE(L) as shown in (54).

(54) H linked to the particle in *kagami-ga*ˀ

kagami-ga | H`L	NOFLOP-ω	ALIGN (H, A)	PARSE(L)
(a) [kagami]ga | | H`L	*!		
(b) [kagami]ga | | H' L		*!	
→ (c) [kagami]ga | H'			*

[142] Note that in an approach based on an underlying diacritic accent or an underlying H' only, this constraint would not be relevant. In such an analysis, however, a constraint that demands the accented (and H-toned) mora to be followed by a L-toned mora could play a similar role. In any case, it does not seem that assuming the L tone is part of the input is of crucial importantce to the claims made in this section.

Although this works fine, the ranking in (54) makes the wrong predictions for forms with particles starting with syllables headed by high vowels. As can be seen in (55), both the candidate in which the H tone is associated to the noun-final syllable (55b), as well as the one in which the H tone is linked to the particle (55c), violate ALIGN(H, A). This means that PARSE(L) becomes decisive, favoring the form that is not attested in Hiroto and Ōhara (1952).[143]

(55) Wrong prediction: H not linked to the particle in *kagami-ni*ˀ

kagami-ni | H'L	NOFLOP- ω	ALIGN (H, A)	PARSE(L)
(a) kagami-ni | | H'L	*!		
→ (b) *kagami-ni | | H' L		*	
←(c) kagami-ni | H'(L)		*	*!

A remedy to this problem is to relativize ALIGN(H, A) to p-words, as in (56).

(56) ALIGN(H, A)-ω
A H tone is not aligned to the left or right edge of a syllable headed by a high vowel within a p-word.

The indexing of a conjoined constraint is no problem, as conjoined constraints are normally defined as applying in a particular domain in any case (Smolensky 1993). In this case, the domain is the p-word. The tableau in (57) demonstrates that adopting this constraint makes the right predictions.

[143] As mentioned above, the winning candidate in (55b) is argued to be grammatical in Okuda (1975). Also, it is the correct output form in Ichihara Japanese (see Section 5.4).

(57) H linked to the particle in *kagami-niˀ*

kagami-ni | HˀL	NOFLOP- ω	ALIGN (H, A)-ω	PARSE(L)
(a) [kagami]ni | | HˀL	*!		
(b) [kagami]ni | | HˀL		*!	
→ (c) [kagami]ni | Hˀ(L)			*

There is one more thing we need to account for: the local nature of the tone shift. Why does the accentual H in words like *ituˀtu* 'five' not shift to a particle in a form like *ituˀtu-ga* 'five-NOM', i.e. **itutu-gaˀ*? It should be clear that a metrical approach in which the left edge of a foot is left-aligned with the sponsor is not available, as this runs counter to the accent shift from the word-final syllable to the particle, which occurs across a p-word-boundary. Here I propose to adopt the constraint in (58c), which is a Boolean disjunction of the tone-sponsor alignment constraints in (58a) and (58b).

(58) a. ALIGN(H, Left, Sponsor, Left):
Align the left edge of a H tone with the left edge of its sponsor.

b. ALIGN(H, Left, Sponsor, Right):
Align the left edge of a H tone with the right edge of its sponsor.

c. ALIGN(H, Left, Sponsor, Left) ∨ ALIGN(H, Left, Sponsor, Right) = ALIGN(H, Sponsor):
Align the left edge of a H tone with the left or right edge of its sponsor. (Satisfied if either of the disjoined constraints is satisfied)

This constraint must of course outrank ALIGN(H, A)-ω, as shown in (59).

(59) H linked to the particle in *ituˀtu-ga*

itutu-ga | HˀL	ALIGN (H, S)	NOFLOP- ω	ALIGN (H, A)-ω	PARSE(L)
(a) itutu-ga | | HˀL			*	
(b) itutu-ga | Hˀ(L)	*!			*

If we assume that both H and L tones are linked in the lexicon, an alternative solution is available. If L tones are linked in the input, we can adopt a constraint that militates against the delinking of L tones: PARSE-LINK(L).[144]

(60) PARSE-LINK(L):
Input associations of a L tone with a TBU are kept.

The assumption that L tones are not only linked in the output, but also in the input, makes sense from the viewpoint of lexicon optimization (Prince and Smolenksy 1993; Inkelas 1995; Krämer 2012; van Oostendorp 2014), according to which the selected input form will be as close as possible to the output form. This constraint does the same job as ALIGN(H, Sponsor).

(61) H linked to the particle in *itu˩tu-ga*

itutu-ga | | H˩L	PARSE-LINK(L)	NOFLOP-ω	ALIGN (H, A)-ω	PARSE(L)
(c) itutu-ga | | H˩L			*	
(d) itutu-ga | H˩(L)	*!			*

It does not matter which of the two analyses we choose.[145] The point is that the local nature of the tone shift can be accounted for in terms of constraints on tonal association.

Altogether, the analysis presented in this section shows that in a tonal approach, the interaction between tone and vowel height can be accounted for in terms of harmonic alignment. Foot structure was shown to play no role in the interaction, even though a word-initial foot may be necessary to account for the absence of long word-initial tonal lapses.

In Chapter 6, pitch alternations in the verbal paradigm of Matsue Japanese are analyzed.

[144] Constraints that penalize the removal of association lines are not only known as PARSE-LINK (Ito, Mester and Padgett 1995), but also known as MAX-PATH (Pulleyblank 1996), NO-DELINK (McCarthy 2000), and *DISASSOCIATE (Yip 2002).

[145] Note that in an approach in which the accent is analyzed as an underlying diacritic accent, the constraint PARSE-LINK(L) is not available. ALIGN(H, Sponsor), on the other hand, could be redefined in terms of a diacritic accent.

5.4. Ichihara Japanese: analysis

In 5.2, I demonstrated that it is not clear whether the peak delay of H tones not immediately preceded by a L tone (i.e. "spreading H tones") described by Fujiwara (1972, 1979) must really be assumed to be phonological. On the other hand, the environments in which peak delay tends to occur are the same as the environments that are relevant to the assignment of phrasal tones (LH.. vs. LLH..) and the accent shift of final accent to the following particle, for which there is less reason to doubt its phonological status. Therefore, let us start with these two phenomena, and come back to the question of H-spreading later.

5.4.1. The assignment of phrase-initial LH

For Ichihara Japanese, I again posit tonal surface representations that involve tonal underspecification.

(62)	Example	Tone pattern	Surface tones	Gloss
a.	kaˈbuto	HˈLL	HˈLØ	'helmet'
	komuˈgi	LHˈL	LHˈL	'wheat'
	nioiˈ	LHHˈ	LHHˈ	'smell'
	nezumi	LHH	LHØ	'mouse'
b.	tobira	LLH	LLH	'door'
	musumeˈ	LLHˈ	LLHˈ	'daughter'

The phrase-initial association of tones in words with accent on the second syllable or further to the right, as well as unaccented words, can be analyzed in exactly the same way as in Matsue Japanese. In order to avoid repetition, I refer to 5.3.2.1.

5.4.2. Accentual H tones

In the analysis proposed for Matsue Japanese, the behavior of word-final accentual H tones that show a surface shift from the word-final high vowel to particle-initial position was analyzed by means of the constraints NOFLOP-ω and ALIGN(H A)-ω. The same analysis can be adopted for Ichihara Japanese, but with one difference: the domain in which ALIGN(H, A) is relevant is the p-prase and not the p-word, as in (63). The reason for this is that a noun-final H tone associated with a high vowel does not shift to the right when the noun is followed by a particle with a high vowel in the initial syllable, such as the dative *-ni*. For ease of comparison with the Matsue dialect, the same example (*kagami*ˈ 'mirror') is used.

(63) H linked to the noun-final high vowel in *kagami-ni*ˀ (‘mirror-DAT’)

kagami-ni | H’L	NOFLOP- ω	ALIGN (H, A)-Φ	PARSE(L)
(a) kagami-ni | | H’L	*!		
→(b) kagami-ni | | H’ L		*	
(c) kagami-ni | H’(L)		*	*!

As we can see in (63), the association of the H tone to the particle *-ni* incurs the same violation as the candidate in which it is associated with the noun-final syllable, with the result that PARSE(L) becomes decisive.

I mentioned in 6.2 that in the fieldwork data I collected, the accent shift could be observed in the case of bimoraic/disyllabic words, but not in longer words. Whether the difference with the data in Fujiwara (1972, 1979) is of a synchronic or diachronic nature is not clear. If the patterns observed in my field data is representative of some variety of Ichihara Japanese, what could be the reason why words of two syllables and words of more syllables would behave differenty in terms of accent shift?

What comes to mind is a role for metrical structure. Under an analysis in which feet are bimoraic, a bimoraic/disyllabic word will consist of exactly one foot. As the only foot of the p-word, this foot is the initial foot as well as the final foot. Now, suppose that the constraint ACCENT-TO-HEAD is undominated, and there is a constraint that promotes initial trochaic feet, which is informally defined in (64).

(64) INITIAL-TROCHEE
P-word-initial feet are trochaic.

If this constraint has the same ranking as ALIGN(H, A) and *LAPSE-L and there are no relevant higher-ranked constraints, lower-ranked constraints become decisive when both of these constraints are violated. This can be seen in (65), where PARSE(L) is decisive in the selection of the non-shifting candidate in (65b). Assuming that the accent must be part of foot, accented particles are footed. Because the constraint *LAPSE-L is involved, associated boundary tones are also specified for all candidates.

(65) H linked to the noun-final high vowel in *mimi-ni*ˀ (‘ear-DAT’)

mimi-ni H˙L	NOFLOP- ω	ALIGN (H, A)	INIT TROCH	*LAPSE- L	PARSE (L)
(a) (mimi)ni H˙L	*!	*			
→(b) (mimi)ni L H˙ L		*	*		
(c) (mimi)(ni) L H˙(L)		*		*	*!

While this is true when the particle is *-ni*, when *-ga* follows, the absence of a violation of ALIGN(H, A) makes the candidate in which the tone has shifted the optimal form.

(66) H linked to the particle in *mimi-ga*ˀ (‘ear-NOM’)

mimi-ga H˙L	NOFLOP- ω	ALIGN (H, A)	INIT TROCH	*LAPSE- L	PARSE (L)
(a) (mimi)ga H˙L	*!	*			
(b) (mimi)ga L H˙L		*	*!		
→(c) (mimi)(ga) L H˙(L)				*	*

Furthermore, as we can see in (67), the proposed ranking is also able to account for the absence of an accent shift in forms longer than two syllables. In this example, the particle consists of two moras, which means the constraint PARSE(L) is not relevant. However, a more general version of NOFLOP ensures the selection of the form in (67b), which is faithful to the underlying association of the H tone.

(67) H linked to the noun-final high vowel in *kagami˺-kara* (‘mirror-ABL’)

kagami-kara H˺L	NOFLOP-ω	ALIGN (H, A)	INIT TROCH	*LAPSE-L	PARSE (L)	NOFLOP
(a) (kaga)(mi)(kara) H˺ L	*!		*			
→(b) (kaga)(mi)(kara) L H˺ L		*				
(c) (kaga)(mi)(kara) L H˺L				*		*!

In this example, the phrasal tone is assumed to be a spread accentual tone, which is satisfies ACCENT-TO-HEAD in (67b) (and in (67c)) because it is associated to the head of a foot. At least in the case of unaccented words, however, the inserted boundary H tone must be assumed to be associated to the non-head of a left-headed foot. An example of this is *sake-ga*^{0} {LHØ} in (68).

(68) Phrasal tones in *sake-ga*^{0} ‘alcohol-NOM’

sake-ga	ALIGN-L (Φ, L)	INIT TROCH	*LAPSE-L	PARSE (L)	H-TO-HEAD
→a. (sake)ga LH					*
b. (sake)ga L H		*!			
c. (sake)ga L H			*!		
d. (sake)ga H	*!		*	*	*

The association of the H tone to a foot non-head in (68) suggests that we need to distinguish between a H tone that is linked to an accented syllable and other H-toned syllables. Theoretically, this could mean two things. First, it could be that the accentual H tone has a different representation in the grammar (e.g. a H tone linked to a mora or syllable with a diacritic accent mark). In terms of constraints, this means that there is a fundamental difference between ACCENT-TO-HEAD, a constraint that refers to accentual H tones, and a more general H-TO-HEAD, which refers to H tones in general. This is the approach that is taken in (68). Alternatively, there could be a difference in the constraint

ranking between the lexical and postlexical grammar: H-TO-HEAD is highly ranked in the lexical grammar, but less high in the postlexical grammar. What the two approaches have in common is hat lexical H tones and phrasal H tones are not the same things, which is an important conclusion. However, I hasten to add this conclusion is based on a tentative analysis of observations of a small amount of field data.

Let us now move on to the delayed realization of accentual H tones that are not associated to a *CI* in a *CI.CA* sequence in the input. The examples from (12) and (13) above are repeated in (69).

(69)	*Input*	*Output tones*	
a.	aˀme-ga	HˀHL	‘rain-NOM’
	aˀsahi-ga	HˀHLL	‘morning sun-NOM’
b.	aˀki-ga	HˀLL	‘autumn-NOM’
	aˀki-ni	HˀLL~HˀHL	‘autumn-DAT’
c.	akaˀ-gai	LHˀHL	‘blood clam’
	temaˀ-cin	LHˀHL	‘wages for labour’
	amaˀ-mizu	LHˀHL	‘rain water’
d.	ao-zoˀra	LHHˀH	‘blue sky’

If the peak delay is phonetic, there is nothing to explain phonologically. However, let us consider the question what kind of analysis would be possible if the peak delay constitutes a case of H-tone spreading.

It is unlikely that H-spreading is the result of constraints on metrical structure. This can be concluded from words like *akaˀ-gai* ‘blood clam’ in (69c), where the H tone spreads from the second to the third syllable. The foot structure of this compound would be *(akaˀ)(gai)*, in which the H tone spreads across feet. In words like *aˀsa-hi* ‘morning sun’ (69a), on the other hand, the H tone would spread within the foot.

The absence of H-spreading in e.g. *kawaˀ-ga* LHL ‘river-NOM’ could be an indication that the driving force behind it is the pressure for (an accentual) H tone to be as close to the right edge of the p-word as possible: ALIGN-R(H, ω).[146] After all, the accent is already rightmost in words with final accent, where this constraint is fully satisfied. Still, if this is the constraint that triggers H-spreading, there must also be a constraint that prevents

[146] Taking the behavior of phrasal tones in unaccented words – which do not seem to spread in the fieldwork data – into account, it is necessary for the constraint to refer to accentual H tones as opposed to non-accentual H tones.

iterative spreading. The constraint proposed in (58c), repeated in (70), serves our purposes here.

(70) ALIGN(H, Sponsor):
Align the left edge of a H tone with the left or right edge of its sponsor.

If ALIGN(H, Sponsor) >> ALIGN-R(H, ω), H tones will maximally spread one mora to the right, as in (71). In order to indicate that only the first syllable to which the accentual H tone is associated is accented, the diacritic hook is placed on the segmental material.

(71) Bound tone spreading in *aˀsahi-ni* 'morning sun-DAT'

aˀsahi-ni | HL	ALIGN (H, Sponsor)	ALIGN-R (H, ω)
a. [aˀsahi]$_{ω}$ni | | HL		
→b. [aˀsahi]$_{ω}$ ni \ / | H L		*
c. [aˀsahi]$_{ω}$ ni \ / | H L	*!	

The conditions for H tone spreading are more complicated in monomoraic nouns. In monomoraic words with a non-high vowel, spreading of H does not occur (frequently), whereas it does in monomoraic words with a high vowel. The absence of a shift from the high vowel to the following particle in (72b-i) can be explained in terms of positional faithfulness (Beckman 1998); faithfulness to accent in word-initial position is stronger than in other positions. However, why would the vowel spread from the noun-final syllable to the particle? What is more, in the latter case, the H even spreads to the dative particle *-ni* (72b-ii).

(72) a. haˀ-ga HL 'tooth-NOM'
b.i. hiˀ-ga HH 'fire-NOM'
ii. hiˀ-ni HH 'fire-DAT'

One possibility is that p-words consisting of a monomoraic foot are allowed when this foot is headed by a syllable with a non-high vowel, but not when headed by a syllable with a high vowel. We would then have the representations in (73).

(73) a.	[(haˀ)]-ga	HL	‘tooth-NOM’
b.i.	[(hiˀ-ga)]	HH	‘fire-NOM’
ii.	[(hiˀ-ni)]	HH	‘fire-DAT’

As the particle now forms the rightmost syllable in the p-word, the alignment approach is able to account for the fact that the H tone spreads to the right.

The discussion in this section shows that if we want to treat H-spreading in Ichihara Japanese as phonological, we can. However, I would like to stress again that more research is necessary in order to confirm whether what is going on is really phonological.

Next we will turn to a dialect for which we can be sure that the tone-vowel height interaction is genuinely phonological in nature, the Kanazawa dialect.

5.5. Kanazawa Japanese: analysis

The Kanazawa data from (23) is repeated in (74), with the proposed surface tonal representations added in the column “Output tones”.

(74)	Condition	Example	Pitch pattern	Output tones	Gloss
a.	..KV_2(..)	kata	LH	{LH}	‘shoulder’
		aki	LH	{LH}	‘autumn’
b.	..GI_2 #	umi	HH	{HØ}	‘sea’’
		mizu-gi	HHH	{HØØ}	‘swimsuit’
c.	..$GI_2.CA_3$	suzume	HHH	{LHØ}	‘mouse’
d.	..$GI_2.CI_3$	nezumi	LHH	{HØØ}	‘sparrow’
e.	CV_1V_1	koori	HHH	{HØØ}	‘ice’
f.	CV_1N_1	oNna	HHH	{HØØ}	‘woman’
g.	CV_1Q_1	koQpu	LLH	{L(Ø)H}	‘cup’
h.	$CV_1C\underset{\circ}{V}_2$	as̥ita	LLH	{L(Ø)H}	‘woman’

As demonstrated in 5.2.3, both phrasal and lexical H tones are subject to the same constraints. In this section we will focus on the behavior of phrasal tones.

Ignoring the pattern in which the second mora is voiceless, Nitta (2001: 169) summarizes the conditions on the two types of phrase-initial pitch patterns in terms of output conditions that could be called schemas. I have taken the liberty of restating his formulations in terms of the abbreviations and representations adopted in this thesis, where C = any consonant, G = $C_{[+voi]}$ or $C_{[+son]}$, I = $V_{[+high]}$, A = $V_{[-high]}$, X = any moraic

segment, # = morpheme boundary.

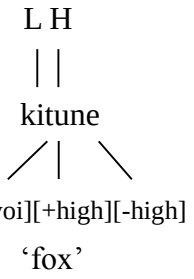
(75)	*Context*	*Tonal Pattern*
a.	(C)V.GI.CA(…)	[H..
	(C)V.GI#	[H..
	(C)$VX_{[+son]}$(…)	[H..
b.	Elsewhere	[LH..

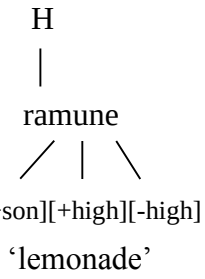
The behavior of $CVX_{[+son]}$ is like in Tokyo and Matsue, and will not be further discussed in this chapter. We will focus on the behavior of CV.GI.CA and CV.GI# as opposed to words with other types of structure.

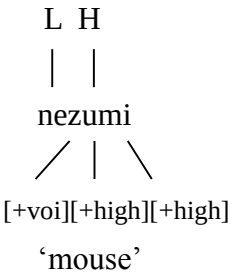
First of all, note that phrases with an initial tonal lapse are not allowed. Therefore, we can assume that *LAPSE-LEFT is undominated. Output forms that violate this constraint will not be considered in the analysis.

Let us now take a closer look at the actual features involved. Capital *G* in (76) stands for [+voi] or [+son]. Treating *G* as simply [+voi] is not possible because sonorant consonants must be assumed to be unspecified for voicing in Japanese.[147] What [+voi] and [+son] onsets have in common, however, is that they are voiceless. In terms of features, they can be said to lack the feature [-voi]. If we adopt this feature for voiceless consonants, the assignment of phrasal H tones in unaccented trimoraic words with different feature structures is as in (76).

[147] This assumption is made for sonorant consonants in Japanese in general. The evidence for this comes from the absence of a blocking effect on *rendaku* that is observed for voiced obstruents (Vance 2015: 406). Vance (2015: 406), referring to Mester and Ito (1989: 277-279), states that rendaku voicing must involve the monovalent feature [voice] rather than [+voice].The reason for this is that, rendaku is blocked even if a syllable with a voiceless obstruent in its onset intervenes between the obstruent onset of the initial syllable of the second element of a compound (the target for rendaku) and a voiced obstruent further in this second element. The example given by Vance (2015: 407) is *ko-hituzi* ‘lamb’ (‘small’ + ‘sheep’), which does not surface as ***ko-bituzi*** even though there is a voiceless obstruent between the two voiced obstruents in the second element. This argument holds if we assume a classical autosegmental version of the Obligatory Contour Principle (OCP; Goldsmith 1976), according to which identical elements that are adjacent are prohibited. However, in a markedness-based approach to the OCP, the blocking of rendaku can be accounted for by means of the self-conjunction of a constraint that bans the feature [+voi] (Ito and Mester 2003). Thus, the feature [+voi] may appear once, but not twice in a morpheme. The domain of morpheme proposed by Ito and Mester (2003) can be replaced by p-word.

(76) a. 

b. 

c. 

The representations in (76) show that syllables with [+high] vowels and [+son] (76b) or [+voi] (76c) onsets are avoided by H tones when followed by a syllable with a [-high] vowel. If the onset of a syllable with a [+high] vowel has the the feature [-voi] as in (77a), however, the association of a H tone will not be blocked, regardless of the features of the vowel in the same syllable and the vowel in the next syllable. In other words, it seems that H tones have an affinity for [-voi] onsets. This tendency can be captured in terms of the following alignment constraint.

(77) ALIGN(H, Left, [-voi], Right) = ALIGN(H, [-voi])
The left edge of a H tone is aligned with right edge of [-voi] ($C_{[-voi]}$).[148]

This constraint is natural from a phonetic point of view, and resembles the constraint H→¬[+voi], which bans the association of a H tone to stretches of material that include voiced segments (see Tang 2008).

Next, let us proceed with the environment [(C)V.GI#, which is easier to tackle than the more complicated environment [(C)V.GI.CA. Compare the words in (78), which are taken from Nitta (2001). Both words lack an onset in their first syllable, and both words have a high vowel in their first and second syllable. This means that the difference in tone patterns must be related to difference in voicing of the onset of the second syllable: [-voi] in (78a), and [+son] in (78b).

(78)

	Structure	Example	Tone Pattern	Gloss
a.	V.CI	iki^{0}	LH	‘breath’
b.	V.GI	umi^{0}	HØ	‘sea’

Because the features of the words in (78) only differ in terms of voicing (and sonority),

[148] Because H tones on syllables with devoiced vowels, which may be said to have the feature [–voi], are ruled out by a higher ranked constraint that says a TBU must be specified with [+son], the H can actually be argued to be aligned with a syllable with the feature [-voi].

ALIGN(H, [-voi]) can indeed be thought of as a potentially important constraint.

Now consider the forms in (79), where "+" stands for a "weak morpheme boundary" and "#" for a "strong morpheme boundary". The morpheme *mi* in (79a) is a Sino-Japanese (SJ) root meaning "taste", whereas the suffix *-mi* in (79b) is a nominalizing suffix that can be attached to adjectival roots.

(79) a.	sibu+mi	LHØ	'severity'
b.	sibu#mi$_{<SJ>}$	HØØ	'astringent taste'

Now, because (80b) behaves like the bimoraic word in (79b), we can be quite sure that the suffix in this word is somehow "invisible". The suffix in (80a), however, must be visible, as it takes the same tonal pattern as the simplex word *nezumi* (LHØ). Both words in (80) can be considered a single p-word, so the difference must lie in a lower level of constituent structure.

Because the words in (79) consist of a single bimoraic noun, we may assume these words are parsed into a single foot. Furthermore, this must also hold for the stem in (80b), since it behaves identically to the word in (79b). If the different tone patterns of the word in (80) are the result of a difference in foot structure, the only possible conclusion is that that the word in (80a) has a word-final foot, as opposed to (80b), which has a word-initial foot. The situation is illustrated in (81).

(80) a.i.	(i^{L}ki^{H})	'breath'
ii.	(u^{H}mi)	'sea'
b.i.	si^{L}(bu^{H}mi)	'severity'
ii.	(si^{H}bu)mi	'astringent taste'

A possible explanation for the difference in foot structure is that the SJ root *mi* is mapped onto a foot, whereas the suffix *-mi* is not. Thus, the footing of the words in (80b-ii) is as in (81).

(81)	(si^{H}bu)(mi)	'astringent taste'

The next step in is to account for the mapping of the SJ morpheme onto a foot. In their overview of Sino-Japanese phonology, Ito and Mester (2015a) propose the constraint in (82), to which I added an informal definition. The constraint basically says that the edges of a Sino-Japanese root must be aligned with the edges of a foot.

(82) MATCH-ROOT$_{SJ}$-TO-FT
A Sino-Japanese root is mapped onto a foot.

Because the foot is a member of the Prosodic Hierarchy, the direct mapping of a root onto a foot does not violate the Indirect Reference Hypothesis.[149] We will return to the analysis of the minimal pair below, after some other necessary constraints have been introduced.

The next challenge is to find out what exacly causes the H tone to get linked to different positions in the structures in (80). Comparing the forms in (80b), the only difference is the position of the syllable *bu* within the foot: in the syllable on the left in (80b-i), and in the syllable on the right in (80b-ii). Because the syllable on the left in both forms in (80b) is the position to which the H tone is aligned, we may hypothesize that this is a foot head. Thus, in (80b-i) the syllable *bu* is a foot head, whereas it is not in (80b-ii). From these facts we can deduce that the constraints in (83) play a role. The constraint in (83b) is a reformulation of H-TO-HEAD in terms of alignment.

(83) a. TROCHEE
Feet are left-headed.
b. ALIGN(H, Left, Head$_{Ft}$, Left) = ALIGN(H, HD)
The left edge of a H tone is aligned with the left edge of a foot-head.

Returning to the forms in (80a), in the form *(i^{L}ki^{H})*, one or the other of the two constraints in (83) is violated; the foot is either left-headed with a H tone on the non-head (*(i^{L}ki^{H})*), or right-headed with a H tone on the foot-head (*(i^{L}ki^{H})*). In the form *(u^{H}mi)*, on the other hand, both constraints are satisfied. What is more, the fact that this output form is chosen must be related to the violation of ALIGN(H, [-voi]) in the unattested form **(u^{L}mi^{H})*, as

[149] Clearly, the approach is not in line with the idea of two separate prosodic and metrical hierarchies. In an alternative approach, we could assume that the different tone patterns of *sibu+mi* 'severity' {LHØ} and *sibu#mi*$_{<SJ>}$ 'astringent taste' {HØØ} are the result of differences in the mapping of m-stems onto p-stems. If the native nominalizing suffix *-mi* is analyzed as a stem-forming suffix rather than a word-forming one, the newly created stem could be mapped onto a p-stem. As for *sibu#mi*$_{<SJ>}$, on the other hand, the two bound morphemes will both be parsed into p-stems, both being content rather than grammatical morphemes. A constraint that aligns feet with p-stem boundaries would rule out a parsing like *[|si(bu||mi)|] 'astringent taste', where both |*sibu*| and |*mi*| are p-stems. An analysis along these lines is not pursued here, because it is not clear whether in *sibu#mi*$_{<SJ>}$ we can speak of stems in the sense of Aronoff (1994), i.e. as "that sound form to which a given affix is attached or upon which a given non-affixal realization rule operates" (Aronoff 1994: 39). In *sibu#mi*$_{<SJ>}$, it seems more likely that we are dealing with two roots (a root being "what is left after all morphological structure has been wrung out of a form"; Aronoff 1994: 40), and not with two stems. Of course, an analysis in terms of the p-root would be possible, but in the absence of clear evidence for a role of the p-root, I assume SJ roots are mapped onto feet.

this is the only relevant difference between the two forms. What seems to be wrong with *$(\underline{u}^{L}mi^{H})$* is that H is both associated to a foot non-head, and to a syllable that lacks a voiceless onset. From this it can be deduced that the following Boolean disjunction is at work.

(84) ALIGN(H, HD) ∨ ALIGN(H, [-voi]) = ALIGN(H, HD ∨[-voi])
The left edge of a H tone is aligned to the left edge of a foot-head, or to the right edge of the feature [-voi] ($C_{[-voi]}$).
(Satisfied if one of the two conditions is satisfied; the domain is the foot)

When we reformulate the constraint in prose, it boils down to the following: within a foot, align H to a foot-head or to a syllable with a voiceless onset.

The evaluation of the two types of bimoraic words is illustrated in (85) and (86), where I assume that (simplex) words with two adjacent monomoraic feet are not allowed: *(μ)(μ). In (85), all three candidates violate a different constraint, but the candidate in (85b) that violates the constraint ALIGN-L (Φ, L), which was introduced in (44), wins. This shows us that ALIGN-L (Φ, L) is ranked below the other two constraints. As before, foot heads are indicated by underlining.

(85) *umi^{0}* 'sea'

umi	ALIGN(H, HD∨[-voi])	TROCHEE	ALIGN-L (Φ, L)
a. ($u^{L}\underline{mi}^{H}$)		*!	
→b. ($\underline{u}^{H}mi$)			*
c. ($\underline{u}^{L}mi^{H}$)	*!		

In (86), ALIGN(H, HD ∨[-voi]) is not violated, the result of which is that a violation of ALIGN-L (Φ, L) (86b) is not tolerated. The fact that the constraint ALIGN(H, HD) is violated by the winning candidate in (86c) suggests that it is ranked below ALIGN-L (Φ, L).

(86) *iki*^{0} ‘breath’

iki	ALIGN(H, HD∨[-voi])	TROCHEE	ALIGN-L (Φ, L)	ALIGN (H, HD)
a. ($i^{L}\underline{ki}^{H}$)		*!		
b. ($\underline{i}^{H}$ki)			*!	
→c. ($\underline{i}^{L}ki^{H}$)				*

In summary, the two different phrasal tone patterns of bimoraic words do not reflect a difference in foot structure, but a difference in tonal alignment and association: a H tone may only be linked to a *GI* sequence if this syllable functions as a foot head (85). H tones are allowed on high vowels that are no foot heads, but only if the high vowel is preceded by a voiceless onset (86).

Next we may return to the trimoraic minimal pair discussed above. In (87a), the constraint MATCH-ROOT$_{SJ}$-TO-FT is not relevant as there is no Sino-Japanese morpheme in the input. The winning candidate in (87a) starts with a phrasal L tone, and at the same time satisfies the tonal alignment and foot structure constraints. In (87b), on the other hand, the candidate without a phrasal L tone (87b-iv) is selected because it is more important to satisfy ROOT$_{SJ}$-TO-FT than to satisfy ALIGN-L (Φ, L). In the absence of evidence for monomoraic feet are not the result of an explicit constraint like ROOT$_{SJ}$-TO-FT, leftover syllables like those in the winning candidate in (87a) are left unparsed.

(87) a. *sibumi*^{0} ‘severity’

sibu- , -mi	ROOT$_{SJ}$ -TO-FT	ALIGN(H, HD∨[-voi])	TROCHEE	ALIGN-L (Φ, L)
→i. $si^{L}(\underline{bu}^{H}mi)$				
ii. $(si^{L}\underline{bu}^{H})mi$			*!	
iii. $(\underline{si}^{L}bu^{H})mi$		*!		
iv. $(\underline{si}^{H}bu)mi$				*!

b. *sibumi*^{0} ‘astringent taste’

sibu-, -mi$_{<SJ>}$	ROOT$_{SJ}$ -TO-FT	ALIGN(H, HD∨[-voi])	TROCHEE	ALIGN-L (Φ, L)
i. $si^{L}(\underline{bu}^{H}mi)$	*!			
ii. $(si^{L}\underline{bu}^{H})(mi)$			*!	
iii. $(\underline{si}^{L}bu^{H})(mi)$		*!		
→iv. $(\underline{si}^{H}bu)(mi)$				*

If foot structure is responsible for the different tonal patterns in this minimal pair, it is natural to expect that the same difference is also involved in other trimoraic words, which show the same variant tone patterns. Thus, in the analysis developed so far, the difference between *suzume* {HØØ} 'sparrow' and nezumi {LHØ} 'mouse' most probably results from the foot parsings in (88).

(88) a. (su^{H}zu)(me) or (su^{H}zu)me, *su(zu^{H}me), *su(zu^{H}me)
b. ne^{L}(zu^{H}mi) or (ne^{L})(zu^{H}mi), *(ne^{L}zu^{H})mi, *(ne^{L}zu^{H})mi

The challenge is again to motivate the different types of foot structure. This time we cannot refer to morphological or morpho-prosodic structure; rather, it must be related to differences in segmental structure. What is wrong with **su(zu^{H}me)* as opposed to *ne^{L}(zu^{H}mi)*? One thing is certain: it must be related to the non-high vowel in the third syllable in **su(zu^{H}me)*. Most likely, the problem is that the non-high vowel is not dominated by a foot head. This idea can be formalized in terms of the constraint in (89), which can be seen as the alignment version of the constraint *NONHEAD$_{Ft}$/*a,e•o* (see de Lacy 2007).

(89) ALIGN(A, Left, Head$_{Ft}$, Left) = ALIGN(A, HD)
The left edge of [-high] (V$_{[-high]}$) is aligned with the left edge of a foot head.

The selection of the output of *suzume* 'sparrow' is presented in (90). Notice that a monomoraic foot with a non-high vowel is preferred to parsing the syllable with the non-high vowel in the non-head of a foot (90b) and to leaving it unparsed (90d).

(90) *suzume^{0}* 'sparrow'

suzume	ALIGN (A, HD)	ALIGN(H, HD∨[-voi])	TROCHEE	ALIGN-L (Φ, L)
a. su^{L}(zu^{H}me)		*!	*	
b. su^{L}(zu^{H}me)	*!			
c. (su^{L}zu^{H})(me)		*!		
d. (su^{H}zu)me	*!			
→ e. (su^{H}zu)(me)				*

Finally, the tableaux in (91) and (92) show that the forms *nezumi* 'mouse' and *kitune* 'fox' are also correctly selected under the proposed constraint ranking. Interestingly, for these words the difference in foot structure does not result in a difference in tone patterns.

(91) *nezumi*^{0} 'mouse'

nezumi	ALIGN (A, HD)	ALIGN(H, HD∨[-voi])	TROCHEE	ALIGN-L (Φ, L)
→a. (ne^{L})(zu^{H}mi)				
b. (ne^{H}zu)mi				*!
c. (ne^{L}zu^{H})mi	*!		*	
d. (ne^{L}zu^{H})mi		*!		

(92) *kitune*^{0} 'fox'

kitune	ALIGN (A, HD)	ALIGN(H, HD∨[-voi])	TROCHEE	ALIGN-L (Φ, L)
a. ki^{L}(tu^{H}ne)	*!			
b. (ki^{H}tu)(ne)				*!
c. (ki^{L}tu^{H})(ne)			*!	
→ d. (ki^{L}tu^{H})(ne)				

The most important generalization within the proposed analysis is that a H tone on a syllable of the shape *GI* is only possible if this H tone is associated to a foot head. This is even the case in words like *mizu*ˀ 'water' (93), in which the accent is on the second syllable, but the H tone spreads to the first syllable to satisfy ALIGN(H, HD∨[-voi]).

(93) *mizu*ˀ 'water': H spreading in final accented words ending in *GI*

mizuˀ HL	ALIGN(H, HD∨[-voi])	TROCHEE	ALIGN-L (Φ, L)
a. (mi^{Hi}zu$^{H\text{ˀ}iL}$)		*!	
→b. (mi^{Hi}zu$^{H\text{ˀ}iL}$)			*
c. (mi^{L}zu$^{H\text{ˀ}L}$)		*!	
d. (mi^{L}zu$^{H\text{ˀ}L}$)	*!		

An interesting consequence of the analysis is that the foot head does not coincide with the H-Sponsor (or the underlying diacritic accent), but with the syllable to its left. This

representation does not violate H-TO-HEAD if this constraint is analyzed in terms of alignment, above in (83b).[150] Evidence for the head status of the left syllable in words like that in (90) comes from Uwano and Nitta (1983), who argue that the phonetic realization of words like this is not [HF] but rather [HM] or [H-ML].[151]

This concludes the analysis of phrasal tone assignment in Kanazawa Japanese. The analysis may appear to be complicated, because on top of metrical structure, it was necessary to posit an alignment constraint that is a Boolean disjunction. However, by doing so, the use of complex constraints like those in (94), which would be needed in a purely tonal approach, can be avoided.

(94) a. *H/$C_{[+voi]}V_{[+high]}$ / __$CV_{[-high]}$
*H/$C_{[+son]}V_{[+high]}$ / __$CV_{[-high]}$
b. *H/$C_{[+voi]}V_{[+high]}$ / __#
*H/$C_{[+son]}V_{[+high]}$ /__#

A purely tonal approach based on constraints like those in (94) would be uninsightful, because not only is it necessary to refer to both [+voi] and [+son], but it also fails to explain why the environments in (94a) and (94b) trigger the same tonal patterns. In the analysis proposed here, on the other hand, this environment can be formulated as "a mora of the shape *GI* in the non-head part of a foot".

5.6. Devoicing of high vowels followed by non-high vowels

In this chapter, we have focused on the special behavior of syllables with high vowels followed by syllables with non-high vowels in terms of H tone association. However, there is data from vowel devoicing that suggests that the prosodic "weakness" of syllables with high vowels followed by syllables with non-high vowels is of a more general nature. In a number of Tohoku dialects, only high vowels followed by non-high vowels devoice (Inoue 1968; Ōhashi 2002; Saito 1992). The data in (95) is from the Ichinoseki dialect spoken in Iwate Prefecture (adapted from Saito 1992). The condition of devoicing is that the high vowel is preceded by a voiceless onset and followed by a syllable with a voiceless onset and a non-high vowel.

[150] If the relevant constraint would turn out to be ACCENT-TO-HEAD can be defined as follows: the left edge of an accentual H tone is aligned with the left edge of a foot-head. Note that the winning candidate does violate ACCENT-TO-HEAD if this constraint is defined in terms of a diacritic accent.

[151] Another consequence of the proposed analysis is that in words like *asi*ˀ 'foot', the accentual H tone is not left-aligned with the head of a foot. In words like this, spreading is not necessary because the second syllable has an onset with the feature [-voi], as is the case in unaccented words like *iki* 'breath' in (86).

(95)	*Input*	*Output*	*Gloss*	*Comments*
a.	hito	ɸʲïto	‘person’	Devoicing of *CI.CA*
b.	kutu	kɯ̈ᵈzɯ̈	‘mouth’	No devoicing of *CI.CI*

In (95a), *CI* is followed by *CA*, the result of which is devoicing of the high vowel. This in turn results in the failure of intervocalic voicing to occur. In (96b), on the other hand, *CI* is followed by another *CI*, and therefore does not devoice. In the absence of a devoiced vowel, intervocalic voicing gets the chance to apply.

The data in (95) is interesting, but, as pointed out by Inoue (1968), there are exceptions to the intervocalic voicing rule. This would imply that whether an intervocalic obstruent is voiced or not is part of the lexical entry. If so, the absence of devoicing can be explained by the fact that the vowel is not flanked by two voiceless consonants, which is a necessary condition for high vowel devoicing. In other words, the different contexts in terms of vowel height may not necessarily be relevant in the synchronic grammar.

On the other hand, even if intervocalic voicing is lexicalized in Tohoku Japanese, we might still prefer a grammar that reflects the distributional properties of words in terms of surface constraints. In that sense, the argument that the alternations in (95) are not synchronic is not a very strong one. Also, even in other dialects, if only just a tendency, high vowels are more easily devoiced before non-high vowels than before high vowels (Byun 2012). The data from vowel devoicing thus suggests that *CI* followed by *CA* (i.e. *CI.CA*) is prosodically weak in a more general sense. Therefore, a careful comparative study of the environments in which H tones are avoided in dialects with tone-vowel height interaction and the environments in which high devoicing is claimed to be sensitive to relative sonority is an important topic for further research.

5.7. Summary

In this chapter, I have shown that H tones interact directly with vowel height in terms of alignment constraints. Alhough H tones were shown to be sensitive to certain prosodically prominent positions, it proved to be necessary to make use of tone-vowel height alignment constraints.

In all three dialects, foot structure was shown to play a role in the assignment of phrasal tones. In Matsue Japanese, phrase-initial tonal lapses of more than one foot are avoided. In Ichihara Japanese, a constraint that favors initial trochees was shown to cause the accent to shift from the final syllable in disyllabic words to a following particle with a non-high vowel. In Kanazawa Japanese, foot structure was shown to play an important role in accounting for the phrasal tone patterns. Alignment constraints that align H tone

or non-high vowels with foot-heads were shown to be necessary. Also, monomoraic feet were shown to play a crucial role in the tonal system.

Apart from the evidence for a role of metrical structure in the tonal grammar, constraints that refer to tone play an important role in these dialects. What is more, because tone-vowel interaction can be observed both in lexical processes as well in phrasal contexts, we can be sure that the constraints regulating the interaction are active throughout the phonological grammar. This is evidence for an approach to pitch accent in which tone is active lexically, regardess of the underlying representation in the lexicon.

6. Tonal patterns in verbs: prosodic constituency and surface correspondence

Verbal paradigms are potentially an interesting area for investigating the interaction between phonology and morphology. This also holds for verbal paradigms in Japanese, in which – depending on the dialect – considerably complex accentual alternations can be observed.

As discussed in Bermúdez-Otero (2011) and Scheer (2011), it is possible to distinguish between two different types of morpho-syntactic conditioning in phonology: representational and procedural. Bermúdez-Otero (2011) summarizes the differences in the following way. In representational morpho-syntactic conditioning, morphological boundaries or prosodic constituents are projected from morphological constituents.[152] In procedural morpho-syntactic conditioning, on the other hand, the structure that is visible to the phonological grammar is controlled by the morpho-syntactic grammar, by means of serial interaction with the phonological grammar (as in cyclic models), or by making paradigmatically related forms available to the phonology. The distinction between the two approaches to procedural effects is based on the distinction between derivational and non-derivational models. In derivational models, phonology applies after each round of morphology, as in Lexical Phonology (Kiparsky 1982) and Stratal OT (Bermúdez-Otero 2012). In non-derivational models, on the other hand, reference is made to "uniform exponence" (Kenstowicz 1996) or "paradigm uniformity" (Steriade 2000).

Raffelsiefen (2005) stresses the importance of distinguishing between representational "boundary effects" and "paradigm uniformity" (PU) effects. In English, not only clues from stress, but also segmental clues like aspiration, lenition, and vowel reduction can be used in determining the prosodic structure of words. However, Japanese differs from English in that there are not many clues other than the location of accent and the presence or absence of geminates – the presence of which is, as mentioned in Chapter 2 (2.3.2), evidence for the absence of a p-word boundary (Ito and Mester 1996) – that can help us to determine the possible prosodic domains in a word or phrase. Therefore, distinguishing between boundary effects and PU effects in Japanese can be expected to involve even more subtle kinds of evidence than in English.

In this chapter, the relative roles of boundary effects or "domain effects" on the one hand, and paradigm uniformity on the other hand are discussed. It will be shown that the p-stem is necessary to account for the accent patterns in verbs across dialects. Also, evidence will be provided in favor of a paradigm uniformity approach as opposed to a cyclic approach.

[152] See Scheer (2011) for an alternative approach in his CVCV-model.

6.1. Tokyo Japanese verbal accent

6.1.1. Data and issues

Like other topics in Japanese phonology, most theoretically-oriented studies on Japanese verbal accent have focused on Tokyo Japanese.[153] Still, verbal accent is relatively understudied compared to for instance nominal compound accentuation.

In terms of accentedness, there are two classes of verbal roots in Tokyo Japanese: accented roots, and unaccented ones. Furthermore, in terms of segmental make-up, there are again two types of roots: consonant-final roots (C-verbs) and vowel-final roots (V-verbs). We will start with the difference in accent patterns between accented C-verbs and accented V-verbs, and come back to the accented vs. unaccented distinction later.

Exampes of verbal forms of both types of accented roots are provided in (1).[154] The number between parentheses indicates the accented location counted in moras from the right edge of the root (ignoring the root-final consonant of C-verbs).

(1) Verbs with accented roots

Exponence	Morpheme	C-verbs *syaber-* 'chat'	V-verbs *tabe-* 'eat'
nonpast	-u/-ru	sya.b**e**ˀ.r-u (-1)	ta.b**e**ˀ.-ru (-1)
imperative	-e/-ro	sya.b**e**ˀ.r-e (-1)	ta.b**e**ˀ.-ro (-1)
provisional	-eba/-reba	sya.b**e**ˀ.r-e.ba (-1)	ta.b**e**ˀ.-re.ba (-1)
past	-ta	sya.b**e**ˀQ.-ta (-2)	t**a**ˀ.be.-ta (-2)
gerund	-te	sya.b**e**ˀQ.-te (-2)	t**a**ˀ.be.-te (-2)
conditional	-tara	sya.b**e**ˀQ.-tara (-2)	t**a**ˀ.be.-ta.ra (-2)

As we can see in (1), if we count from the left of the word, the accent of nonpast and past forms in C-verbs (i.e. forms with a root ending in a consonant) is located on the same mora. However, when counting from the right edge of the word, the accent is located on different moras. In V-verbs (i.e. verbs with a root ending in a vowel), the accent is not located on the same mora, either counting from the left, or counting from the right. Nevertheless, a clear pattern is visible: in nonpast forms, the accent is located on the final mora of the root, whereas in past forms it falls on the penultimate mora of the root.

The difference between the "non-past group" and the "past group" has been explained in quite a number of different ways in the literature. McCawley (1968) posits a basic stem-penultimate accent and a rule of accent attraction for the nonpast, imperative, and provisional forms. Martin (1975), Okuda (1980), and Hayata (1997) propose similar rules.

[153] Section 6.1 is partly based on Poppe (2012).

[154] The data in this chapter are based on the forms listed in the accent dictionaries NHK (1998) and Kindaichi and Akinaga (2001), as well as on data that I elicited from native speakers.

However, based on the observation that if you know whether a root is accented or unaccented, the location of accent is predictable, many scholars assume that the location of verbal accent is computed by rule.[155] Concretely, this means that accented verbs have a feature [+accented] (a 'floating accent') or a floating tone (H or HL), whereas unaccented verbs lack such a specification. Advocates of an analysis with a floating accent include Kitagawa (1986), Haraguchi (1991), Kubozono (2008), Nishiyama (2010), and Yamaguchi (2010a/b), among others. However, there is no consensus with respect to the question how the accent location is computed.

In verbs with unaccented roots, an accent occurs on the suffix in the provisional and conditional forms (2).[156] Furthermore, as we will see below, in certain constructions an accent may appear on the monomoraic suffixes as well (e.g. *aketa*ˀ*-ga* 'open-PAST-but').

(2) Verbs with unaccented roots

Exponence	Morpheme	V-verb *ake-* 'open (tr.)'
nonpast	-u/-ru	a.ke.ru
imperative	-e/-ro	a.ke.ro
provisional	-eba/-reba	a.ke.r**e**ˀ.ba
past	-ta	a.ke.-ta
gerund	-te	a.ke.-te
conditional	-tara	a.ke.-t**a**ˀ.ra

There seems to be a general consensus that the accent in the provisional and conditional forms are part of the suffix. In Poppe (2012), I proposed that these accents are inserted by the grammar. Below I will argue, however, that it is better to assume they are indeed part of the underlying specification of the suffix. However, I do follow the proposal made in Poppe (2012) that the accent that can be observed in certain verb forms with unaccented roots followed by multiple suffixes or particles is a default boundary-marking accent (see below in (23)).

6.1.2. Previous analyses

Kitagawa (1986) and Zamma (1992) propose an analysis in terms of level ordering, in which the concatenation of the nonpast marker *-(r)u* precedes accent assignment, and the concatenation of the past marker *-ta* is ordered after accent asignment. Their approach can be summarized as in (3). The basic accent assignment rule in this analysis associates

[155] As I mentioned earlier in this thesis, the same is true for adjectives, which in the interest of space are not discussed in any detail in this thesis. See McCawley (1968) and Zamma (1992) for more details.
[156] The distinction between C-roots and V-roots is not relevant for unaccented verbs.

the accent with the penultimate (-2) mora.

(3) Level-ordering (see Kitagawa 1986, Zamma 1992)

		continuative	*nonpast*	*past*
Lexicon		tabe [+acc]	tabe [+acc]	tabe [+acc]
Level 1	morphology	-	tabe-ru	-
	accent rule (-2)	taˈbe	tabeˈ-ru	taˈbe
Level 2	morphology	-	-	taˈbe-ta
Output		taˈbe	tabeˈru	taˈbeta

This analysis in the spirit of Lexical Phonology (Kiparsky 1982) works,[157] but if possible, we would like to do without level ordering. In 7.4, empirical arguments from the Matsue dialect against an approach based on level-ordering will be presented.

Haraguchi (1991) and Nishiyama (2010) also assume a basic penultimate accent. Instead of making use of level ordering, their analyses make use of extrametricality: the past suffix *-ta* is invisible for rules of accent assignment. However, it turns out that once we consider the suffixes like *-tara* and *-(r)eba*, the approach based on extrametricality only works if we assume extrametricality only for verbs with accented roots. This is illustrated in (4), which is based on the analysis in Nishiyama (2010).

(4)

Accented verbs	*Unaccented verbs*
tabeˈru	akeru
tabeˈre<ba>	aker**é**ba
taˈbe<ta>	aketa
taˈbe<tara>	aket**á**ra

Nishiyama (2010) introduces the notion of "selective extrametricality" to explain how the provisional conditional forms could get an accent on the penultimate mora. The idea is that particular suffixes are visible to processes related to metrical and prosodic structure in particular cases only. The notion of selective extrametricality has a number of problems. First, extrametricality is nothing but a stipulation, and does not explain anything. Second, the reason why extrametricality only applies to verbs with accented roots remains unclear, as Nishiyama (2010) admits.

Drawing a parallel with compound accentuation, Kubozono (2008) proposes that the

[157] Still, as Zamma (1992) points out, a complete analysis of all verbal and adjectival accent seems to require the need for certain suffixes to belong to different levels depending on the construction.

accent of accented verbs is aligned with the morpheme boundary. The problem with this proposal, however, is that in the forms *ta˺be-ta* (**tabe˺-ta*) 'eat-PAST' and *ta˺be-tara* (**tabe˺-tara*) 'eat-COND' the accent is not aligned with the morpheme boundary.

Building on the proposal by Kubozono (2008), Yamaguchi (2010a/b) demonstrates that *ta˺be-ta* (**tabe˺-ta*) and *ta˺be-tara* (**tabe˺-tara*) can be explained in terms of paradigm uniformity. The analysis is schematically summarized in (5).

(5)

-u/-ru	-ta
sya.b**e**˺.r-u (-1) ↕ PU ta.b**e**˺.-ru (-1)	sya.b**e**˺t.-ta (-2); *sya.be**t**˺.-ta (-1) ↕ PU t**a**˺.be.-ta (-2); (*tabe˺ta (-1))

The analysis can be described as follows. In the ideal case, the accent is assigned to the rightmost mora of the root, for which Yamaguchi (2010b: 6) proposes the constraint in (6).

(6) ALIGN-R (accent, root)
Align the right edge of the accented mora with the right edge of the root.
("Assign one violation mark for every mora which stands between μ_h-σ˺ and the right edge of the root"; Yamaguchi 2010b:6).

In the past form of C-verbs, root-final accent is impossible because it would result in an accent on the first half of a geminate (**syabeQ˺-ta*). An accent in a marked location like this is prevented by the constraint in (7).

(7) *CVX˺
The accent is not located on the right branch of a rhyme.
("Assign one violation mark for every element on the right branch of a rhyme […] which is the head mora of the accented syllable") (Yamaguchi 2010b: 6).

In order to avoid an accent on *Q*, the accent is assigned one mora to the left of the rightmost mora in the root, resulting in *syabe˺Qta*. Finally, a constraint on "uniform exponence" (UE; Kenstowicz 1996; van de Weijer 1999), causes the accent in past forms of V-verbs to be associated to the penultimate mora instead of the final mora, even though in V-verbs there is no root allomorph with a final *Q*. Both the informal and formal definition in (8) are taken from Yamaguchi (2010b: 5-6).

(8) UNIFORM EXPONENCE (UE)-AFFIX:
"An […] affix […] has the same realization for property P in its various contexts of occurrence." (Yamaguchi 2010: 5)
("Assign one violation mark for every pair of {V-final root + suffix α, C-final root + suffix α} where the two forms have the head mora of the accented syllables (μ_h-σ') in different positions […]") (Yamaguchi 2010b: 6)

As shown by Yamaguchi (2010b), under the ranking *CVX' >> UE-AFFIX >> ALIGN-R(accent, root), the past forms come out as in (9).[158]

(9)

tabe-[acc] -ta, syaber[acc] -ta	*CVX'	UE-AFFIX	ALIGN-R
a. {ta.be'.ta, sya.bet'.ta}	*!		
→b. {ta'.be.ta, sya.be't.ta}			**
c. {ta.be'.ta, sya.be't.ta}		* !	*

As for verbs with unaccented roots, Yamaguchi (2010b) shows that by adopting a constraint that protects roots from changes in its status as accented or unaccented (FAITH-ACCENT$_{Root}$), the accent patterns of unaccented verbs can also be accounted for. The conditional suffix *-tara* has a floating accent in Yamaguchi's (2010b) analysis.

(10)

ake-tara[acc], syaber-tara[acc]	FAITH-ACC$_{Root}$	UE-AFFIX	ALIGN-R
a. {a'.ke.ta, sya.be't.tara}	*!		
b. {a.ke.ta'ra, sya.bet.ta'ra}	*!		**
→c. {a.ke.ta'ra, sya.be't.tara}		*	*

While Yamaguchi's (2010b) analysis is insightful, there are a number of issues that remain unresolved. As pointed out by Yamaguchi (2010b) herself, her analysis makes the wrong predictions for verbs with roots that end in a long vowel or a diphthong and a final consonant, such as *to'or-u* 'pass-NONPAST' in (11).

[158] I have taken the liberty of making some changes in the lexemes used in the input and output forms, as well as the format of the tableau. The ranking of *CVX' above UE-Affix can be confirmed by comparing pairs like {*mi'ta, syabe'tta*}, where although in the first form only final accent is possible, in the second form the accent does not fall on the first half of a geminate.

(11)

-u/-ru	-ta
sya.b**e**ˀ.r-u, tabeˀ-ru (-1) ↕ No PU t**o**ˀo.r-u (-2); *to**o**ˀ.r-u (-1)	sya.b**e**ˀt.-ta (-2); taˀ.be.-ta (-2) ↕ No PU t**o**ˀo.t.-ta (-3); *to**o**ˀt.-ta (-2)

A second problem is that it is not clear what forms would serve as the analogical 'base' in longer verbal complexes consisting of more than two morphemes. For example, it is counterintuitive that the accent location in a form like *tabe-sase-ráre-ta* 'was forced to eat' is computed by comparing this form to e.g. *syabeˀt-ta*.

Finally, a problem involving the Indirect Reference Hypothesis that is assumed in this thesis is that ALIGN-R(accent, root) refers to a phonological feature and a morpho-syntactic constituent, rather than a prosodic constituent and a morpho-syntactic one. It is this third problem that we will take up first, coming back to the problem of paradigm uniformity later.

6.1.3. The role of the p-stem in verbal accentuation

By adopting the Indirect Reference Hypothesis, we are forced to define the domain of accent assignment in terms of prosodic constituents. As all verb forms involve suffixation, we may redefine the root as a stem. Therefore, our main candidate for the domain of accent assignment of underlying accents in verbs is the p-stem.

Let us hypothesize that in verbs with accented roots the accent is consistently aligned within the prosodic domain across all forms. Under this hypothesis, p-stems would have the following forms.

(12)

	C-verbs	*V-verbs*
a. nonpast	[syabeˀru]$_{\psi}$	[tabeˀru]$_{\psi}$
b. provisional	[[syabeˀre]$_{\psi}$ ba]	[[tabeˀre]$_{\psi}$ ba]
c. past	[[syabeˀt]$_{\psi}$ ta]	[[taˀbe]$_{\psi}$ ta]
d. conditional	[[syabeˀt]$_{\psi}$ tara]	[[taˀbe]$_{\psi}$ tara]

In all forms, the accentual H tone falls on the penultimate mora. This fact could be the result of the alignment of the HL complex as such with the p-stem boundary, or the alignment of a trochaic foot to the right of the p-stem. Here I adopt the latter option, for reasons that will become clear later. Apart from a constraints TROCHEE and ACCENT-TO-HEAD, which together force the accent to be realized in the head of a trochee, I adopt the constraints in (13a) and (13b). The constraint ALIGN-R(ψ, F) (13a) forces a foot to be right-aligned with the p-stem (13a), and according to the constraint RIGHTMOST(ψ) (13b)

the rightmost foot in the p-stem must be accented (13b). Whnever it is not necessary to refer to foot structure within the analysis, I use the cover constraint PENULT-ACCENT(ψ) in (13c).[159]

(13) a. ALIGN(P-stem, Right, Foot, Right) = ALIGN-R(ψ, F)
The right edge of a p-stem is aligned with the right edge of a foot.
b. RIGHTMOST(ψ)
An accent is associated to the rightmost foot of the p-stem.
c. PENULT-ACCENT(ψ) = PENULT(ψ) (cover constraint)
The accent in a p-stem is associated with the penultimate mora.

Returning to the prosodic structure in (12), the inner domain in (12c) and (12d) ends in the first half of a geminate, which shows that this inner domain indeed cannot be a p-word, and thus must be a p-stem. The mapping from m-stem to p-stem can be captured by the constraint in (14), which can be seen as a combination of ALIGN-LEFT(M-Stem, P-stem) and ALIGN-RIGHT(M-Stem, P-stem) (Downing 1998a/b, 1999).

(14) M-STEM ≈ P-STEM
The left and right edges of an m-stem coincide with the left and right edges of a p-stem.
(One violation is assigned for each segment between the boundaries of the p-stem and the m-stem)

The question that must be answered now is why the p-stem boundaries are located where they are in (12).

One of the major arguments for prosodic constituents that correspond to morpho-syntactic ones is that in many cases mismatches can be observed between morpho-syntactic constituents and the constituents relevant to phonological computation. While a perfect match between the two types of stems is the ideal, a mismatch occurs when phonotactic constraints prohibit this. In the case of m-stem to p-stem mapping, we are dealing with exactly such a mismatch. In (Tokyo) Japanese, no prosodic unit may end in a consonant other than the moraic nasal /N/ or the first half of a geminate /Q/, which

[159] It will be shown below that in certain cases, the rightmost foot in the p-stem does not result in penultimate accent. Note that the generalization that p-stems in the default case have penultimate accent is not in contradiction with the default antepenultimate accent in loanwords, etc. The reason for this is that loanwords are simplex words, and thus lack a morpho-syntactic stem that maps onto a p-stem.

means the stem of a C-verb cannot map onto a p-stem without violating phonotactic constraints. This means that if the mismatch is minimalized, there are two logically possible ways in which the morphological structure of a form like *syaber-u* can be translated into phonological structure: [[syabe]ru] or [syaberu]. In the first parsing, the morphologically complex nature of the form is mirrored in the phonology, albeit not in a perfect way. In this parsing, not including material in the p-stem that does not belong to the m-stem (the root-final /r/) is more important than parsing all morphological material into the p-stem. In the second parsing, it is the other way around: parsing at least all m-stem material into the p-stem is more highly valued than not parsing material into the p-stem that belongs to other morphemes (the /u/ of which the non-past allomorph consists). Thus, there are two constraints with conflicting interests: one that says that the p-stem may not be larger than the m-stem, and one that says it may not be smaller than the m-stem. These constraints are the m-stem~p-stem correspondence constraints in (15), which are adapted from Downing (1998a/b, 1999); see also Kim (1997, 2010).

(15) DEP(M-P):
Every element of the p-stem has a correspondent in the m-stem.
(One violation for each segment in the p-stem that is not part of the m-stem)
MAX(M-P):
Every element of the m-stem has a correspondent in the p-stem.
(One violation for each segment in the m-stem that is not part of the p-stem)

When DEP(M-P) outranks MAX(M-P), in the case of a mismatch the p-stem will be smaller than the m-stem. On the other hand, when MAX(M-P) outranks DEP(M-P), the p-stem will be larger than the m-stem when there is a mismatch. Obviously, Tokyo Japanese must adopt the latter ranking, for else we would expect the form *[[syaˀbe]ru], which is not attested in Tokyo Japanese (but which is attested in Shizuoka Japanese, as we will see below).

Now that we can account for the incorporation of material from the affix into the p-stem, we can take up the question why the constraint DEP(M-P) is violated in V-verbs even though perfect alignment and thereby satisfaction of this constraint as well as M-STEM ≈ P-STEM is an option.[160] The answer to this issue must lie in paradigm uniformity:

[160] In the Japanese data the right edges of the m-stem and the p-stem may show a mismatch, but their left edges always coincide. This means that if we split M-STEM ≈ P-STEM, ALIGN-L(M-Stem, P-stem) can be assumed to be top-ranked, with ALIGN-R(M-Stem, P-stem) having the same (or a lower) ranking as DEP(M-P). Actually, reference to ALIGN-R(M-Stem, P-stem) is not necessary because the two correspondence constraints together decide how much material will be included in the p-stem.

if a certain affix is (partly) incorporated into the p-stem in at least one output form, it must be incorporated in the same way in other output forms that correspond to the same morphological structure. This idea can be translated into the "uniform alignment" constraint[161] in (16a). In (16b), a concrete example is given.

(16)a. UNIFORM ALIGNMENT-RIGHT(ψ) = UA-R(ψ, ψ)
In two output forms containing the affix *α*, align the right edge of the p-stem to the right edge of the p-stem.

b.	[syaber] + (r)u	↔	[syaberu]	
	↕		↕	UA-R
	[tabe] + (r)u	↔	[taberu]	
			*[tabe]ru	

This constraint can be seen as a paradigm uniformity constraint that demands that an affix is incorporated in the same way within the p-stem every time it occurs as an output form. In other words, it checks the alignment between the edges of prosodic domains in output forms that are paradigmatically related. In the case at hand, the constraint outranks M-STEM ≈ P-STEM, resulting in a single p-stem in the non-past forms of V-verbs.

Alternatively, we could assume that certain morphemes are subcategorized with a p-stem specification. One possible way to do this is to include the subcategorization in the lexical entry of the form, as in (17) (see Inkelas 1989/1990).

(17) a.	nonpast	$[X]_{Stem}$ + (r)u	→	$[X(r)u]_{\psi}$
b.	imperative	$[X]_{Stem}$ + ro	→	$[X(r)o]_{\psi}$
		$[X]_{Stem}$ + e	→	$[Xe]_{\psi}$
c.	provisional	$[X]_{Stem}$ + (r)eba	→	$[X(r)e]_{\psi}$ ba

Formalizing this subcategorization approach in terms of morpheme-specific alignment constraints is a less attractive option. First, it is not clear how to formulate an alignment constraint for the provisional suffix *-(r)eba*, of which only the first vowel is included in

[161] As this constraint refers to the alignment of two prosodic categories, it is not in conflict with the Indirect Reference Hypothesis. In a parallel architecture, it will be clear from the output form whether two forms share a morpheme. The uniform alignment constraint is similar to Krämer's (2009) OO-ANCHOR-PCat constraint according to which "[t]he edges of a prosodic unit (syllable, foot) coincide in corresponding representations R_1 and R_2" (Krämer 2009: 229). As shown by McCarthy and Prince (1995, 1999), anchoring can be seen as the correspondence of elements adjacent to some edge.

the p-stem. A possible solution to this problem would be to analyze *-(r)eba* as consisting of a stem-forming suffix *-e*, and an inflectional suffix *-ba*, as in the traditional school of Japanese grammar.[162] However, even if we adopt this analysis, another problem for the morpheme-specific alignment-based subcategorization approach is that a constraint like that in (18) violates the strong version of the Indirect Reference Hypothesis (Bermúdez-Otero 2012), in which interface alignment constraints may only refer to M-Cat and P-Cat, and not to specific morphemes.

(18) ALIGN(-*(r)u*]$_{\text{Affix}}$, Right, p-stem, Right):
The right edge of the nonpast suffix is aligned with the right edge of a p-stem.

In the interest of the Indirect Reference Hypothesis, and the general constraint-based approach taken in this thesis, I adopt the constraint ALIGN-R(ψ, ψ).

The tableau in (19) shows how the most important constraints discussed are responsible for the p-stems appearing in nonpast and past forms of accented verbs. Because the exact ranking of the cover constraint PENULT(ψ) is not relevant here, it is placed outside the ranking of the first three constraints.[163] The ranking of MAX(M-P) and UA-R(ψ, ψ) above DEP (M-P) is of crucial importance in Tokyo Japanese, penalizing the forms in (19c) and (19f) respectively. The candidate consisting of the forms with the least number of violations of DEP (M-P) that does not violate the uniform alignment constraint UA-R(ψ, ψ) and the accent placement cover constraint (19e) is selected as the winner. Note that candidates with an accent on a moraic obstruent are not considered.

[162] See Vance (1987) and Shibatani (1990) for a discusson of traditional Japanese approaches to the Japanese verbal paradigm.

[163] In all dialects considered in this chapter, RIGHTMOST(ψ), which is also satisfied by a p-stem-final monomoraic or trimoraic foot, can be assumed to be undominated. The constraint ALIGN-R(ψ, F), on the other hand, is violated in for instance past forms of C-verbs in the Ichihara dialect (see (36b) below) and certain past forms of C-verbs in the Matsue dialect (see (65d) below). ACCENT-TO-HEAD is not violated in the dialects discussed in this chapter. Finally, TROCHEE outranks IAMB in all dialects except Maisaka Japanese.

(19) Verbal accent in Tokyo Japanese

tabe-[acc] -ru,-ta tukur-[acc] -u, -ta	MAX (M-P)	UA-R (ψ, ψ)	DEP (M-P)	PENULT (ψ)
a. [ta.beˀ.ru], [ta.beˀ.ta] [tu.kuˀ.ru], [tu.kuˀt.ta]			**,**(!), *, **	 *(!)
b. [taˀ.be.ru], [taˀ.be].ta [tuˀ.ku.ru], [tu.kuˀt].ta			**, *	*! *
c. [taˀ.be.]ru, [taˀ.be].ta [tu.kuˀ.ru], [tu.kuˀt].ta		*!	 *	
d. [ta.beˀ.ru], [ta.beˀ.].ta [tu.kuˀ.ru], [tu.kuˀt].ta			**, *	*!
→e. [ta.beˀ.ru], [taˀ.be].ta [tu.kuˀ.ru], [tu.kuˀt].ta			**, *	
f. [taˀ.be.]ru, [taˀ.be].ta [tuˀ.ku].ru, [tu.kuˀt].ta	*!			

We will now take a look at independent evidence for the proposed analysis from verbs with unaccented roots.

The data in (20) consists of contracted forms in which the gerund marker *-te* is followed by an inflected form of the animate existential root *i-*. Martin (1975: 514-517) gives a whole list of contracted forms, but as he does not give a full list of concrete examples, I collected actual examples of fully realized forms and corresponding contracted forms myself. The forms (20) belong to the paradigm of the root *ak-* 'open (intr.)', of which the stem *ai-* is an allomorph.

(20)	*Input*	*Fully realized forms*	*Contracted forms*	
a.	ai-teˀ ǀ i-teˀ	ai-te i-ta^0	ai-t**e**ˀ-ta	*ai-te-ta^0
b.	ai-teˀ ǀ i-taˀra	ai-te i-taˀra	ai-t**e**ˀ-tara	*ai-te-taˀra
c.	ai-teˀ ǀ i-reˀba	ai-te i-reˀba	ai-te-reˀba	
d.	ai-teˀ ǀ i-ruˀ	ai-te i-ru^0	ai-te-ru^0	

According to Martin (1975), all monomoraic particles have an underlying accent. He claims that this accent is suppressed before a juncture (indicated by a vertical bar "ǀ" in (20)) as in the "fully realized forms" in (20a/b/c/d/) and the contracted forms in (20c/d). Martin (1975) treats the contracted forms in (20a) and (20b) as exceptions, because "[n]ormally the oxytonic accent of the gerund of an atonic verb will be suppressed by the underlying juncture that separates the gerund from the auxiliary" (Martin 1975: 514-515). However, the question remains why the forms in (20a) and (20b) are exceptions.

Note that the different accent patterns of the contracted forms of unaccented words

cannot be explained if we assume the monomoraic suffixes are underyingly unaccented (21). The contracted forms in (21c/d) are as we would expect them to be based on the input forms. However, an accent appears on the final mora of the gerund in (20a/b), in which the underlying accent is even ignored in the latter form. It is not clear where the accent comes from.

(21)		*Input*		*Fully realized forms*	*Contracted forms*	
	a.	ai-te	i-te	ai-te i-ta^{0}	ai-t**e**ˀ-ta	*ai-te-ta^{0}
	b.	ai-te	i-taˀra	ai-te i-taˀra	ai-t**e**ˀ-tara	*ai-te-taˀra
	c.	ai-te	i-reˀba	ai-te i-reˀba	ai-te-reˀba	
	d.	ai-te	i-ru	ai-te i-ru^{0}	ai-te-ru^{0}	

Let us now see what an approach in which p-stems are subject to paradigm uniformity or prosodic subcategorization has to offer. In (22), the prosodic structure according to an analysis based on paradigm uniformity is shown.[164] The examples are inflected forms of the verb *ake-ru* ‘to open (tr.)’, the transitive counterpart of *ak-u*.

(22)	*Regular inflections*		*Contracted forms of gerund+ i-*
a.	[[ake]$_{\psi}$ ta^{0}]$_{\omega}$	↔	[[[ake]$_{\psi}$ teˀ]$_{\omega}$ ta]$_{\omega}$
b.	[[ake]$_{\psi}$ taˀra]$_{\omega}$	↔	[[[ake]$_{\psi}$ teˀ]$_{\omega}$ tara]$_{\omega}$
c.	[[akeru^{0}]$_{\psi}$]$_{\omega}$	↔	[[[ake]$_{\psi}$ te ru^{0}]$_{\psi}$]$_{\omega}$
d.	[[ereˀ]$_{\psi}$ ba]$_{\omega}$	↔	[[[ake]$_{\psi}$ te reˀ]$_{\psi}$ ba]$_{\omega}$

As the contracted forms in (22a) and (22b) show, a single p-stem that corresponds to the m-stem *ake-* is assigned, with the following suffix closing the first p-word. This p-word is the head of a recursive p-word that is closed by a second suffix. The recursive p-word triggers the insertion of an accent, which is assigned to the innermost p-word. This junctural accent overrides the underlying accent of the suffix in (22b). In (22c) and (22d), the UA-R(ψ, ψ) constraint triggers the alignment of the p-stem with the right edge of the suffix *-(r)u* and the right edge of the first syllable of *-(r)eba*. The result of this is that the string *ake-te* which is contained in *ake-te-ru* and *ake-te-reˀba* cannot be mapped onto a p-word, as a p-word cannot be dominated by a p-stem. Therefore, no recursive p-word

[164] The contracted forms in (22c) and (22d) are analyzed as containing a recursive p-stem.The reason for this is that in accented verbs, the accent must be located in innermost p-stem, as in the form [[[taˀbe]$_{\psi}$ te ru^{0}]$_{\psi}$]$_{\omega}$ ‘eat-GER-NONPAST’.

arises, and hence no junctural accent is inserted. The same holds for (22d), in which the underying accent of the suffix can be faithfully realized.

Before we move on to take a look at the same constructions in other dialects, it is interesting to note that the default accent is also inserted when inflected forms ending in monomoraic suffixes are followed by particles, as in the constructions in (23).

(23) a. ake-ruˀ-ga ‘[I] open [it] but’
ake-taˀ-ga ‘[I] opened [it] but’
b. ake-ruˀ-kara ‘because [I] open’
ake-teˀ-kara ‘after opening’

The accent on the suffixes in (23) can also be treated as a default boundary marker. Because the default accent only appears in recursive p-words, this means that a particle is incorporated into a recursive p-word (24), rather than being directly adjoined to the phrasal node.

(24) a. $[[\text{ake-ru}ˀ]_{\omega}\ \text{ga}]_{\omega}$
$[[\text{ake-ta}ˀ]_{\omega}\ \text{ga}]_{\omega}$
b. $[[\text{ake-ru}ˀ]_{\omega}\ \text{kara}]_{\omega}$
$[[\text{ake-te}ˀ]_{\omega}\ \text{kara}]_{\omega}$

Note that not all particles may be parsed into a recursive p-word with the preceding verb. For instance, the pragmatic marker *ne* does not trigger the insertion of an accent, which means that this particle is not parsed into a recursive p-word headed by the verb.

One may wonder why particles that follow verbs may be incorporated into a recursive p-word, whereas this is generally not allowed in noun-particle constructions (except for dominant suffixes; see Chapter 2). I have no conclusive answer to this question, but I suspect it is related to differences in the ability of constructions to undergo grammaticalization. Suffix-particle combinations may grammaticalize more easily than noun-particle combinations, which means that it is easier for particles to “enter” the domain of the verb to which the suffix belongs than to enter the domain of a noun. Related to this, faithfulness constraints specific to nouns may be higher ranked than faithfulness constraint specific to verbs, because verbs appear in different constructions all the time.

Summarizing, the ability to account for accentual alternations in regular inflections and in contracted gerund-existential constructions is evidence for the proposal that the p-stem plays a role in the accentual alternations in the verbal paradigm. No other approach

is able to account for these alternations in a straightforward way. What is more, the analysis makes the right micro-typological predictions, as we will see now.

6.1.4. The role of the p-stem in Shizuoka and Ichihara Japanese

I already mentioned that in the Shizuoka dialect (the dialect of Shizuoka City), the nonpast form of accented V-verbs has an accent on the penultimate mora of the stem. The data in (25) is taken from Yamaguchi (2010a). Similar data can be found in Hori and Takigawa (2014).

(25)		Shizuoka Japanese	Tokyo Japanese	
a.	V-Verbs:	taꜜbe-ru	tabeꜜ-ru	'eat-NONPAST'
		taꜜbe-ta	taꜜbe-ta	'eat-PAST'
b.	C-verbs:	tukuꜜr-u	tukuꜜr-u	'make-NONPAST'
		tukuꜜt-ta	tukuꜜt-ta	'make-PAST'

According to Yamaguchi (2010a), in Shizuoka the form *taꜜbe-ru* is favored above *tabeꜜ-ru* because it is possible to perfectly align a trochaic foot with the right edge of the root: *(taꜜbe)-ru*. In the non-past form of C-verbs, this is not possible, and thus the accent is aligned as closely as possible to the right edge of the root: *tu(kuꜜr-u)*. Thus, in Shizuoka Japanese, the paradigm uniformity constraint is not ranked high enough to enforce identity between *tukur-u* and *tabe-ru*, hence *{*tu(kuꜜ-ru), ta(beꜜ-ru)*} but {*(taꜜbe)ru, tu(kuꜜru)*}.

In the analysis based on the p-stem, on the other hand, the prosodic structure must be as in (26).[165]

(26)		Shizuoka Japanese	Tokyo Japanese	
a.	V-Verbs:	$[[taꜜbe]_{\psi}\ ru]_{\omega}$	$[[tabeꜜru]_{\psi}]_{\omega}$	'eat-NONPAST'
		$[[taꜜbe]_{\psi}\ ta]_{\omega}$	$[[taꜜbe]_{\psi}\ ta]_{\omega}$	'eat-PAST'
b.	C-verbs:	$[[tukuꜜru]_{\psi}]_{\omega}$	$[[tukuꜜru]_{\psi}]_{\omega}$	'make-NONPAST'
		$[[tukuꜜt]_{\psi}\ ta]_{\omega}$	$[[tukuꜜt]_{\psi}\ ta]_{\omega}$	'make-PAST'

The difference between the Shizuoka forms in (26a) and (26b) can be considered a result of the fact that DEP(M-P) is ranked at least as highly as UA-R (ψ, ψ). Because UA-R (ψ, ψ) does not outrank DEP(M-P), as it does in Tokyo Japanese, the candidates in (27a) and

[165] Below in (36) foot structure is added to these representations. For Shizuoka Japanese (and Tokyo Japanese) this foot structure is be the same as that proposed by Yamaguchi (2010a).

(27d), in which the accent location is identical to that of Tokyo Japanese, fail to surface. As in Tokyo Japanese, however, MAX(M, P) is top-ranked, excluding (27e). The winning candidate in (27b) beats that in (27c) due to the fact that the accent is penultimate in the p-stem in all forms in the former, but not in the latter.

(27) Verbal accent in Shizuoka Japanese

tabe-$_{[acc]}$ -ru,-ta tukur-$_{[acc]}$ -u, -ta	MAX (M-P)	DEP (M-P)	UA-R (ψ, ψ)	PENULT (ψ)
a. [ta.beˀ.ru], [ta.beˀ.ta] [tu.kuˀ.ru], [tu.kuˀt.ta]		**!,**, *, **		*
→b. [taˀ.be.]ru, [taˀ.be].ta [tu.kuˀ.ru], [tu.kuˀt].ta		 *	*(-*ru*)	
c. [ta.beˀ.]ru, [ta.beˀ.].ta [tu.kuˀ.ru], [tu.kuˀt].ta		 *	*(-*ru*)	*!
d. [ta.beˀ.ru], [taˀ.be].ta [tu.kuˀ.ru], [tu.kuˀt].ta		**!, *		
e. [taˀ.be.]ru, [taˀ.be].ta [tuˀ.ku].ru, [tu.kuˀt].ta	*!			

The prediction that this analysis makes is that in Shizuoka Japanese, the gerund form will project a p-word in the contracted gerund-existential constructions in all cases, with the result that in all forms a junctural accent will be inserted. According to Kyoko Yamaguchi (p.c., May 2012), a native speaker of the Shizuoka dialect, this is actually what happens. Thus, the relation between regular inflections and contracted gerund-existential forms is as in (28).

(28) *Regular inflections* *Contracted forms of gerund + i-* (Shizuoka)

a. $[[ake]_{\psi}\ ta^{0}]_{\omega}$ ↔ $[[[ake]_{\psi}\ teˀ]_{\psi}]_{\omega}\ ta]_{\omega}$

b. $[[ake]_{\psi}\ taˀra]_{\omega}$ ↔ $[[[[ake]_{\psi}\ teˀ]_{\psi}]_{\omega}\ tara]_{\omega}$

c. $[[ake]_{\psi}\ ru^{0}]_{\omega}$ ↔ $[[[[ake]_{\psi}\ teˀ]_{\psi}]_{\omega}\ ru]_{\omega}$

d. $[[ake]_{\psi}\ reˀ\ ba]_{\omega}$ ↔ $[[[[ake]_{\psi}\ teˀ]_{\psi}]_{\omega}\ re\ ba]_{\omega}$

Shizuoka Japanese thus provides us with additional evidence for the analysis based on the p-stem. Moreover, when we take a look at the data of Ichihara Japanese that I collected myself, the evidence becomes even more convincing.

The Ichihara Japanese forms in (29) all have the accent on the penultimate mora, except in the past form of C-verbs, in which moraic obstruent in the penultimate position is avoided by the accent.

(29)		Ichihara Japanese	Tokyo Japanese	
a.	V-Verbs:	tabeˀ-ru	tabeˀ-ru	‘eat-NONPAST’
		tabeˀ-ta	taˀbe-ta	‘eat-PAST’
b.	C-verbs:	tukuˀr-u	tukuˀr-u	‘make-NONPAST’
		tukuˀt-ta	tukuˀt-ta	‘make-PAST’

The penultimate accent in the past form of V-verbs suggests that the past suffix *-ta* is incorporated into the p-stem. Other evidence for the incorporation of the past suffix into the p-stem comes from tone-vowel height interaction in words with a stem-final high vowel, e.g. *miˀ-ru* ‘see-nonpast’ vs. *mi-taˀ* ‘see-past’.[166] Thus, the prosodic structure of verbs in Ichihara Japanese is as in (30).

(30)	Ichihara Japanese	Tokyo Japanese
a.	$[[\text{tabeˀru}]_{\psi}]_{\omega}$	$[[\text{tabeˀru}]_{\psi}]_{\omega}$
	$[[\text{tabeˀre}]_{\psi}\ \text{ba}]_{\omega}$	$[[\text{tabeˀre}]_{\psi}\ \text{ba}]_{\omega}$
	$[[\text{tabeˀta}]_{\psi}]_{\omega}$	$[[\text{taˀbe}]_{\psi}\ \text{ta}]_{\omega}$
	$[[\text{tabeˀta}]_{\psi}\ \text{ra}]_{\omega}$	$[[\text{taˀbe}]_{\psi}\ \text{tara}]_{\omega}$
b.	$[[\text{tukuˀru}]_{\psi}]_{\omega}$	$[[\text{tukuˀru}]_{\psi}]_{\omega}$
	$[[\text{tukuˀre}]_{\psi}\ \text{ba}]_{\omega}$	$[[\text{tukuˀre}]_{\psi}\ \text{ba}]_{\omega}$
	$[[\text{tukuˀtta}]_{\psi}]_{\omega}$	$[[\text{tukuˀt}]_{\psi}\ \text{ta}]_{\omega}$
	$[[\text{tukuˀtta}]_{\psi}\ \text{ra}]_{\omega}$	$[[\text{tukuˀt}]_{\psi}\ \text{tara}]_{\omega}$

In all Ichihara Japanese forms in (30), the p-stem consists of the m-stem plus maximally one syllable. The incorporation of material that does not belong to the m-stem can again be thought to be due to paradigmatic pressures. Judging from the accent patterns of Ichihara Japanese in (30), it seems that there is a preference for the accent to be located in the same location within the stem throughout all forms. Some of the paradigmatic correspondences are illustrated in (31).

(31)
$$\begin{array}{ccc} [[\text{tabeˀru}]_{\psi}]_{\omega} & \leftrightarrow & [[\text{tabeˀta}]_{\psi}]_{\omega} \\ \updownarrow & & \\ [[\text{tukuˀru}]_{\psi}]_{\omega} & \leftrightarrow & [[\text{tukuˀtta}]_{\psi}]_{\omega} \end{array}$$

[166] Observe that in Ichihara Japanese the accent does not fall on the suffix in *tukuˀtta* in (30). In this, Ichihara Japanese differs from Matsue Japanese, where the accent even avoids high vowels before the first half of a geminate. In Matsue Japanese the p-stem also incorporates both suffixes. However, the gerund-existential constructions involve the existential *or-* rather than *i-*, which is accented in isolation and therefore cannot be compared with *i-*.

The correspondence between the accented syllables in (31) can be analyzed as the result of the alignment of accent across paradigmatically related forms. This paradigmatic relatedness is relevant both among verbs that share a root (32a), as well as among verbs that share an affix (32b). The constraints are defined in terms of uniform exponence in (32). Allomorphs are assumed to be treated as the same form by the two constraints.

(32) a. UNIFORM EXPONENCE-ROOT = UE-ROOT
In two output forms that share the root *R*, the accent is located on the same mora.
b. UNIFORM EXPONENCE-AFFIX = UE-AFFIX
In two output forms that share the affix *A*, the accent is located on the same mora.

In the tableau in (33), these two constraints are collapsed into one UE-constraint in order to determine their relative importance. Apart from violation marks, the locus of violation (i.e. root, affix, and in case of affixes, which affix) is indicated. The "Tokyo output forms" in (33c) seem to be ruled out by the constraint UE-ROOT. The forms in (33b), on the other hand, seem to be penalized by UA-AFFIX, because the accent is aligned differently in the non-past forms. However, the attested output forms in (33a) also violate UE-AFFIX because the accent is aligned differently in the two past forms. This shows that not all suffixes behave identically with respect to UE-AFFIX, and therefore suggests that this constraint needs to be split into two constraints that refer to different sub-paradigms. The constraint referring to forms that include the non-past suffix (and other forms of the "non-past group") must be ranked higher than the one referring to past forms, for otherwise the forms in (33b) would be optimal in Ichihara Japanese due to its lower number of violations of DEP(M, P). Evidence for the distinction between two sub-paradigms for two groups of suffixes will be given in the discussion of Matsue Japanese in 6.2.2. Finally, like in all other dialects, m-stems that are smaller than p-stems are not allowed (33d).

(33) Verbal accent in Ichihara Japanese

tabe-[acc] -ru,-ta tukur-[acc] -u, -ta	MAX (M-P)	UE	UA-R (ψ, ψ)	DEP (M-P)	PENULT (ψ)
→a. [ta.beˀ.ru], [ta.beˀ.ta] [tu.kuˀ.ru], [tu.kuˀt.ta]		*AFFIX *-ta*		**,**, *, **	*
b. [taˀ.be.ru], [taˀ.be].ta [tu.kuˀ.ru], [tu.kuˀt].ta		*!AFFIX *-(r)u*		**, *	*
c. [ta.beˀ.ru], [taˀ.be].ta [tu.kuˀ.ru], [tu.kuˀt].ta		*!ROOT		**, *	
d. [taˀ.be.]ru, [taˀ.be].ta [tu.kuˀ.ru], [tu.kuˀt].ta		*AFFIX *-(r)u*	*!	 *	
e. [taˀ.be.]ru, [taˀ.be].ta [tuˀ.ku].ru, [tu.kuˀt].ta	*!	*!ROOT			

Because the suffixes *-(r)u* and *-ta* and the first syllable of the suffixes *-(r)eba* and *-tara* are parsed into a p-stem in regular forms, the prediction is that they are also parsed into a p-stem in contracted forms. Hence, the prediction is that the contracted gerund-existential forms will only consist of a single p-word, and thus are not assigned a default boundary accent. As we can see from the constracted forms of the Ichihara dialect that I collected (34), this prediction is supported by the data.

(34) *Regular inflections* *Contracted forms of gerund+ i-* (Ichihara)

a. $[[\text{aketa}]_{\psi}]_{\omega}$ ↔ $[[[\text{ake te}]_{\psi}\ \text{ta}^{0}]_{\psi}]_{\omega}$

b. $[[\text{aketaˀ}]_{\psi}\ \text{ra}]_{\omega}$ ↔ $[[[\text{ake te}]_{\psi}\ \text{taˀ}]_{\psi}\ \text{ra}]_{\omega}$

c. $[[\text{akeru}]_{\psi}]_{\omega}$ ↔ $[[[\text{ake te}]_{\psi}\ \text{ru}^{0}]_{\psi}]_{\omega}$

d. $[[\text{akereˀ}]_{\psi}\ \text{ba}]_{\omega}$ ↔ $[[[\text{ake te}]_{\psi}\ \text{reˀ}]_{\psi}\ \text{ba}]_{\omega}$

All Ichihara Japanese forms in (34) consist of a single p-word. Therefore, no junctural accents are assigned, and the underlying accent of the suffixes *-taˀra* and *-(r)eˀba* are faithfully realized.

The comparison of the Tokyo, Shizuoka, and Ichihara data provides us with evidence for a role of the p-stem in the verbal paradigm. Furthermore, it shows that the alignment of suffixes with p-stems is similar across all occurrences of the suffix. This is an important generalization, regardless of whether we formalize it by means of subcategorization schemas or by making use of the Uniform Alignment constraint.

While the accent patterns were not presented in terms of foot structure, in all three dialects, the accent in regular inflections in the three dialects can be restated as in (35).

(35) The rightmost foot of the p-stem carries the accent.

That this generalization is correct is shown in (36), where the attested output forms of basic inflected verbs in the three dialects discussed in this section is presented.

(36)	Ichihara Japanese	Tokyo Japanese	Shizuoka Japanese
a.	[[ta(beˀru)]$_{\psi}$]$_{\omega}$	[[ta(beˀru)]$_{\psi}$]$_{\omega}$	[[(taˀbe)]$_{\psi}$ ru]$_{\omega}$
	[[ta(beˀta)]$_{\psi}$]$_{\omega}$	[[(taˀbe)]$_{\psi}$ ta]$_{\omega}$	[[(taˀbe)]$_{\psi}$ ta]$_{\omega}$
b.	[[tu(kuˀru)]$_{\psi}$]$_{\omega}$	[[tu(kuˀru)]$_{\psi}$]$_{\omega}$	[[tu(kuˀru)]$_{\psi}$]$_{\omega}$
	[[tu(kuˀt)ta]$_{\psi}$]$_{\omega}$	[[tu(kuˀt)]$_{\psi}$ ta]$_{\omega}$	[[tu(kuˀt)]$_{\psi}$ ta]$_{\omega}$

Note that the constraint RIGHTMOST(ψ) is satisfied by all forms. The constraint ALIGN-R(ψ, F), however, is violated in the Ichihara past form in (36b).

In the following sections, it will be shown that the generalization in (35) actually holds for two other dialects as well: Matsue Japanese and Maisaka Japanese. Furthermore, I will present additional evidence for the p-stem and the role of paradigm uniformity.

6.2. Additional evidence for the p-stem and paradigm uniformity

6.2.1. P-stem-antepenultimate accent and accent in longer forms in Tokyo Japanese

In 6.1.2 it was mentioned that there are exceptional verbs in which the accent is associated with the antepenultimate rather than the penultimate mora of what I have argued to be the p-stem. An example of such a verb in given in (37c). In verbs with a root ending in -VVC like *toˀoru*, the accent falls on the antepenultimate mora of the p-stem, which must be accounted for. Importantly, however, not all verbs with roots ending in a -VVC sequence show this accent pattern. As we can see in (37d), there are "innovative verbs" (Tsujimura and Davis 2011) in which the accent pattern of verbs with roots ending in -VVC that do show regular p-stem-penultimate accent.

(37)	*Nonpast*		*Past*		
a.	ta.beˀ.-ru	(-2)	taˀ.be.-ta	(-2)	'eat'
b.	sya.beˀ.ru	(-2)	sya.beˀt.-ta	(-2)	'chat'
c.	toˀo.r-u	(-3)	toˀot-ta	(-3)	'pass(intr.)'
	haˀiru	(-3)	haˀit-ta	(-3)	'enter'
d.	tyaaˀ.r-u	(-2)	tyaaˀt.-ta	(-2)	'go out to have tea'
	o.da.kyuuˀ.r-u	(-2)	o.da.kyuuˀt.-ta	(-2)	'go to the suburbs of Tokyo (by taking a Odakyu train)'

What is peculiar about the nonpast forms in (37d) is that the accent falls on a syllable

non-head, which in general is not allowed in Tokyo Japanese (due to *H/NONHEAD$_{\sigma}$ or *CVX˺).[167] According to Tsujimura and Davis (2011), this is one of the pieces of evidence for positing a special construction that is responsible for the coinage of innovative verbs.

(38) Innovative verbs (Tsujimura and Davis 2011: 807)

*

. . . (C)V(C)V - r]$_{Vroot}$

Now, in Construction Morphology (Booij 2010), the lexicon is thought to be a network of relations, with new constructions arising as the result of schema interaction. In other words, the root-final accent must have its origin in the patterns that can be observed in existing words. It may be clear from (37) what this pattern is: penultimate accent. Let us assume then that the accent pattern of innovative verbs is the result of constraints that trigger penultimate accent, which in (39) is again collapsed into the cover constraint called PENULTACCENT(ψ), which is ranked above *H/NONHEAD$_{\sigma}$.[168]

(39) PENULTACCENT(ψ) >> *H/NONHEAD$_{\sigma}$.

The fact that innovative verbs get a default accent can also be confirmed by the results of a nonce-word experience conducted by Allen (2013). The table in (40) shows the percentage of a total of 100 forms ending in -VVC to which antepenultimate accent was assigned.

[167] According to two of my informants, two speakers of the standard variety of Japanese in their early twenties, it is not possible to pronounce the words in (37d) with a glottal stop that breaks the long vowel into two. Although superheavy syllables are generally avoided in Tokyo Japanese (Kubozono 1999), as shown by Tanaka (2000), superheavy syllables with a long vowel as in the past forms in (37d) are actually attested. Long vowels followed by a dependent mora can be regarded as being parsed into superheavy syllables due to a high degree of dependency of the second mora; it needs to receive the vowel feature from the head mora. Still, for the argument it would not matter if the past forms in (37d) were analyzed as *tya.a˺Q.ta* and *o.da.kyuu˺.Q.ta*.

[168] The insertion of an accent in innovative verbs seems to be related to type frequency. As shown by Akinaga and Sakamoto (2010), there are about twice as many verbs with accented roots as with unaccented roots.

(40) Antepenultimate (-3) accent in nonce words (Allen 2013: 113)

V1 \ V2	a	o	e	i	u
a	11.6%	6.3%	22.5%	16.9%	28.8%
o	4%	12.5%	6.5%	30.2%	18.2%
e	3.6%	7%	19.1%	7.1%	34%
u	5.8%	9.3%	10.2%	14.6%	16.9%
i	13.7%	3.6%	16.7%	9.6%	7%

It is clear from (40) that antepenultimate accent is infrequently assigned in nonce words that end in -VVC, which forms evidence for the penultimate pattern being assigned by default.

If the accent pattern of innovative verbs is predictable, this means the accent pattern in non-innovative verbs ending in -VVC is exceptional. Intuitively, we may propose that the accent of verbs like *toˈoru* is part of the form, e.g. /toˈor-/. This is actually what is proposed by Oshima (2014).

The tableau in (41) shows how the analysis would work for the simple non-past form. For ease of reference, UE-SUFFIX is replaced by "cvˈru". With FAITH-IO(Accent) ("the accent in the input is faithfully realized in the output") ranked above the other constraints, underlying accents are protected from being moved around.

(41) Input: /toˈor-(r)u ('pass-NONPAST')

toˈor-(r)u	FAITH-IO (Acc)	cvˈru	*H/NONHD$_\sigma$
→a. toˈoru		*	
b. tooˈru	*!		*

While the analysis works for forms that consist of a root and a single inflectional suffix, for longer forms such as that in (42), the wrong predictions are made. The actual grammatical form in (42b) loses out to the ungrammatical form in (42a) due to a violation of the same faithfulness constraint that was decisive in (41).

(42) Input: /toˈor-(s)ase-(r)u ('pass-CAUS-NONPAST')

toˈor-(s)ase-(r)u	FAITH-IO (Acc)	cvˈru	*H/NONHD$_\sigma$
→a. *toˈoraseru		*	
←b. toorase ˈru	*!		*

If we assume the accent falls on the rightmost foot of the p-stem, the problem can be resolved. The constraints responsible for this will be RIGHTMOST(ψ), Align-R(ψ, F) and FT-BIN. In the tableau in (43), only the first of these constraints is incorporated.

(43) Input: toˀor-(s)ase-(r)u (‘pass-CAUS-NONPAST’)

toˀor-(s)ase-ru	RIGHTMOST(ψ)	FAITH-IO (Acc)	*H/NONHD$_\sigma$
a. [(toˀo)ra(seru)]$_\psi$	*!		
b. [(toˀo)(rase)ru]$_\psi$	*!		
→c. [toora(seˀru)]$_\psi$		*	*

RIGHTMOST(ψ) is satisfied in all forms consisting of a stem and an inflectional form only, e.g. the conditional form *toˀoreba* in (44).

(44) Input: /toˀor-(s)ase-(r)u (‘pass-CAUS-NONPAST’)

	RIGHTMOST(ψ)
[(toˀo)re]ba	✓

Furthermore, the absence of accent shifts in forms ending in *-tara* and *-tari* can also be explained, as these suffixes fall outside the p-stem.

Notice that if the accentuation is indeed foot-based, this either implies that forms like *tyaaˀru* and *odakyuuˀru* violate syllable integrity (45a) or have trimoraic feet (45b).[169] Because trimoraic forms may be independently necessary to account for certain loanword accent patterns (see Chapter 3), it may be better to assume that (45b) is the winning candidate, and that (45a) is ruled out by Syllable Integrity.

(45) Input: /tyaar$_{[acc]}$-(r)u (‘go to have teaˀ-NONPAST’)

tyaar$_{[acc]}$-(r)u	RIGHTMOST(ψ)	cvˀru	*H/NONHD$_\sigma$
(→) a. tya(a.ˀru)			*
→b. (tyaaˀ.ru)			*
c. (tyaˀa.)ru		*!	

An alternative analysis in terms of OO correspondence is also available. As we can see in (46), if the faithfulness constraint is not FAITH-IO but FAITH-OO, which is not relevant

[169] This means that FT-BIN is ranked below cvˀru.

for the causative form, the attested output forms are generated.

(46) Difference beteen the nonpast and causative nonpast

toor$_{[acc]}$-(r)u [toˀoru]	FAITH-OO (Acc)	cvˀru	*H/NONHD$_{\sigma}$
→a. toˀoraseru		*	
b. tooraseˀru	*!		
toor$_{[acc]}$-(s)ase-ru [toˀoru]	FAITH-OO (Acc)	cvˀru	*H/NONHD$_{\sigma}$
c. toˀoraseru		*!	
→d. toraseˀru			

When a new word enters the language, there is no stored output form, which means FAITH-OO is not relevant.

(47) Input: /tyaar$_{[acc]}$-(r)u (‘go to have tea’-NONPAST’)

tyaar$_{[acc]}$-(r)u	FAITH-OO	cvˀru	*H/NONHD$_{\sigma}$
→a. tyaaˀ.ru			*
b. tyaˀa.ru		*!	

The challenge for the OO-based analysis is to explain why it is not relevant for forms with stem-forming derivational suffixes. The only possible answer is that FAITH-OO may target stems rather than only words (Alderete 1999). Thus, whereas in {toˀoru, toˀoreba, toˀore} an inflectional suffix is attached to an identical stem, this does not hold for *toraseˀru*, in which the nonpast suffix *-(r)u* is attached to the extended stem *toorase-*. In other words, FAITH-OO could be defined as a uniform exponence constraint referring to a stem rather than a root or a word.

The idea that FAITH-OO is relevant only to stems could be either stipulated like this, or it could be assumed to follow from Burzio’s theory of OO correspondence in which lexical representations are subject to different degrees of attraction (Burzio 2002, 2005a/b). Burzio (2002) defines such gradient attraction as follows.

(48) Gradient Attraction

a. The overall structure of a word *w* (in both its phonological and semantic components) is influenced by that of other words in the lexicon to which *w* is independently similar, and which can be thought of as "attractors" of *w*.

b. Attraction is stronger where independent similarity is greater.

Based on the idea of Gradient Attraction, Burzio (2002, 2005a/b) proposes the Representational Entailment Hypothesis in (49) (taken from Burzio 2005b).

(49) Representational Entailments Hypothesis (REH):
Mental representations of linguistic expressions are sets of entailments. E.g. a representation consisting of A and B corresponds to the entailments: A⇒B, B⇒A (if A then B; if B then A).

According to Burzio, the ranking of FAITH-OO constraints is the result of the number of shared entailments between two forms. The idea is examplified in (50), where we can see that (50a) and (50b) are more similar to each other than to (50c) in a number of ways. In terms of gradient attraction, (50a) and (50b) attract each other more strongly than they are attracted to (50c).

(50) The ranking of OO-FAITH computed in terms of entailment violations

Verbal forms	A. Root: segments	B. Root: accent	C. Root: semantics	D. Suffix: semantics
a. toˀor-u	too ✓	toˀo..	✓	¬CAUS ✓
b. toˀor-eba	too ✓	toˀo..	✓	¬CAUS ✓
c. toor-aseˀ-ru	too ✓	too.. ×	✓	CAUS ×

Put simply, the more entailment violations between two forms, the lower the ranking of FAITH-OO. Based on the entailments in (50), Faith-OO will be ranked higher for the pair {toˀoru, toˀoreba} than for {toˀoru, toˀoraseru}.

In the analysis based on FAITH-OO, the absence of a rightward shift in forms with stem-antepenultimate accent in Tokyo Japanese can be explained without making reference to prosodic structure other than the accent location. However, whatever the merits of Burzio's REH, an analysis based on the p-stem in combination with metrical structure is more straightforward. Moreover, the p-stem was shown to be necessary to

account for a number of accent patterns in not only Tokyo Japanese, but also three other dialects. Furthermore, in the next section, it will be shown that the p-stem plays a similar role in Matsue Japanese.

6.2.2. Sub-paradigms in Matsue Japanese

Accentual alternations in the verbal paradigm in Matsue Japanese provide us with additional evidence for an approach based on both prosodic structure and paradigm uniformity. Consider the forms in (51), which are adapted from Hiroto and Ōhara (1952) and which were confirmed during the fieldwork conducted in November 2012. Notice that there are two provisional forms, of which the the first one is the more colloquial one.

(51) a.	*Pattern I*	'build'	b.	*Pattern II*	'rise'	
	..μˀCu	tateˀru		..C**i**ˀ(r)u	ok**i**ˀru	nonpast
	..μˀV	tateˀe		..C**i**ˀi	ok**i**ˀi	imperative
	..μˀCa	tateˀra		..C**i**ˀ(r)a	ok**i**ˀra	provisional (i)
	..μˀCeba	tateˀreba		..C**i**ˀ(r)eba	ok**i**ˀreba	provisional (ii)
	..μˀte	tateˀte		..Ci(t)t**e**ˀ	okit**e**ˀ	gerund
	..μˀta	tateˀta		..Ci(t)t**a**ˀ	okit**a**ˀ	past
	..μˀtara	tateˀtara		..Ci(t)t**a**ˀra	okit**a**ˀra	conditional

The verb forms above show that there are two types of accented verbs in Matsue Japanese: those in which the accent is realized on the same mora throughout the paradigm (Pattern I), and those in which the accent alternates depending on the vowel height of the stem-final vowel (Pattern II). The difference between the two patterns is not one of different input patterns. The forms in (52a) and (52b) show that in all cases the accents must be the property of the root, and not the suffix.

(52) a.	'put on'		b.	'sell'	
	kabutt**e**ˀ	LLØHˀ		utte^{0}	LLH
	kabutt**a**ˀ	LLØHˀ		utta^{0}	LLH
	kabutt**a**ˀra	LLØHˀL		uttara^{0}	LLHØ

The difference between the two provisional forms and the gerund, past, and conditional forms in (51b) is of crucial importance. In the latter three forms, the accent avoids the stem-final high vowel, associating to the suffix-initial syllable. Within both of the two provisional forms, on the other hand, the accent falls on a high vowel followed by a non-

high vowel, which in Matsue Japanese is generally not allowed (see Chapter 5, section 5.3.2.1).

The difference in tone-vowel interaction between the two groups of forms suggests that there are two "sub-paradigms" (see Burzio and Tantalou 2006): one for forms with suffixes that group with the non-past suffix, and one for forms with suffixes that group with the past suffix. The paradigmatic relations among the roots (horizontal) and affixes (vertical) are illustrated in (50), to which I have added the "alternative" marker *-tari*, and the continuative form, which in the case of V-verbs has zero exponence (and in C-verbs is realized by *-i*.).

(53) a.	okiˀ-ru	———	okiˀ- reba	———	okiˀ-i	———	okiˀ	UE-ROOT
	|		|		|		|	UE-AFFIX
	tateˀ-ru	———	tateˀ-reba	———	tateˀ-e	———	tateˀ	UE-ROOT
b.	oki-teˀ	———	oki-taˀ	———	oki-taˀra	———	oki-taˀri	UE-ROOT
								No UE-AFFIX
	tateˀ-te	———	tateˀ-ta	———	tateˀ-tara	———	tateˀ-tari	UE-ROOT

Interestingly, uniform exponence of affixes can be observed in (53a), but not in (53b). For the suffixes in (53a), it appears that more importance is attached to keeping the accent in the same position across forms with the same suffix and different stems. Within the two sub-paradigms, uniform exponence is observed for the stem. However, across the two sub-paradigms, only verbs whose root-final vowel is not a high vowel show uniform exponence of the root.

Now, because the presence or absence of uniform exponence across the roots of the different sub-paradigms is predictable on the basis of the phonological make-up, it is ad hoc to assume two differently ranked UE-ROOT constraints for verbs with roots ending with a final high vowel and those without such a vowel. Rather, it seems that the absence of this uniform exponence across stems of the different sub-paradigms is the result of a difference in the strength of UE-AFFIX. Because the accent is not shifted to the suffix in the forms of the nonpast group ('Group I'), violations of UE-AFFIX by suffixes of the past-tense group ('Group II') entail violations of UE-ROOT. Thus, we can posit two different OO constraints for the two groups.[170]

[170] The fact that the suffixes of one of the groups all begin with a coronal stop, combined with similarity in accentual behavior (i.e. accent attraction in verbs with stems ending in a high vowel) makes them an easily recognizable subparadigm.

(54) a. UE-AFFIX$_{I}$ (Accent)
In two output forms that share the affix *α* (*α= -u/-ru ~ -eba/-reba ~ -e/V*), the accent is located on the same mora.

b. UE-AFFIX$_{II}$ (Accent)
In two output forms that share the affix *α* (*α= -te, -ta, -tara, -tari*), the accent is located on the same mora.

The first of the two constraints, UE-AFFIX$_{I}$ must be ranked above ALIGN(H, A) (see Chapter 5, section 6.3.2.1), because the accent does not avoid high vowels in Group I. UE-AFFIX$_{II}$, on the other hand, must be ranked below ALIGN(H, A), because the accent does avoid the stem-final high vowel in Group II.

(55) UE-AFFIX$_{I}$ >> ALIGN(H, A) >> UE-AFFIX$_{II}$

Apart from uniform exponence constraints that compare the output forms of constructions containing the same suffixes, we also need a constraint that compares output forms with the same stem: UE-ROOT. The constraint UE-ROOT prevents the accent in forms with stems that do not end in high vowels from shifting to the suffix (**tateta*ˀ), which would be preferable from the point of view of UE-AFFIX$_{II}$. However, the constraint is not strong enough to block the association of the accent to the suffix in verbs with stems that do end in a high vowel. This means that the constraint is ranked between ALIGN(H, A) and UE-AFFIX$_{II}$, as in (56).

(56)

oki-$_{[acc]}$ -(r)eba,-ta tate-$_{[acc]}$ -(r)eba, -ta	UE-AFFIX$_{I}$	ALIGN (H, A)	UE-ROOT	UE-AFFIX$_{II}$
→a. okiˀreba, okitaˀ tateˀra, tateˀta		*	*	*
b. okiˀreba, okiˀta tateˀra, tateˀta		*,*!		
c. okirebaˀ, okitaˀ tateˀra, tateˀta	*!		*	*
d. okiˀreba, okitaˀ tateˀra, tatetaˀ		*	* *!	

Now, in order for the continuative form to come out as *okiˀ* in this ranking, the continuative affix – in this case the segmentally empty allomorph – must be thought to belong to the first group of affixes, and therefore be subject to UE-AFFIX$_{I}$. However, the

actual situation is more complicated. Consider the difference between the continuative forms of C-verbs and V-verbs in (57).

(57)	a.	huˈr-u	HˈL	‘fall-NONPAST’
		huˈr-i	HˈL	‘fall-CONT’
	b.	okiˈ-ru	LHˈL	‘rise-NONPAST’
		okiˈ	LHˈ	‘rise-CONT’

The two continuative forms show differences in exponence in both segmental and accentual terms. The form in (57a) has a suffix *-i*, and the accent is realized on the penultimate syllable of the derived form. In contrast, the continuative form in (57b) has no segmental features, and an accent on its final syllable. Although we could say they are identical in terms of accent in the sense that in both forms the accent falls on the root- or stem-final vowel, this account has a serious problem: how does a constraint like UE-AFFIX know that it must check whether the accent is on the root-final vowel? Surely, we could stipulate this in the definition of the constraint. However, it seems rather ad hoc to assume a constraint on uniform exponence in which an affix, i.e. non-root material, directly refers to the root or stem. In the approach based on parallel representations, the identity in accent location should hold for the output phonological forms (“P-Structure”), in which no morphological information is directly available. Therefore, UE-AFFIX, which checks whether the accent is aligned in the same way in the two different output forms, has no direct access to the morphological make-up (M-Structure) in (58). This means that UE-AFFIX is violated by the two continuative forms, which is shown in (58).

(58)	M-Structure		P-Structure	
	[[hur]$_{\text{Stem}}$ i]	↔	Output: huˈri]$_{\psi}$]$_{\omega}$	penultimate accent
	‘fall-CONT’		|	≠
	[oki]$_{\text{Stem}}$ Ø]	↔	Output: okiˈ]$_{\psi}$]$_{\omega}$	final accent
	‘rise-CONT’			

The challenge is to explain why the accent does not fall on the non-high vowel in *oki* (**oˈki*). The non-identity in (58) suggest that UE-AFFIX is not ranked highly for the continuative. This means that we are dealing with UE-AFFIX$_{\text{II}}$. However, if we are dealing with UE-AFFIX$_{\text{II}}$, it should be possible to violate UE-ROOT as well and thereby satisfy ALIGN(A, H), as in for example the past form. However, this does not happen.

A possible solution to this dilemma is to posit the relativized UE-ROOT constraint in (59).

(59) MINIMAL UNIFORM EXPONENCE ROOT = MIN-UE-ROOT
The accent in an output form O_1 consisting of a root R_1 and an affix A_1 is aligned with the accent of a form O_2 that contains the same root R_1 and an affix A_2.

According to the constraint in (59), there must be minimally one other form in the paradigm that has the accent in the same position. In fact, constraints like that in (59) for which even the exact identity of the affixes A_1 and A_2 are spelled out have been proposed by Burzio and Tantalou (2007), who call them constraints on "subparadigm uniformity". Although adopting such subparadigm uniformity constraints would be possible for Matsue Japanese, the advantage of MIN-UE-ROOT is that it is more general; it does not spell out which form this should be, and only promotes minimal uniformity within the paradigm. Thus, in the case of an input form /oki$_{[acc]}$/, the only options are a stem with an accent on the final high vowel (as in the nonpast form *oki'ru*), or an accentless stem (as in the past form *okita*'). The second form is not an option, as it would result in accent deletion. Hence, the form with an accent on the final syllable is selected: *oki*'.

If we adopt MIN-UE-ROOT, the continuative forms can be assumed to be subject to UE-AFFIX$_{II}$, which makes sense from both a historical as well as a cross-dialectal point of view.[171]

The tableau in (60) illustrates the interaction of the constraints introduced above. The candidates in (60) consist of C-verbs and V-verbs with a final high vowel (which in C-verbs is of course followed by the root-final consonant) and without a final high vowel. For both C-verbs and V-verbs, an affix of Group I, an affix of Group II, and the continuative (which also belongs to Group II) are considered as candidates. The lower-ranked constraint UE-AFFIX$_{II}$ is not considered as it is not relevant for the candidates considered. The most important violations are marked in grey. It is important to realize that adding candidates does not influence the results of the evaluation.

[171] The forms of Group II are historically derived from the continuative suffixed by the gerund *-te* and different inflections of the existential *ar-*. What is more, as was mentioned in 7.1.2, the accent in the continuative form in Tokyo Japanese falls on the same mora as that in the past form and the other forms of Group II.

(60)

oki-[acc] -(r)eba, Ø/-i, -ta hur-[acc] -(r)eba, Ø/-i, -ta	MIN-UE-ROOT	UE-AFFIX$_I$	ALIGN (H, A)	UE-ROOT
→a. okiˀreba, okiˀ, okitaˀ tateˀreba, tateˀ, tateˀta huˀreba, huˀri, huttaˀ kaˀkeba, kaˀki, kaˀita			*,* *,*	* *
b. okiˀreba, oˀki, okitaˀ tateˀreba, taˀte, tatetaˀ huˀreba, huˀri, huttaˀ kaˀkeba, kaˀki, kaitaˀ	*! *		*, *,*	
c. okiˀreba, okiˀ, okiˀta tateˀreba, tateˀ, tateˀta huˀreba, huˀri, huˀtta kaˀkeba, kaˀki, kaˀita			*,*,* *,*!,*	
d. okireˀba, okiˀ, okiˀta tateˀreba, tateˀ, tateˀta hureˀba, huˀri, huttaˀ kaˀkeba, kaˀki, kaˀita		*! *	*,* *,*	* *

In (60b), there are fewer violations of ALIGN(H, A) than in the winning candidate, but MIN-UE-ROOT is crucially violated. In (60c), there is more stem-uniformity, but at the cost of more violations of ALIGN(H, A) than in the winning candidate. The candidate in (60d) loses out because of fatal violations of UE-AFFIX$_I$, which compares the suffixes belonging to Group I such as *-(r)eba* across words with different roots.

In the tableau in (60), only forms in which the accent falls on one of the last two syllables of the word are considered. The next step is thus to account for the absence of pre-penultimate accent.

Let us assume that the domain of accent assignment in verbs is the p-stem, as in the dialects discussed above. In Matsue Japanese this makes sense, because the accent does not consistently fall on the same syllable in some morphological domain, e.g. the stem-final syllable. If we assume the p-stem includes the m-stem plus an additional syllable, as in Ichihara Japanese, the generalization is that the accent falls on one of the two final syllables of the p-stem.

(61) [okiˀru]$_\psi$, [okiˀre]$_\psi$ba, [okiˀ]$_\psi$, [okitaˀ]$_\psi$, [okitaˀ]$_\psi$ra ‘rise’
[tateˀru]$_\psi$, [tateˀre]$_\psi$ba, [tateˀ]$_\psi$, [tateˀta]$_\psi$, [tateˀta]$_\psi$ra ‘build’
[huˀru]$_\psi$, [huˀre]$_\psi$ba, [huˀri]$_\psi$, [huttaˀ]$_\psi$, [huttaˀ]$_\psi$ ra ‘fall’
[kaˀku]$_\psi$, [kaˀke]$_\psi$ba, [kaˀki]$_\psi$, [kaˀita]$_\psi$, [kaˀita]$_\psi$ra ‘write’

If the p-stem is the domain of accent assignment, our next task is to account for the absence of pre-penultimate accent. In a metrical analysis, we first need a constraint that says the p-stem must be right-aligned with a foot: ALIGN-R(ψ, F). Apart from this, we need the earlier mentioned constraint (13b) that forces the accent to be associated to the rightmost foot of the p-stem: RIGHTMOST(ψ). An important characteristic of RIGHTMOST(ψ) is that it is not violated in the past forms of C-verbs even if the foot is not right-aigned with the p-stem, as in *(kaˀi)ta*. Thus, RIGHTMOST(ψ) can be given a high ranking. If we add to these constraints the well-known constraint FT-BIN, we have the ingredients necessary to explain the difference between *kaˀita* and *okitaˀ*.

The tableau in (62) shows how UE-ROOT and FT-BIN conspire to rule out **(kai)(taˀ)* in (62c).[172] A form with the same foot structure but the accent in the non-rightmost foot loses out due to RIGHTMOST(ψ). Note that apart from the underlying form in the input, in which two root allomorphs are given, a stored output form is given to which the constraint UE-ROOT can refer.

(62)

kak[acc]~kai[acc]-ta [kaˀku]	RIGHT MOST(ψ)	UE-ROOT	ALIGN-R (ψ, F)	FT-BIN
→a. [(kaˀi)ta]ψ			*	
b. [(kaˀi)(ta)]ψ	*!			*
c. [(kai)(taˀ)]ψ		*		*!

The tableau in (63) illustrates how RIGHTMOST(ψ) and ALIGN-R(ψ, F) help in ruling out candidates with the accent on the pre-penultimate syllable in (63b) and (63c). ALIGN(H, A) penalizes the candidate that assigns an accent to the penultimate syllable (63d). As a result, the final-accented form in (63a) emerges as the winner.

(63)

oki[acc]-ta [okiˀru]	RIGHT MOST(ψ)	ALIGN (H, A)	UE- ROOT	ALIGN-R (ψ, F)	FT-BIN
→a. [(oki)(taˀ)]ψ			*		*
b. [(oˀki)(ta)]ψ	*!		*		*
c. [(oˀki)ta]ψ			*	*!	
d. [o(kiˀta)]ψ		*!			

[172] Forms that violate Syllable Integrity as in **ka(itaˀ)* are not considered.

For V-verbs with roots ending in non-high vowels, the candidate with a final trochaic foot does not violate any of these constraints.

(64)

tate$_{[acc]}$-ta [tateˈru]	RIGHT MOST(ψ)	ALIGN (H, A)	UE-ROOT	ALIGN-R (ψ, F)	FT-BIN
a. [(tate)(taˈ)]$_{ψ}$			*!		*
b. [(taˈte)(ta)]$_{ψ}$	*!		*		*
c. [(taˈte)ta]$_{ψ}$			*!	*	
→d. [ta(teˈta)]$_{ψ}$					

The proposed ranking, combined with the constraints proposed above, will assign prosodic structure in verbs as in (65). In the interest of visibility, the level of the p-word is omitted.

(65)a. [o(kiˈru)]$_{ψ}$, [o(kiˈre)]$_{ψ}$ba, [(okiˈ)]$_{ψ}$, [(oki)(taˈ)]$_{ψ}$, [(oki)(taˈ)]$_{ψ}$ra ‘rise’

b. [ta(teˈru)]$_{ψ}$, [ta(teˈre)]$_{ψ}$ba, [(tateˈ)]$_{ψ}$, [ta(teˈta)]$_{ψ}$, [ta(teˈta)]$_{ψ}$ra ‘build’

c. [(huˈru)]$_{ψ}$, [(huˈre)]$_{ψ}$ba, [(huˈri)]$_{ψ}$, [(hut)(taˈ)]$_{ψ}$, [(hut)(taˈ)]$_{ψ}$ ra ‘fall’

d. [(kaˈku)]$_{ψ}$, [(kaˈke)]$_{ψ}$ba, [(kaˈki)]$_{ψ}$, [(kaˈi)ta]$_{ψ}$, [(kaˈi)ta]$_{ψ}$ra ‘write’

The foot-based analysis works well, but to be fair, a tonal alternative seems available. In a tonal approach, the absence of pre-antepenultimate accent in forms like *okitaˈ* (**o*ˈ*kita*), and the presence of this pattern in forms like *kaˈita* (**kaitaˈ*) could be due to a preference for the H tone to align to the heavy syllable in *kaˈita*, rather than to the right edge of the p-stem (i.e. on the suffix).[173]

In any case, in this section I have shown that a p-stem-based analysis of Matsue Japanese verbal accentuation is able to account for the facts. If we also introduce foot structure, a straightforward analysis of p-stem-penultimate accent is available. While a

[173] Interestingly though, a tonal analysis is only possible if the grammar is able to distinguish between lexical tones and postlexical H tones. The reason for this is as follows. Recall that heavy syllables ending in sonorant moras receive a H tone in Matsue Japanese. This means that in phrase-initial position, the first syllable of *kaita* would be pronounced on a high pitch even if the accent fell on the suffix. Thus, in a model in which lexical and postlexical representations are not distinguished, we would thus expect the position of the accent to depend on whether the verb is phrase-initial or not, which is not the case. In phrase-initial position, the output would be **kaitaˈ* [HHH], whereas in phrase-non-initial position, it would be *kaˈita* [HLØ]. As this kind of variation is not observed, we can conclude that a tonal analysis is only possible in an approach that distinguishes between a lexical and postlexical round of computation, or a model in which the H tone on the accented syllable is represented in a different way from a H tone associated to non-accented syllables.

tonal analysis does not seem impossible, the foot-based approach can account for all dialects in the same way, including the exceptional antepenultimate patterns. Therefore, I concude that a foot-based approach to verbal accent is superior to a purely tonal one.

In the next section, it is shown that an alternative analysis based on level ordering within the lexical phonology instead of paradigmatic relations is not available.

6.2.3. Against level ordering

The case of Matsue Japanese is of course conclusive evidence for paradigm uniformity only if alternatives are less attractive. In this section, it is shown that an approach based on level ordering is only possible under some very specific conditions, which make it unattractive.

An analysis based on level ordering (66) is only possible under the following two assumptions: (i) the conditional should consist of a stem followed by *-ba* rather than a stem followed by *-(r)eba*, and (ii) phonology applies according to the principle of "level economy", according to which phonology only applies when a morphological operation has taken place at the level in question (Inkelas and Orgun 1995), and (iii) the phonology of nouns and verbs differs even at the postlexical level.

Analyzing the conditional form as consisting of a stem followed by *-ba* is necessary to get the accent in the right place (penultimate) at Level 1. A rule of (or constraints on) accent assignment will assign an accent to the penultimate syllable, oblivious of sonority. At Level 2, sonority does become relevant, but accents that are already assigned may not shift. In forms in which the accent location is not determined yet, on the other hand, a constraint against the association of the accent with a high vowel will ensure the accent is assigned to the suffix. This derivation is shown in (66).

(66) Level-ordering with level economy

		nonpast	*conditional*	*past*
Lexicon		oki$_{[acc]}$	oki$_{[acc]}$	oki$_{[acc]}$
Level 1	morphology	oki-ru	oki-re	-
	accent ("sonority-insensitive")	okiˀru	okiˀre	-
Level 2	morphology	-	okiˀre-ba	oki-ta
	accent ("sonority-sensitive")	-	-	oki-taˀ
Output		okiˀru	okiˀreba	okitaˀ

From (66), it should also be clear why level economy is necessary: if the stem *oki-* were inserted at Level 1 for the past form, the floating accent would have docked onto the stem,

making any further modifications impossible. Finally, the need to assume nouns and verbs are subject to different phonologies is related to the fact that in noun-particle combinations, in the variety of Matsue Japanese described by Hiroto and Ōhara (1952) discussed in Chapter 5, the accent shifts from a noun-final high vowel to the particle. The problem is, however, that case particles are attached syntactically rather than the morphologically,[174] which means that the accent shift is part of the postlexical phonology. Hence, we need to assume that nouns and verbs are subject to different postlexical phonologies, which is a very unattractive idea. Therefore, we may conclude that there is no convincing explanation for the different behavior of the suffixes belonging to Group I and II other than paradigmatic relations.

6.3. Maisaka Japanese

In Chapter 4 it was shown that iambic feet play an important role in the nominal accent system of the Maisaka dialect. In this section it is shown that the metrical analysis can be naturally extended to verbs.

In Chapter 4 the foot structure of verbs was already introduced in order to show that verbs are subject to the same constraints as nouns. It was also pointed out that the foot containing the rightmost vowel of the m-stem is accented in this dialect. As in the other dialects discussed in this section, the generalization can be restated in terms of the p-stem: the rightmost foot in the p-stem carries the accent. This generalization can be accounted for by a high ranking of ALIGN-R(ψ, F) and RIGHTMOST(ψ).

The prosodic structure of inflected V-verbs and C-verbs of different conjugations are presented in (67) and (68), respectively. The parsing of particles is left unspecified.

(67)	*mi*-[acc] ‘see’	*tabe*-[acc] ‘eat’	*atume*-[acc] ‘gather(tr.)’
non-past	[[(miˈru)]$_{\psi}$]$_{\omega}$	[[(tabeˀ)]$_{\psi}$ ru]$_{\omega}$	[[(atu)(meˈru)]$_{\psi}$]$_{\omega}$
transitional	[[(miruˀ)]$_{\psi}$]$_{\omega}$ to	[[(tabeˀ)]$_{\psi}$ ru]$_{\omega}$ to	[[(atu)(meruˀ)]$_{\psi}$]$_{\omega}$to
provisional	[[(miryaˈa)]$_{\psi}$]	[[(tabeˀ)]$_{\psi}$(ryaa)]$_{\omega}$	[[(atu)(meryaˈa)]$_{\psi}$]$_{\omega}$
past	[[(miˈta)]$_{\psi}$]$_{\omega}$	[[(tabeˀ)]$_{\psi}$ ta]$_{\omega}$	[[(atu)(meˈta)]$_{\psi}$]$_{\omega}$
conditional	[[(mitaˀ]$_{\psi}$ ra]$_{\omega}$	[[(tabeˀ)]$_{\psi}$(tara)]$_{\omega}$	[[(atu)(metaˀ)]$_{\psi}$ ra]$_{\omega}$
gerund-foc	[[(miteˀ)]$_{\psi}$]$_{\omega}$ mo	[[(tabeˀ)]$_{\psi}$te]$_{\omega}$ mo	[[(atu)(meteˀ)]$_{\psi}$]$_{\omega}$ mo]

[174] See Vance (1993) for evidence, but see Sells (1995) for a different view.

(68)	*kak*-[acc] ‘write’	*aruk*-[acc] ‘walk’	*atumar*-[acc] ‘gather(intr.)’
non-past	[[(kaˀku)]$_{\psi}$]$_{\omega}$	[[(aruˀ)]$_{\psi}$ ku]$_{\omega}$	[[(atu)(maˀru)]$_{\psi}$]$_{\omega}$
transitional	[[(kakuˀ)]$_{\psi}$]$_{\omega}$ to	[[(aruˀ)]$_{\psi}$ ku]$_{\omega}$ to	[[(atu)(maruˀ)]$_{\psi}$]$_{\omega}$ to
provisional	[[(kakyaˀa)]$_{\psi}$]$_{\omega}$	[[(aruˀ)]$_{\psi}$(kyaa)]$_{\omega}$	[[(atu)(maryaˀa)]$_{\psi}$]$_{\omega}$
past	[[(kaˀi)]$_{\psi}$ ta]$_{\omega}$	[[(aruˀi)]$_{\psi}$ ta]$_{\omega}$	[[(atu)(maˀt)]$_{\psi}$ ta]$_{\omega}$
conditional	[[(kaˀi)]$_{\psi}$(tara)]$_{\omega}$	[[(aruˀi)]$_{\psi}$(tara)]$_{\omega}$	[[(atu)(maˀt)]$_{\psi}$(tara)]$_{\omega}$
gerund-foc	[[(kaˀi)]$_{\psi}$ te]$_{\omega}$ mo	[[(aruˀi)]$_{\psi}$ te]$_{\omega}$ mo	[[(atu)(maˀt)]$_{\psi}$ te]$_{\omega}$mo

The data in (67) and (68) shows that depending on the length of the stem, material not belonging to the m-stem is parsed into the p-stem (e.g. all forms of *mi-* ‘to see’), or material belonging to the m-stem is left out of the p-stem (as in e.g. *aruˀk-u* ‘walk-NONPAST’). The fact that there is misalignment in both directions shows us that both DEP(M-P) and MAX(M-P) can be violated. Let us assume that a top-ranked left-alignment constraint (69) is responsible for the creation of p-stems and interacts with DEP(M-P) and MAX(M-P) in the construction of p-stems.

(69) ALIGN-L(M-Stem, P-stem) = ALIGN-L(St, ψ)
The left edge of an m-stem is aligned with the left edge of a p-stem.

The constraint in (69) is assumed to be undominated, and will not be further discussed below.

Maisaka Japanese differs from the nearby dialect of Shizuoka in that the stem-final consonant is allowed to be left out of the p-stem. This means that MAX (M-P), the constraint that militates against segments of the m-stem not being parsed into the p-stem, is crucially violated in forms like *aruˀku*. The selection of the form *aruˀk-u* ‘walk-NONPAST’ is presented in (70), where two highly ranked foot-alignment constraints, RIGHTMOST(ψ) and ALIGN-R(ψ, F), ensure a foot is aligned with the right edge of a p-stem. Candidates without an initial binary foot are not considered; see Chapter 4 for the constraints that ensure words are assigned an initial binary foot. As discussed in Chapter 4, all else being equal, the accented foot in Maisaka Japanese will be iambic because of the high ranking of the constraint IAMB (rather than TROCHEE). Candidates with trochaic feet will not be considered, an therefore IAMB is left out of the tableau in (70), as is NONFINALITY(μ), which bans accents on the phrase-final mora.

(70)

aruk[acc]-u	RIGHT MOST(ψ)	ALIGN-R (ψ, F)	FT-BIN	MAX (M-P)
→a. [[(aruˀ)]ψ ku]ω				*
b. [[(aruˀ)ku]ψ]ω		*!		
c. [[(aruˀ)(ku)]ψ]ω	*!		*	
d. [[(aru)(kuˀ)]ψ]ω			*!	

Next we turn to the question of why in the provisional form the accent does not fall on the final syllable *aruˀkyaa*. Importantly, DEP(M-P), the constraint against the insertion of non-stem material into the p-stem, cannot be responsible for this, because the p-stem contains two additional segments in forms like *atumeˀru*. This suggests DEP(M-P) is ranked below MAX(M-P). This means that under the current ranking, the unattested form in (71) will be wrongly selected.[175]

(71)

atume[acc]-ru	RIGHT MOST(ψ)	FT-BIN	MAX (M-P)	DEP (M-P)
←a. [[(aruˀ)]ψ (kyaa)]ω			*!	
b. [[(aruˀ)(kyaa)]ψ]ω	*!			
→c. *[[(aru)(kyaˀa)]ψ]ω				***

Clearly, the constraint that rules out the wrong winner in (71c) must be one that refers to the accent location across words that share a stem. This idea can be formalized in a number of ways. For example, it is possible to posit a constraint that aligns the right edge of p-stems across words sharing the same lexical root (rather than the same suffix, as the constraint in (16)). Or, we could adopt the constraint OO-FAITH(Fˀ) discussed in Chapter 4, which penalizes morphologically related words with an accent in a different foot. Here I adopt the latter, because the same constraint can be used for nouns, making the sponsor-alignment constraint discussed in Chapter 4 unnecessary. The constraint OO-FAITH(Fˀ) must outrank at least MAX (M-P) in order for the correct output form in (72a) to be generated.

[175] The number of violations of DEP(M-P) in (71c) is two or three, depending on whether consonant-glide sequences are analyzed as one or two segments. To keep the counting transparant, I here assume the onset contains two segments (see Vance 1987 for discussion).

(72)

aruk$_{[acc]}$-yaa [(aruˀ)ku]	RIGHT MOST(ψ)	OO-FAITH(Fˀ)	FT-BIN	MAX (M-P)	DEP (M-P)
→a. [[(aruˀ)]$_{ψ}$ (kyaa)]$_{ω}$				*	
b. [[(aruˀ)(kyaa)]$_{ψ}$]$_{ω}$	*!				
c. [[(aru)(kyaˀa)]$_{ψ}$]$_{ω}$		*!			***

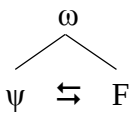
The derivation of inflected forms of verbs with shorter or longer roots works in the same way, although the number of violations of the lower ranked MAX(M-P) and DEP(M-P) will differ per form.

Summarizing, the analysis proposed Maisaka Japanese nouns in Chapter 5 can be extended to verbs if we introduce the concept of p-stem, which was shown to be independently necessary in the analysis of four other dialects.

6.4. Summary

In this chapter, three important findings were obtained. First, we found evidence for the p-stem playing an important role in the morpho-accentual grammar. Second, it was confirmed that by adopting the foot, verbal accent in Tokyo Japanese and Matsue Japanese can be analyzed in straightforward ways. In fact, it was shown that the general "accent rule" for verbs in all dialects can be formulated as putting an accent on the final foot of the p-stem. The different accent patterns are thus caused by differences in the way in which m-stems are projected onto p-stems, and in the case of Maisaka Japanese, a difference in the type of foot that is preferred.

The evidence for the interaction between these two constituents is interesting, because (at least) depending on the construction (e.g. underived nouns vs. verbs), they both must be assumed to be dominated by the p-word. This means that either there is a single prosodic hierarchy that is construction-specific (with the p-stem only appearing between the p-word and the foot in constructions in which a morpho-syntactic stem is involved), or the hierarchy splits into two at the level of the p-word, which dominates the construction-specific p-stem and the foot. If we adopt the latter option, we can still account for the fact that words and stems are subject to different minimality conditions (two syllables vs. two moras; see Chapter 2). Through edge alignment, the p-stem can be thought to map onto a foot, which consists of minimally two moras. The relations between the p-word, the p-stem, and the foot are shown in (74). Relations of dominance are indicated by straight lines. The arrows indicate a relation in terms of edge alignment in which neither of the two constituents dominates the other.

(74) 

Third, evidence for paradigm uniformity was presented. Although paradigm uniformity had been claimed to be relevant in Tokyo and Shizuoka Japanese by Yamaguchi (2010a/b), the role of prosodic structure had not been sufficiently considered. In this chapter, I showed that paradigm uniformity and prosodic structure both play important roles in the accent system of verbs across Japanese dialects.

Conclusion

In this thesis, I have discussed several issues related to word-prosodic structure in Japanese. In this conclusion, the main findings are summarized and suggestions for topics that deserve to be looked at more carefully in future studies are given.

After motivating the present study in the first chapter, in Chapter 2 I started with an overview of claims that have been made in the literature about word-prosodic structure in Japanese. Special attention was given to foot structure, and it was shown that even in prosodic morphology, in which the role of the foot is usually taken for granted, there are some issues that require further investigation. It was argued that a potential role for the p-stem in minimality effects and other prosodic-morphological phenomena should be taken seriously, and that the relation between the foot, the syllable, and the mora needs to be clarified before we can decide on the right analysis. I also pointed out that before introducing the p-stem into prosodic morphology, independent evidence for this category, and, if possible, its relation with the foot, would be welcome. As for the p-word, I presented evidence in favor of an approach in which particles are parsed outside the p-word, without forming a recursive p-word with its host. However, combinations of particles do form recursive p-words, the result of which is the insertion of a default accent.

Chapter 3 presented a critical review of foot-based approaches to Tokyo Japanese accent, followed by a purely tonal analysis. It was shown that while the proposed foot-based analyses are insightful, the generalizations concerning default loanword accent and non-phrasal compounds can be restated in terms of tone. For two phenomenon, phrasal compound accent and the emergence of unaccentedness, a non-metrical tonal approach is only possible if we seek for explanations outside the phonological grammar, which speaks in favor of the metrical approach. Still, because foot-based analyses of the different phenomena are not (yet) in harmony with each other, there remain a number of unresolved issues as to what this foot structure looks like.

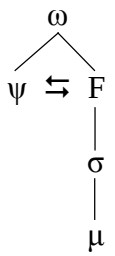
In Chapter 4, conclusive evidence for a metrical approach to Japanese pitch accent from the Maisaka dialect was presented. It was shown that the accentual alternations in this variety of Japanese can only be convincingly accounted for in a foot-based approach. Moreover, evidence from this dialect for covert feet and the co-existence of trochees and iambs was found. Interestingly, the default foot of Maisaka Japanese turned out to be the iamb, rather than the trochee, which has been claimed to be the default foot at least for Tokyo Japanese. Another interesting finding was that case particles may be incorporated into the p-word so that a word- and/or foot-minimality constraint can be satisfied.

In Chapter 5, I started with a description of tone-vowel interaction in three dialects: Matsue, Ichihara, and Kanazawa Japanese. After discussing the data and the most

important descriptive generalizations, analyses were proposed for each of the dialects. In all three dialects, constraints on tonal alignment were shown to play an important role. At least in Matsue and Ichihara Japanese, the interaction of H tone with vowel height cannot be reduced to the interaction of H tone and vowel height with metrical structure. In Kanazawa Japanese, on the other hand, the interaction between phrasal H tones and vowel height was shown to be mediated by the foot. In all three dialects, both accentual and phrasal H tones interact with vowel height, which suggests that the H tone is not simply a feature of the postlexical phonology or phonetic implementation. Still, it does seem that some constraints must be assumed to be sensitive to accentual H tones only, which suggests that the representation of the accentual H tone may differ from that of H tones linked to TBUs that are not accented. The question whether this means that the pitch accent itself is represented in the lexicon as a diacritic rather than in terms of tone remains unanswered.

Chapter 6 contains an analysis of accentual alternations in the verbal paradigms in the dialects of Tokyo, Shizuoka, Ichihara, Matsue, and Maisaka. An important finding was that in order to account for the accent patterns, reference to both the p-stem as well as paradigm uniformity are necessary. Also, it was shown that a foot-based analysis to accent in verbs enables us to account for the accent location in a unified way. More concretely, I showed that the default accent pattern in all dialects can be formulated in terms of a trochaic foot that is right-aligned with the p-stem. The differences in accent patterns are related to the differences in the way in m-stems are projected onto p-stems. The evidence for the interaction between these two constituents is interesting, because they must both be assumed to be dominated by the p-word.

Related to this, in the analysis of Kanazawa Japanese in Chapter 5, it was argued that Sino-Japanese roots are mapped onto feet (as in Ito and Mester 2012). This means that the distinction between interface constituents and rhythmic constituents (Inkelas 1989, 1993) is blurred. This is no problem, however, because the p-word was shown to dominate both the p-stem and the foot. Therefore, rather than keeping the two hierarchies strictly separated, it may be better to look at them as an integrated whole, as in (1). Relations of dominance are indicated by straight lines. The arrows between the p-stem and the foot level indicate that these levels are related by edge alignment.

(1) 

Although the p-stem was shown to be necessary in the analysis of verbal accent, the question of the potential role the p-stem in other parts of the grammar needs further exploration, as does the role of the p-stem in dialects other than the ones discussed in Chapter 6.

The evidence for a role for the foot in the accentual and tonal systems of the different dialects suggests that a foot-based approach to Tokyo Japanese is on the right track. It is hoped that at least some of the ideas developed in this thesis will help to solve the remaining issues in the accentuation of Tokyo Japanese and other dialects.

Over the course of this thesis, I have shown that by studying variation across closely related dialects, new light can be shed on both theoretical and typological issues. By studying different dialects at the same time, we can gain a better view of the language as a whole. In this particular case, I have shown that the same type of prosodic structure, regulated by the same kinds of constraints, plays an important role across dialects. Differences across the dialects emerge from the different ways in which the prosodic constituents are aligned with morpho-syntactic constituents, as well as differences in the alignment of accentual and phrasal tones within the different constituents.

References

Abe, Y. (1981). Tone harmony schema: a theory of pitch accent. In: D. Wheeler and W. Chao (eds.), *University of Massachusetts Occasional Papers*. 1-26.

Akinaga, K. (1966). Sanagi akusento no teiki suru mono [Problems related to Sanagi accent]. *Kokubungaku Kenkyū.* 33. 74-88.

Akinaga K. and K. Sakamoto (2010). "Shin Meikai Nihongo Akusento Jiten" kara no hōkoku [A report on the "New Meikai Japanese Accent Dictionary"]. Tokyo: Akusento Shiryô Kenyûkai.

Akinlabi, A. and M. Liberman (2001). Tonal complexes and tonal alignment. In: M-J. Kim and U. Strauss (eds.), *NELS 31*. Amherst, MA: GLSA. 1-20.

Alderete, J. (1995). Faithfulness to prosodic heads. Ms., University of Massachusetts, Amherst.

Alderete, J. (1999). Morphologically governed accent in Optimality Theory. University of Massachusetts, Amherst doctoral dissertation.

Alderete, J. (2001). Dominance effects as transderivational faithfulness. *Phonology* 18. 201-253.

Amano, S. and T. Kondo (1999). *Nihongo no goi tokusei* [Lexical properties of Japanese]. Tokyo: Sanseido.

Anderson, J.M. and C. Jones (1974). Three theses concerning phonological representations. *Journal of Linguistics* 10. 1-26.

Anderson, J.M. and C.J. Ewen (1987). *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Archangeli, D. (1981) Diabolus in Japonica: Tree, tone, and accent. Talk given at Trilateral Conference on Formal Phonology, University of Texas, Austin.

Archangeli, D. and D. Pulleyblank (1984). *Extratonality and Japanese accent*. Ms, MIT and McGill University.

Archangeli, D. and D. Pulleyblank (1994). *Grounded phonology*. Cambridge, MA/London: The MIT Press.

Arisaka, H. (1941). Akusento no kata no honshitsu ni tuite [On the nature of accent patterns]. *Gengo Kenkyū* 7. 83-92.

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.

Asano, M. (2002). Studies in Japanese prosody. Harvard University doctoral dissertation.

Ayusawa, T., S. Kitazawa and S. Toki (1998). A study of Japanese rhythm by means of TEMAX and its application in language teaching [in Japanese]. *Journal of the Phonetic Society of Japan* 2:1. 34-40.

Backley, P. and K. Nasukawa (2013). The role of L in the pitch accent system of Tokyo Japanese. In: The Phonological Society of Japan (ed.), *Phonological Studies* 17. 37-44.

Baković, E. (2011). Opacity and ordering. In: John Goldsmith, Jason Riggle and Alan Yu (eds.), *The handbook of phonological theory* (2nd edition). Malden, MA: Wiley-Blackwell. 40-67.

Becker, M. and P. Jurgec (to appear). Interactions of tone and ATR in Slovenian. In: W. Kehrein and B. Köhnlein (eds.), *Tone and segmental structure*. Tübingen: Niemeyer.

Beckman, J. (1998). Positional faithfulness. University of Massachusetts, Amherst doctoral dissertation.

Beckman, M. (1982). Segment duration and the mora in Japanese. *Phonetica* 39. 113-135.

Beckman, M. (1986). *Stress Accent and Non-Stress Accent*. Dordrecht: Floris Publications.

Bekku, S. (1977). *Nihongo no rizumu* [The rhythm of Japanese]. Tokyo: Kôdansha.

Bennett, R.T. (2012). Foot-conditioned phonotactics and prosodic constituency. University of California, Santa Cruz doctoral dissertation.

Berg, T. and C. Koops (2014). Phonotactic constraints and sub-syllabic structure: A difficult relationship. *Journal of Linguistics*.
DOI: http://dx.doi.org/10.1017/S002222671400022X

Bermúdez-Otero, R. (1999). Constraint interaction in language change: quantity in English and Germanic [Opacity and globality in phonological change]. University of Manchester and Universidad de Santiago de Compostela doctoral dissertation.

Bermúdez-Otero, R. (2011). Cyclicity. In: M. van Oostendorp, C.J. Ewen, E. Hume and K. Rice (eds.), *The Blackwell companion to phonology*. Malden, MA: Wiley-Blackwell. 2019-2048.

Bermúdez-Otero, R. (2012). The architecture of grammar and the division of labor in exponence. In: J. Trommer (ed.), *The morphology and phonology of exponence*. Oxford: Oxford University Press. 8-83.

Bermúdez-Otero, R. (forthcoming). *Stratal Optimality Theory*. Oxford: Oxford University Press.

Bickel, B., K.A. Hildebrandt and R. Schiering (2009). The distribution of phonological word domains: a probabilistic typology. In: J. Grijzenhout and B. Kabak (eds.), *Phonological domains: Universals and deviations*. Berlin: Walter de Gruyter. 15-46.

Bickmore, L.S. (1996). Bantu tone spreading and displacement as alignment and minimal misalignment. Ms., University of Albany. ROA# 161.

Bloch, B. (1950). Studies in colloquial Japanese IV: Phonemics. *Language* 26. 86-125.

Boersma, P. (1998). Functional phonology: Formalizing the interactions between

articulatory and perceptual drives. Univerity of Amsterdam Phd dissertation.
Boersma, P. and D. Weenink (2011). *Praat, doing phonetics by computer*. Version 5.2.41. www.praat.org.
Booij, G. (2010). *Construction morphology*. New York: Oxford University.
Bruce, G. (1977). *Swedish word accents in sentence perspective.* Lund: Gleerup.
Buckley, E. (2009). Locality in metrical typology. *Phonology* 26. 389-435.
Buckley, E. (2013). Kashaya extrametricality and formal symmetry. *Supplemental Proceedings of Phonology 2013*. Phonology 2013, University of Massachusetts at Amherst, November 9-10.
Burzio, L. (2002). Surface-to-surface morphology: when your representations turn into constraints. In P. Boucher (ed.), *Many morphologies*. Cascadilla Press. 142-177.
Burzio, L. (2005a.). Sources of paradigm uniformity. In: L.J. Downing, T.A. Hall & R. Raffelsiefen (eds.), *Paradigms in phonological theory.* New York: Oxford University Press. 65-106
Burzio, L. (2005b). Lexicon and grammar: Equal but inseperable. Ms., Johns Hopkins University.
Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
Byun, H-G (2012). Nihongo kyōbo'in no museika: Kyōtsūgo fukyū no shihyō to shite [High vowel devoicing in Japanese: As an indicator of the standardization of Dialects]. University of Tokyo doctoral dissertation.
Cassimjee, F. and C.W. Kisseberth (1998). Optimal domains theory and Bantu tonology. In: C. Kisserberth and L.M. Hyman (eds.), *Theoretical aspects of Bantu tone*. Stanford, CA: CSLI. 33-132.
Clark, M. (1986). Japanese as a tone language. In: T. Imai and M. Saito (eds.), *Issues in Japanese linguistics*. Dordrecht: Foris. 53-105.
Clements, G.N. (1985). The geometry of phonological features. *Phonology Yearbook* 2. 225-252.
Cole, J. (1995). The cycle in phonology. In: Goldsmith, John (ed.), The handbook of phonological theory. Cambridge, MA/Oxford: Blackwell. 70-113.
Coleman, J. (1998). Phonological representations: Their names, forms, and powers. Cambridge: Cambridge University Press.
Connell, B. (2011). Downstep. In: M. van Oostendorp, C.J. Ewen, E. Hume and K. Rice (eds.), *The Blackwell companion to phonology*. Oxford: Blackwell. 824-847.
Crowhurst, M. and M.S. Hewitt (1995a). Prosodic overlay and headless feet in Yidin^{y}. *Phonology* 12:1. 39-84.
Crowhurst, M. and M.S. Hewitt (1995b). Directional footing, degeneracy, and alignment.

In: J. Beckman (ed.), *Proceedings of the 25th Annual Meeting of the North East Linguistics Society*. Amherst, Mass.: GLSA.

Crowhurst, M.J. and Mark Hewitt (1997). Boolean operations and constraint interactions in Optimality Theory. Ms., University of North Carolina and Brandeis University. Available online at [ROA-229].

Cutler, A. and T. Otake (1999). Pitch accent in spoken word recognition in Japanese. Journal of the Acoustic Society of America 105:3. 1877-1888.

Czaykowska-Higgins, E. (1998). The morphological and phonological constituent structure of words in Moses-Columbia Salish (Nxaʔamxcin). In: E. Czaykowska-Higgins and M.D. Kinkade (eds.), *Salish languages and linguistics: Theoretical and descriptive perspectives*. Berlin: Mouton de Gruyter. 153-195.

Davis, S. (2005). Capitalistic vs. militaristic: the paradigm uniformity effect reconsidered. In: L.J. Downing, T.A. Hall and R. Raffelsiefen (eds.), *Paradigms in phonological theory*. Oxford/New York: Oxford University Press.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. [Revised and Expended Edition]. New York: Oxford University Press. First edition published in 1997.

Downing, L.J. (1998a). Prosodic misalignment and reduplication. In: Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology* 1997. 83-120. Dordrecht: Kluwer.

Downing, L.J. (1998b). On the prosodic misalignment of onsetless syllables. *Natural Language and Linguistic Theory* 16. 1-52

Downing, L.J. (1999). Prosodic stem ≠ prosodic word in Bantu. In: T.A. Hall & U. Kleinhenz (eds.), *Studies on the phonological word*. Amsterdam: John Benjamins. 73-98.

Downing, L.J. (2006). *Canonical forms in prosodic morphology*. Oxford/New York: Oxford University Press.

Dresher, B.E. and H. van der Hulst (1998). Head-dependent asymmetries in phonology: complexity and visibility. *Phonology* 15:3. 317-352.

Elenbaas, N. and R. Kager (1999). Ternary rhythm and the lapse constraint. *Phonology* 16. 273-329.

Everett, D.L. (1997). Syllable integrity. In: *Proceedings of the Westcoast Conference on Formal Linguistics* 16. 177-190.

Fox, A. (2000). *Prosodic structure and prosodic features: The phonology of 'suprasegmentals'.* Oxford: Oxford University Press.

Fudge, E.C. (1969). Syllables. *Journal of Linguistics* 5. 253-286.

Fujiwara, F. (1972). Chiba-ken Ichihara-shi nanbu hōgen no akusento ni tsuite [On the accent of the southern dialect of Ichihara, Chiba Prefecture]. *Kokugo Kenkyū* 34. 2-34.

Fujiwara, F. (1979). Ichihara-shi no hōgen [The dialect of Ichihara]. *Ichihara shishi bekkan nukizuri* [Offprint of an extra issue of The History of Ichihara City].

Garrett, E. (1999). Minimal words aren't minimal feet. *UCLA Working Papers in Linguistics* 1. 68-105.

Giriko, M. (2009). Nihongo ni okeru go ninshiki to heibangata akusento [Word Recognition and the Unaccented Pattern in Japanese]. Kobe University doctoral dissertation.

Goldsmith, J. (1976). Autosegmental Phonology. MIT doctoral dissertation.

Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.

Goldsmith, J. & F. Sabimana (1986). The Kirundi verb. In: F. Jouannet (ed.), *Modèles en tonologie*. Paris: Editions du CNRS. 19-62.

Gomez-Imbert, E. and M. Kenstowicz (2000). Barasana tone and accent. *International Journal of American Linguistics* 66:4. 419-463.

Gordon, M. (2002). A factorial typology of quantity insensitive stress. *Natural Language and Linguistic Theory* 20. 491-552.

Gordon, M. (2006). *Syllable weight: Phonetics, phonology, typology.* New York/London: Routledge.

Green, A.D. (1997). The Prosodic Structure of Irish, Scots Gaelic and Manx. Cornell University doctoral dissertation,

Gussenhoven, C. (2004). *The Phonology of tone and intonation*. Cambridge Univ. Press.

Gussenhoven, C. (to appear, a). Suprasegmentals. In J.D. Wright (ed.), *International encyclopedia of social and behavioral sciences*. Second Edition. Oxford: Elsevier.

Gussenhoven, C. (to appear, b). On the privileged status of intonational boundary tones: Evidence from Japanese, French, and Cantonese English. In G. Buckley, T. Crane and J. Good (eds.), *Revealing structure: Finding patterns in grammars and using grammatical patterns to elucidate language. A festschrift to honor Larry M. Hyman.* Stanford: CSLI.

Hagberg, L.R. (1993). An Autosegmental theory of stress. Doctoral dissertation, University of Arizona. Published as e-Book by SIL International, 2006.

Halle, M. (1982). Remarks on Metrical Theory. Talk given at GLOW conference, Paris.

Halle, M. and W.J. Idsari (1995). General properties of metrical structure. In: J. Goldsmith (ed.), *The handbook of phonological theory*. Cambridge, MA: Blackwell.

Halle, M., and J.-R. Vergnaud (1987). *An essay on stress.* Cambridge, Mass.: MIT Press.

Hamano, S.S. (1986). The sound-symbolic system of Japanese. University of Florida dissertation.

Han, E. (1994). Prosodic Structure in Compounds. Stanford University dissertation.

Han, M.S. (1962a). The feature of duration in Japanese. *Onsei no Kenkyū* 10. 65-80.
Han, M.S. (1962b). Unvoicing of vowels in Japanese. *Onsei no Kenkyū* 10. 81-100.
Han, M.S. (1994). Acoustic manifestations of mora timing in Japanese. *The Journal of the Acoustic Society of America* 96. 73-82.
Haraguchi, S. (1977). *The tone pattern of Japanese: An autosegmental theory of tonology*. Tokyo: Kaitakusha.
Haraguchi, S. (1984). Some tonal and segmental effects of vowel height in Japanese. In: M. Aronoff et al. (eds.), *Language and sound structure*. Cambridge/London: MIT Press. 145-156.
Haraguchi, S. (1988). Pitch accent and intonation in Japanese. In: H. van der Hulst and N. Smith (eds.), *Autosegmental studies on pitch accent*. Dordrecht: Foris Publications. 123-150
Haraguchi, S. (1991). *A theory of stress and accent*. Dordrecht: Foris Publications.
Haraguchi, S. (1999). Accent. In: N. Tsujimura (ed.). *The handbook of Japanese Linguistics.* Malden, MA and Oxford: Wiley-Blackwell. 1-30.
Haraguchi, S. (2001). The accent of Tsuruoka Japanese reconsidered. In: J.M. van de Weijer and T. Nishihara (eds.), *Issues in Japanese phonology and morphology*. Berlin: Walter de Gruyter. 47-65.
Haraguchi, S. (2003). The phonology-phonetics interface and syllable theory. In: J. van de Weijer, V.J. van Heuven and H. van der Hulst (eds.), *The phonological spectrum. Volume II: Suprasegmental structure*. Amsterdam: John Benjamins. 31-58.
Hasegawa, Y. (1999). Pitch accent and vowel devoicing in Japanese. *Proceedings of the International Congress of Phonetic Sciences*. 523-526. Available as PDF file at http://ist-socrates.bereley.edu/~hasegawa/pub.html.
Hashimoto, D. (2012). Tanjungo o kihon to shite motsu tanshukugo keisei to inritsu kōzō [Truncation and the prosodic structure of simplex words]. *Nihon Gengogakkai Dai 145-kai Taikai Yokōshū*. 46-51.
Hattori, S. (1954/1955). On'inron kara mita kokugo no akusento [Japanese accent as seen from phonology] *Kokugo Kenkyū* 2-3. Reprinted in: M. Tokugawa (ed.) (1980), *Ronshū Nihongo Kenkyū 2: Akusento.* Tokyo: Yūseidoo.
Hattori, S. (1961). Prosodeme, syllable structure, and laryngeal phonemes. *Bulletin of the Summer Institute in Linguistics* (ICU) 1. 1-27.
Hattori, S. (1967). Descriptive linguistics in Japan. In: T.A. Sebeok (ed.), *Current trends in linguistics, vol. 2: Linguistics in East Asia and Southeast Asia*. The Hague: Mouton. 530-584.
Hayata, T. (1997). Akusento to tōn: inritsu no ippanron ni mukete [Accent and tone:

toward a general theory of prosody]. *Ajia, Afurika Bunpoo Kenkyū* 26. 1-22
Hayata, T. (1999a). Accent and Tone: Towards a General Theory of Prosody. In: S. Kaji (ed.), *Cross-Linguistic studies of tonal phenomena: Tonogenesis, typology, and related topics*. Tokyo: ILCAA. 221-234.
Hayata, T. (1999b). *Onchō no taiporojii* [The Typology of Tone]. Tokyo: Taishūkan.
Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253–306.
Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: The University of Chicago Press.
Hirayama, T. (1952). Hokurikudō hōgen no onchō [Tone in the Hokuriku dialects]. In Terakawa, K. et al. (eds.), *Kokugo akusento ronsō*. Tokyo: Hôsei Daigagku Shuppankyoku.
Hirayama, T. (1960). *Zenkoku akusento jiten* [National accent dictionary]. Tokyo: Tōkyōdō Shuppan.
Hirayama, T. et al. (eds.) (1974). *Hōgen taikei henka no tsūjironteki kenkyū* [A Diachronic Study of a Dialect System]. Tokyo: Meiji Shoin.
Hirayama, M. (2009). Postlexical prosodic structure and vowel devoicing in Japanese. University of Toronto doctoral dissertation.
Hiroto, A. and T. Ōhara. (1952). *San'in chihō no akusento* [Accent in the San'in region]. Hōkōsha. Unive
Hiroto, A. (1986). *Hōgen goi no kenkyū* [A study of dialect vocabulary]. Tokyo: Kazama Shobō.
Hombert, J.-M. (1977). Development of tones from vowel height. *Journal of Phonetics* 5. 9-16.
Hombert, J.-M. (1978). Consonant types, vowel quality and tone. In: Fromkin,Victoria A. (ed.), *Tone: A linguistic survey*. New York: Academic Press. 77-111.
Hombert, J.-M.., J.J. Ohala & W.G. Ewan (1979). Phonetic explanations for the development of tones. *Language* 55. 37–58.
Homma, Y. (1981). Duration relationship between Japanese stops and vowel. *Journal of Phonetics* 9. 273-281.
Hori, H. and Takigawa F. (eds.) (2014). *Shizuoka-ken chiyoda hōgen no kijutsu kenkyū* [A descriptive study of the Shizuoka dialect of Japanese]. Shizuoka: Shizuoka Daigaku Jinbun Shakai Kagakubu Gengo Bunka Gakka.
Houghton, P.S. (2013). Switch languages: Theoretrical consequence and empirical reality. Rutgers, The State University of New Jersey doctoral dissertation.
Hulst, H. van der (2009). Brackets and grid marks, or theories of primary accent and

rhythm. In: E. Raimy and C. Cairns (eds.), *Contemporary views on architecture and representations in phonological theory*. Cambridge, MA: MIT Press. 225-245.

Hulst, H. van der (2011). Pitch accent systems. In: M. van Oostendorp, C.J. Ewen, E. Hume and K. Rice (eds.), *The Blackwell companion to phonology*. Oxford: Wiley-Blackwell. 1003-1026.

Hyde, B. (2012). Alignment constraints. *Natural Language and Linguistic Theory* 30. 789-836.

Hyman, L.M. (1993). Conceptual issues in the study of the Bantu verb stem. In: S.S. Mufwene and I. Moshi (eds.), *Topics in African linguistics*. Amsterdam: John Benjamins. 3-34.

Hyman, L.M. (2007). Universals of tone rules: 30 years later. In: C. Gussenhoven and T. Riad (eds.), *Tones and tunes, vol. 1: Studies in word and sentence prosody*. Berlin: Mouton de Gruyter.

Hyman, L.M. (2008). Directional asymmetries in the morphology and phonology of words, with special reference to Bantu. *Linguistics* 46, 309–350

Hyman, L.M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31: 213-238.

Hyman, L.M. (2012). Phonological theory vs. description. LabEx Lecture #2. Paris, Sept. 24, 2012. http://labex-efl.univ-paris3.fr/Hyman%20LabEx%20Lecture%202.pdf [Accessed at 2014-05-11].

Idsardi, W.J. (1992). The computation of prosody. Doctoral dissertation, MIT.

Idsardi, W.J. and Kim, S.-H. (1997). Enclitic tonal patterns in North Kyungsang Korean and Tokyo Japanese. *Proceedings of the Fourth Seoul International Conference on Linguistics*. 305-314.

Idsardi, W.K. and T.C. Purnell (1997). Metrical tone and the elsewhere condition. *Rivista di Lingusitica* 9:1. 129-156.

Igarashi, Y. (2012). Prosodic typology in Japanese dialects from a cross-linguistic perspective. *Lingua* 122:13. 1441-1453.

Igarashi, Y. (2014). Typology of intonational phrasing in Japanese dialects. To appear in Sun-Ah Jun (eds.), *Prosodic typology II*. New York: Oxford University Press. 464-492.

Inaba, S. (1998). Moras, syllables, and feet in Japanese. *Language, Information, and Computation: Proceedings of The 12th Pacific Asia Conference*. 106-117.

Inkelas, S. (1989). Prosodic constituency in the lexicon. Stanford University doctoral dissertation. Published by Garland, New York in 1990.

Inkelas, S. (1993). Deriving cyclicity. In: S. Hargus and E.M. Kaisse (eds.), *Studies in Lexical Phonology. Phonetics and phonology 4*. San Diego: Academic Press. 75-100.

Inkelas, S. (1995). The consequences of optimization for underspecification. *NELS* 25. 287-302.

Inkelas, S. (1998). The theoretical status of morphologically conditioned phonology: a case study from dominance. *Yearbook of morphology* 1997:121-155.

Inkelas, S. (2007). Is grammar dependence real? A comparison between cophonological and indeced constraint approaches to morphologically conditioned phonology. *Linguistics* 45:1. 133-171.

Inoue, H. (1968). Toohoku hōgen no shi'in taikei [Consonant system of the Tohoku dialects]. *Gengo Kenkyū* 52. 80-98.

Ishihara, T. (2006). Tonal alignment in Tokyo Japanese. University of Edinburgh doctoral dissertation.

Ito, C. (2013). Korean accent: Internal reconstruction and historical development. *Korean Linguistics* 15:2. 129-198.

Ito, J. (1990). Prosodic minimality in Japanese. In: *CLS 26-II. Papers from the Parasession on the Syllable in Phonetics and Phonology*. 213-239.

Ito, J. and A. Mester (1992). Weak layering and word binarity. Ms., Linguistic Research Center, University of California, Santa Cruz. Published in T. Honma, M. Okazaki, T. Tabata, and S. Tanaka (eds.), 2003. *A new century of phonology and phonological theory. A festschrift for Professor Shosuke Haraguchi on the occasion of his sixtieth birthday*. Tokyo: Kaitakusha. 26-65.

Ito, J. and A. Mester (1996). Stem and word in Sino-Japanese. In: T. Otake and A. Cutler (eds.), *Phonological structure and language processing: Cross-linguistic studies*. Berlin: Mouton de Gruyter. 13-44.

Ito, J. and A. Mester (2003). *Japanese morphophonemics: Markedness and word structure*. Cambridge, MA: MIT Press.

Ito, J. and A. Mester (2007). Prosodic adjunction in Japanese compounds. In: Y. Miyamoto and M. Ochi (eds.), *Formal approaches to Japanese linguistics: Proceedings of FAJL 4. MIT Working Papers in Linguistics*, vol. 55. MIT Department of Linguistics and Philosophy, Cambridge, MA. 97-111.

Ito, J. and A. Mester (2009). The extended prosodic word. In: J. Grijzenhout and B. Kabak (eds.), *Phonological domains: Universals and deviations.* Berlin: Mouton de Gruyter. 135-194.

Ito, J. and A. Mester (2011). Recursive prosodic phrasing in Japanese. In: W. McCure and M. den Dikken (eds.), *Japanese/Korean linguistics* 18. Stanford, CA: CSLI.

Ito, J. and A. Mester (2012). Unaccentedness in Japanese. Ms., University of Santa Cruz.

Ito, J. and A. Mester (2013). Prosodic subcategories in Japanese. *Lingua* 124. 20-40.

Ito, J. and A. Mester (2015a). Sino-Japanese Phonology. In: H. Kubozono (ed.), *Handbook of Japanese phonetics and phonology*. Mouton de Gruyter. 289-312.

Ito, J. and A. Mester (2015b). Word formation and phonological processes. In: H. Kubozono (ed.), *Handbook of Japanese phonetics and phonology*. Mouton de Gruyter. 363-396.

Ito, J., Y. Kitagawa and A. Mester (1996). Prosodic faithfulness and correspondence: evidence from a Japanese argot. *Journal of East Asian Linguistics* 5. 217-294.

Ito, J., A. Mester and J. Padgett (1995). Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26. 571-613.

Iwai, Y. (1992). Mōra onso oyobi akusento ga rizumu ni ataeru eikyō to shichi-go teikei no rizumu [The influence on rhythm by moraic phonemes and accent and the fixed 5-7 rhythm]. *Gengogaku Kenkyū* 11. 1-24.

Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.

Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Kager, R. (2005). Rhythmic licensing theory: an extended typology. In: *Proceedings of the Third Seoul International Conference on Phonology*. Seoul: The Phonology-Morphology Circle of Korea. 5-31.

Kager, R. and V. Martínez-Paricio (2014). Antepenultimate mora effects: typology and representation. Handout of a talk presented at the Workshop on Word Stress and Accent, 15-17 August 2014, Leiden University.

Kahn, D. (1976). Syllable-based generalizations in English phonology. MIT doctoral dissertation.

Katayama, M. (1998). Optimality Theory and Japanese loanword phonology. University of California, Santa Cruz dissertation.

Kato, Y. (1936). Chūō Izumo hōgen on'inkō: akusento hen [About the phonology of the Central Izumo dialect]. *Hōgen* 6:6. 30-53.

Kawahara, S. (2012). A review of The Phonology of Japanese (by L. Labrune, Oxford University Press, 2012). *Phonology* 29:3. 540-548.

Kawahara, S. (2014). Japanese has syllables: A reply to Labrune (2012). Keio University, Ms.

Kawahara, S., H. Masuda, D. Erickson, J. Moore, Y. Shibuya and A. Suemitsu (2014). Jaw displacement and metrical structure in Japanese: The effect of pitch accent, foot structure, and phrasal stress. *Journal of the Phonetic Society of Japan* 18:2. 54-62.

Kawakami, S. (1961). On the relationship between word-tone and phrase-tone in Japanese. *Onsei no Kenkyū* 9. 169-177. Reprinted in S. Kawakami (1995).

Kawakami, S. (1984). Akusento kenkyū no mondaiten [Problems in accentual studies]. Hirayama Teruo Hakusi Koki Kinenkai (eds.), *Gendai hōgengaku no kadai 2: kijutsuteki kenkyū hen*. 147-166.

Kawakami, S. (1995). *Nihongo akusento ronshû* [A Collection of Papers on Japanese Accent]. Tokyo: Kyûko Shoin.

Kenstowicz, M. (1996). Base-identity and uniform exponence: alternatives to cyclicity. In: J. Durand & B. Laks (eds.), *Current trends in phonology: Models and methods*. Manchester: University of Salford. 363-393.

Kim, N-J. (1997). Tone, segments and their interaction in North Kyungsang Korean: A correspondence theoretic account. The Ohio State University doctoral dissertation.

Kim, N-J. (2010). *Tone and prosodic phrasing in North Kyungsang Korean: An optimality theoretic account*. Saarbrücken: VDM Verlag Dr. Müller.

Kim, S.-H. (1999). The metrical computation in tone assignment. Doctoral Dissertation, University of Delaware.

Kim, S.-H. (2002). The metrical dependency of tone placement in Optimality Theory. *Studies in Phonetics, Phonology and Morphology* 8:1. 17-38.

Kinda'ichi, H. (1942). Kantō chihō ni okeru akusento no bunpu [The distribution of accent in the Kanto region]. In Nihon Hōgen Gakkai (ed.), *Nihongo no akusento*. Tokyo: Chuo Koronsha.

Kindai'ichi, H. and K. Akinaga (2001). Shinmeikai Nihongo Akusento Jiten [Shinmeikai Dictionary of Japanese Accent]. Tokyo: Sanseido.

Kiparsky, P. (1982). *Explanation in phonology*. Foris Publications, Dordrecht.

Kiparsky, P. (2000). Opacity and cyclicity. *The Linguistic Review* 17. 351-366.

Kisseberth, C.W. (2001). Bantu perspectives on Yukihiko Nakai's "The History of the Kyoto Accent". In: S. Kaji (ed.), *Proceedings of the symposium "Cross-linguistic studies of tonal phenomena: Tonogenisis, Japanese Accentology, and Other Topics"*. Tokyo: ILCAA. 145-151.

Kitagawa, Y. (1986). Subjects in Japanese and English. University of Massachusetts, Amherst doctoral dissertation.

Kitahara, M. (1998). The interaction of pitch accent and vowel devoicing in Tokyo Japanese. In D.J. Silva (ed.), *Japanese/Korean Linguistics 8*. Stanford, CA: CSLI Publications. 303–315.

Kitahara, M. (2001). Category structure and function of pitch accent in Tokyo Japanese. Indiana University doctoral dissertation.

Kitazawa, S., H. Sugiura, D. Shitaoka and S. Kobayashi (1996). Studies with the TEMAX for assessment of the rhythmic components in speech. IEICE Techincal Report, SP 86

(364). 51-58.
Klamer, M. (1998). *A grammar of Kambera*. Berlin: Mouton de Gruyter.
Komatsu, M. and T. Arai (2009). Modulation spectrum and rhythmic units of Japanese. *Journal of the Phonetic Society of Japan* 13:3. 85-89.
Kondo, M. (2006). Within-word prosodic constraint on coarticulation in Japanese. *Language and Speech* 49:3. 393-416.
Kondo, M. and T. Arai (1998). Prosodic constraint on V-to-V coarticulation in Japanese: A case of bimoraic foot. In: The Phonological Society of Japan (ed.), *Phonological Studies* 1. Tokyo: Kaitakusha. 75-82.
Krämer, M. (2009). *The phonology of Italian*. Oxford/New York: Oxford University Press.
Krämer, M. (2012). *Underlying representations*. Cambridge: Cambridge University Press.
Kubozono, H. (1985). Speech errors and syllable structure. *Linguistics and Philology* 6. 220-243.
Kubozono, H. (1988). The organization of Japanese prosody. University of Edinburgh doctoral dissertation. Published by Kurosio, Tokyo, 1993.
Kubozono, H. (1994). Nihongo no onsetsuryō ni tsuite [On syllable weight in Japanese]. *Kokugogaku* 178. 7-17.
Kubozono, H. (1995). Constraint interaction in Japanese phonology: evidence from compound accentuation. *Phonology at Santa Cruz* 4. 21-38.
Kubozono, H. (1996). Syllable and accent in Japanese: Evidence from loanword accentuation. *The Bulletin*, The Phonetic Society of Japan 211. 71-82.
Kubozono, H. (1997). Lexical markedness and variation: Anonderivational account of Japanese compound accent. *Proceedings of the 15th West Coast Conference on Formal Linguistics*. 273-287.
Kubozono, H. (1998). Mōra to onsetsu no fuhensei [On the universality of mora and syllable]. *Journal of the Phonetic Society of Japan* 2:1. 5-15.
Kubozono, H. (1999). Mora and syllable. In: N. Tsujimura (ed.), *The handbook of Japanese linguistics*. Oxford: Blackwell. 31-61.
Kubozono, H. (2003). The syllable as a unit of prosodic organization in Japanese. In: C. Féry and R. van de Vijver (eds.), *The syllable in Optimality Theory.* Cambridge: Cambridge University Press. 99–122.
Kubozono, H. (2004). On'in kōzō kara mita tanjungo to gōseigo no kyōkai [The boundary between simplex and complex words as seen from phonological structure]. In: Onsei Bunpō Kenkyūkai [Spoken Language Working Group] (eds.), *Onsei to bunpō* IV. 123-143.
Kubozono, H. (2005). Rendaku: Its domain and linguistic conditions. In: J. van de Weijer

and T. Nishihara (eds.), *Voicing in Japanese*. Berlin: Walter de Gruyter. 5-24.
Kubozono, H. (2006). Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua* 116. 1140-1170.
Kubozono, H. (2008). Japanese Accent. In S. Miyagawa and M. Saito (eds.), *The Oxford Handbook of Japanese Linguistics*. Oxford: Oxford University Press.
Kubozono, H. (2010). Gokeisei to on'in kōzō: tanshukugo keisei no mekanizumu [Word formation and phonological structure: the mechanism of word truncation]. *NINJAL Project Review* 3. 17-34.
Kubozono, H. (2011). Japanese pitch accent. In: M. van Oostendorp, C.J. Ewen, E. Hume and K. Rice (eds). *The Blackwell companion to phonology*. Malden, MA & Oxford: Wiley-Blackwell. 2879-2907.
Kubozono, H. (2012). Varieties of pitch accent systems in Japanese. *Lingua* 122:3. 1395-1414.
Kubozono, H., J. Itō, and A. Mester (1997). On'in kōzō kara mita go to ku no kyōkai: fukugō-meishi no akusento no bunseki [The difference between words and phrases from the point of view of phonological structure: an analysis of compound nouns]. *Bunpō to Onsei*. 147-166.
Kubozono, H. and S. Ogawa (2004). Emergence of unaccented words in Japanese. *Proceedings of Speec Prosody* 2004. 33-36.
Kurisu, K. (1994). Further evidence for bi-moraic foot in Japanese. *Proceedings of the International Conference on Spoken Language Processing, Yokohama* 1. 367-370.
Kurisu, K. (2000). *Richness of the base and root fusion in Sino-Japanese*. Journal of East Asian Linguistics 9. 147-185.
Labrune, L. (2002) The prosodic structure of simple abbreviated loanwords in Japanese: A constraint-based account. *Journal of the Phonetic Society of Japan* 6:1. 98–120.
Labrune, L. (2006). *La phonologie du Japonais*. Leuven/Paris: Peeters.
Labrune, L. (2012a). *The Phonology of Japanese*. Oxford: Oxford University Press.
Labrune, L. (2012b). Questioning the universality of the syllable: evidence from Japanese. *Phonology* 29:1. 113-151.
de Lacy, P. (2007). The interaction of tone, sonority, and prosodic structure. In: P. de Lacy, (ed.), *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press. 281-307.
Ladd, D.R. (2006). Segmental anchoring of pitch movements: Autosegmental association or gestural coordination? *Rivista di Linguistica* 18:1. 19-38.
Lahiri, A. (2001). Metrical patterns. In: E. König and M. Haspelmath (eds.), *Language Typology and Language Universals*. Berlin: Mouton. 1347-1367.

Lawrence, W.P. (1990). Nakijin phonology: Feet and extrametricality in a Japanese dialect. University of Tsukuba dissertation.

Leben, W.R. (1996). Tonal feet and the adaptation of English borrowings into Hausa. *Studies in African Linguistics* 25:2. 139-154.

Lee, D. (2005). Weight-sensitive tone patterns in loanwords of South Kyungsang Korean. *Harvard studies in Korean linguistics* 11. 197-210.

Lee, Y. and M. Goldrick (2008). The emergence of sub-syllabic representations. *Journal of Memory and Language* 59. 155-168.

Lee, H. and J. Zang (2014). The nominal pitch accent of South Kyungsang Korean. *Journal of East Asian Linguistics* 23. 73-111.

Lee, S.J. (2006). Tonal complex: Building structures around H tone. In: S. Kuno et al. (eds.), *Harvard studies in Korean linguistics XI*. Seoul: Hanshin Publishing Co. 224-236.

Levin, J. (1985). A metrical theory of syllabicity. MIT doctoral dissertation.

Liberman, M. (1975). The intonational system of English. MIT doctoral dissertation.

Liberman, M. and A. Prince (1977). On stress and linguistic rhythm. *Linguistic Inquiry* 8:2. 249-336.

Maekawa, H. (1970). Shizuoka hōgen no akusento [Accent in the Shizuoka dialect]. In Hirayama Teruo Hakase Kanreki Kinenkai (eds.), *Hōgen kenkyū no mondaiten* [Problems in Dialectology]. Tokyo: Meiji Shoin. 173-198.

Maekawa, K. (1990). Production and perception of the accent in the consecutively devoiced syllables in Tokyo Japanese. In: *Proceedings of ICSLP*, Volume 90.1. 517-520.

Maekawa, K. (2010). Final lowering and boundary pitch movements in spontaneous Japanese. *Proceedings of DiSS-LPSS Joint Workshop 2010*. (Tokyo, 25-26 September 2010). 47-50.

Maekawa, K. and Y. Igarashi (2007). Prosodic phrasing of bimoraic accented particles in spontaneous Japanese. In: *Proceedings of ICPhS XVI*. Saarbrücken, 6-10 August, 2007. 1217-1220.

Marlo, M.R., L.C. Mwita and M. Paster (2014). Problems in Kuria H tone assignment. *Natural Language and Linguistic Theory.* DOI 10.1007/s11049-014-9251-y.

Martin, S.E. (1952). *Morphophonemics of standard colloquial Japanese*. Baltimore: Linguistic Society of America.

Martin, S.E. (1975). *A reference grammar of Japanese.* New Haven: Yale University Press.

Martínez-Paricio, V. (2013). An exploration of minimal and maximal metrical feet. University of Tromsø doctoral dissertation.

McCarthy, J.J. (1979). On stress and syllabification. *Linguistic Inquiry* 10:3. 443-465.
McCarthy, J.J. (1993). Template form in Prosodic Morphology. *Proceedings of FLSM II*. Bloomington, Ind.: IULC. 187-218.
McCarthy, J.J. and A. Prince (1994). The emergence of the unmarked: Optimality in prosodic morphology. *Proceedings of the North East Linguistics Society* 24. Amherst, MA: GLSA. 333-379.
McCarthy, J.J. (1999). Sympathy and phonological opacity. *Phonology* 16. 331-399.
McCarthy, J.J. (2000). Faithfulness and prosodic circumscription. In: J. Dekkers, F. van der Leeuw, and J. van de Weijer (eds.), *Optimality Theory: Phonology, Syntax, and Acquisition*. Oxford: Oxford University Press. 151-189.
McCarthy, J.J. & A.S. Prince (1986). Prosodic morphology. Ms., University of Massachusetts, Amherst, MA and Brandeis University, Waltham, MA.
McCarthy, J.J. and A. Prince (1993). Prosodic morphology: Constraint Interaction and Satisfaction. Report no. RuCCS-TR-3. New Brunswick, NJ: Rutgers University Center for Cognitive Science. [ROA #482–1201]
McCarthy, J.J. and A. Prince (1995). Faithfulness and reduplicative identity. In: J.N. Beckman, L.Walsh-Dickey (eds.), *University of Massachusetts Occasional Papers in Linguistics 18*: *Papers in Optimality Theory*. 249-384.
McCarthy, J.J. and A. Prince (1999). Faithfulness and identity in prosodic morphology. In: R. Kager, H. van der Hulst, and W. Zonneveld (eds.), *The prosody-morphology interface*. Cambridge: Cambridge University Press. 218-309.
McCawley, J.D. (1968). *The phonological component of a grammar of Japanese*. The Hague: Mouton.
McCawley, J.D. (1978). What is a tone language? In: V.A. Fromkin (ed.), *Tone: A linguistic survey*. New York/San Fransisco/London: Academic Press.
Mester, R.A. (1990). Patterns of truncation. *Linguistic Inquiry* 21:3. 478-485.
Mester, R.A. and J. Ito (1989). Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65. 258-293.
Miyata, K. (1927). Atarashii akusento-kan to akusento hyōkihō [A new view of accent and means of representing accent]. *Onsei no Kenkyū* 1. 18-22.
Miyata, K. (1928). Nihongo no akusento ni kansuru watakushi no kenkai [My views on Japanese accent]. *Onsei no Kenkyū* 2. 31-37.
Mortenson, D.R. (2013). Tonally conditioned vowel raising in Shuijinging Mang. *Journal of East Asian Linguistics* 22. 189-216.
Mutaka, N.M. (1994). *The lexical tonology of Kinande*. Munich: Lincom Europa.
Mutsukawa, M. (2009). *Japanese loanword phonology: The nature of inputs and the*

loanword sublexicon. Tokyo: Hituzi.
Myers, S. (1987). Tone and the structure of words in Shona. University of Massachusetts, Amherst doctoral dissertation.
Myers, S. (1997). OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15. 847-892.
Myers, S. (2000). Boundary disputes: the distinction between phonetic and phonological sound patterns. Burton-Roberts, N., P. Carr and G. Docherty (eds.), *Phonological knowledge: Conceptual and empirical issues*. Oxford: Oxford University Press. 245-272.
Nagahara, H. (1994). Phonological Phrasing in Japanese. UCLA doctoral dissertation.
Nasu, A. (2002). Nihongo onomatope no gokeisei to inritu koozoo. [Word formation and prosodic structure of Japanese mimetics]. University of Tsukuba doctoral dissertation.
Nasu, A. (2009). Tokushu mōra no bunsetsu kōzō to anteido [Given English translation: Segmental structure and stability of subordinate moras in Japanese]. *Bungei Gengo Kenkyū. Gengohen* 56. 53-71.
Neeleman, A. and H. van de Koot (2006). On syntactic and phonological representations. *Lingua* 116:10. 1524-1552.
Nespor, M. and I.Vogel (1986). *Prosodic phonology*. Dordrecht: Foris.
Nespor, M. and I. Vogel (1989). On clashes and lapses. *Phonology* 6. 69-116.
Nesset, T. (2008). *Abstract phonology in a concrete model: Cognitive linguistics and the morphology-phonology interface*. Berlin/New York: Mouton de Gruyter.
Neustupný, J.V. (1966). Nihongo akusento wa kōtei akusento ka [Is Japanese accent a pitch accent?]. *Onsei Gakkai Kaihō* 121. 1-7.
NHK (1985). NHK Nihongo hatsuon akusento jiten [NHK pronunciation and accent dictionary of Japanese]. Tokyo: NOgawaHK.
NHK (1998). NHK Nihongo hatsuon akusento jiten [NHK pronunciation and accent dictionary of Japanese]. Tokyo: NHK.
Nishihara, T., J. van de Weijer & K. Nanjo (2001). Against headedness in compound truncation: English compounds in Japanese. In: J. van de Weijer and T. Nishihara (eds.), *Issues in Japanese phonology and morphology*. Berlin: Mouton de Gruyter. 299-324.
Nishimura, K. (2013). Morphophonology of Japanese compounding. University of Tokyo doctoral dissertation.
Nishiyama, K. (2010). Penultimate accent in Japanese and the verb-noun distinction. *Lingua* 120:10. 2356-2366.
Nitta, T. (1989). Kanazawa hōgen no dōshi akusento [Accent of verbs in the Kanazawa dialect] (given translation). *Nihonkaiiki kenkyōsho hōkoku* 21. 149-200.
Nitta, T. (1994). Tsuruoka hōgen no akusento [Accent in the Tsuruoka dialect]. *Tsuruoka*

hōgen no kijutsuteki kenkyū: Dai 3ji Tsuruoka chōsa hōkoku 1. Tokyo: Shûei Shuppan. 81-140.

Nitta, T. (2001). The accent systems in the Kanazawa dialect: the relationship between pitch and sound segments. In: S: Kaji (ed.), *Cross-linguistic Studies of tonal phenomena: Tonogenesis, Japanese accentology, and other phenomena*. Tokyo: ILCAA. 153-185.

Oda, K. (2006). The accentuation patterns of nominal compounds in Japanese. *Toronto Working Papers in Linguistics* 26. 23-64.

Odden, D. (1994). Adjacency parameters in phonology. *Language* 70:2 289-330.

Odden, D. (2001). Comments on "The accent system in the Kanazawa dialect: the relationship between pitch and sound segments". In: S. Kaji (ed.), *Cross-linguistic studies of tonal phenomena*. Tokyo: ILCAA. 187-193.

Odden, D. (2007). Features impinging on tone. Handout of a presentation at the Workshop on Segments and Tone, Meertens Instituut & University of Amsterdam, June 7-8.

Ogawa, S. (2004). Sino-Japanese word accent and syllable structure. In: The Phonological Society of Japan (ed.), *Phonological Studies* 7. 41-48.

Ogawa, S. (2008). Accentual asymmetry in trimoraic Sino-Japanese words. In: H. Kubozono (ed.), *Asymmetries in phonology: An East-Asian perspective*. Tokyo: Kurosio. 183-191.

Okuda, K. (1975). *Accentual systems in the Japanese dialects: A generative approach*. Tokyo: Bunka Hyoron Publishing.

Okuda, K. (1980). Seisei bunpō kara mita ji no akusento [The accent of auxiliaries from the point of view of generative grammar]. *Kokugo kyōiku kenkyū*. 473-482

Oostendorp, M. van (2007). Derived environment effects and consistency of exponence. In S. Blaho, P. Bye and M. Krämer (eds.), Freedom of Analysis? Berlin: Walter de Gruyter. 123-148.

Oostendorp, M. van (2008). Variation in generative grammar. Ms., Meertens Instituut, Amsterdam.

Oostendorp, M. van (2013). *σ* strikes back: A defense of headedness and constituency in phonology. *The Linguistic Review* 30(2). 347-371.

Oostendorp, M. van (2014). Selective lexicon optimization. *Lingua* 142. 76-84.

Oostendorp, M. van and B. Hermans (1999). Introduction. In: M. van Oostendorp and B. Hermans (eds.), *The derivational residue in phonological Optimality Theory*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 1-27.

Oostendorp, M. van and J. van de Weijer (2005). Phonological alphabets and the structure of the segment. *The internal organization of phonological segments*. Berlin: Mouton de

Gruyter. 1024.
Ōhashi, Jun'ichi (2002). *Tōhoku hōgen onsei no kenkyū* [A study of the phonetics of the Tohohoku dialects]. Tokyo: Ôfû.
Orgun, C.O. (1996). Sign-Based morphology and phonology: with special attention to Optimality Theory. University of California, Berkeley doctoral dissertation.
Oshima, D.Y. (2014). On the morphological status of -te, -ta, and related forms in Japanese: Evidence from accent placement. *Journal of East Asian Linguistics* 23:3. 233-265.
Otake, T., G. Hatano, A. Cutler and J. Mehler (1993). More or syllable? Speech segmentation in Japanese. *Journal of Memory and Language* 32. 258-278.
Otake, T., G. Hatano and K. Yoneyama (1996). Speech segmentation by Japanese listerners. In: Otake, Takashi and Anne Cutler (eds.), *Phonological structure and language processing*. Berlin: Mouton de Gruyter. 183-201.
Otoguro, R. (2009). Surface-oriented morphological network. *Humanitas* 48. Waseda University Law Association. 219-240.
Otsu, Y. (1980). Some aspects of rendaku in Japanese and related problems. *MIT working papers in linguistics: Theoretical issues in Japanese linguistics*, Vol. 2. 208-227.
Peperkamp, S. (1997). *Prosodic words*. HIL Dissertations 34. The Hague: Holland Academic Graphics.
Pierrehumbert, J. and M. Beckman (1988). *Japanese tone structure*. Cambridge MA/London: MIT Press.
Poppe, C. (2012). Accent and domain structure in Tokyo Japanese: generalizing across lexical classes. In: The Phonological Society of Japan (ed.), *Phonological Studies* 15. Tokyo: Kaitakusha. 43-50.
Poppe, C. (2014). The status of feet in Japanese accentuation. In: The Phonological Society of Japan (ed.), *Phonological Studies* 17. Tokyo: Kaitaikusha. 67-74.
Poser, W.J. (1984). The phonetics and phonology of tone and intonation in Japanese. MIT doctoral dissertation.
Poser, W.J. (1990). Evidence for foot structure in Japanese. *Language* 66:1. 78-105.
Prince, A.S. (1976). 'Applying' stress. Unpublished manuscript, University of Massachusetts, Amherst.
Prince, A.S. (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry* 11:3. 511-562.
Prince, A.S. (1983). Relating to the grid. *Linguistic Inquiry* 14:1. 19-100.
Prince, A.S. and P. Smolensky (1993). *Optimality Theory: Constraint interaction in generative grammar.* Rutgers University Center for Cognitive Science Technical Report

2. Published in 2004 by Blackwell Publishers, Oxford.
Pulleyblank, D.G. (1983). Tone in lexical phonology. MIT doctoral dissertation. Published by D. Reidel Publishing Company, Dordrecht, 1986.
Pulleyblank, D. (1996). Neutral vowels in Optimality Theory: A comparison of Yoruba and Wolof. *The Canadian Journal of Linguistics* 41:4. 295-347.
Purnell, T.C. (1998). Principles and parameters of phonological rules: Evidence from tone languages. University of Delaware doctoral dissertation,
Raffelsiefen, R. (2005). Paradigm uniformity effects versus boundary effects. In: L.J. Downing, T.A.Hall and R. Raffelsiefen (eds.), *Paradigms in phonological theory*. Oxford/New York: Oxford University Press. 211-262.
Revithiadou, A. (1999). *Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems*. LOT Dissertation Series 15 (HIL/Leiden Universiteit). Holland Academic Graphics, The Hague.
Riad, T. (1998). Towards a Scandinavian accent typology. In: W. Kehrein & R. Wiese (eds.), *Phonology and morphology of Germanic languages*. Linguistische Arbeiten 386. 77-109.
Rodier, O.D. (1998). Prosodic domains in Optimality Theory. McGill University dissertation.
Rosen, E.R. (2001). Phonological processes interacting with the lexicon: Variable and non-regular effects in Japanese phonology. University of British Columbia doctoral dissertation.
Rosen, E.R. (2008). Deriving the accentual behavior of Japanese suffixes without special devices. Ms., University of British Columbia.
Saito, K. (1992). Boin museika no ‘hirosa’ to ‘tsuyosa’: Iwate hōgen wo chūshin ni shite [The extent and strength of vowel devoicing: with a focus on the Iwate dialects]. *Kokugogaku Kenkyū* 31. 53-64.
Sakano, N. (1996). *Shichi-go-chō no nazo o toku* [Solving the riddles of seven-five rhythm]. Toyo: Taishūkan.
Samuels, B.D. (2011). *Phonological architecture: A biolinguistic perspective.* Oxford/New York: Oxford University Press.
Sanada, S. (1998). *Toyama-ken no kotoba* [The language of Toyama Prefecture]. Tokyo: Meiji Shoin.
Sato, H. (2002). Gairaigo ni okeru onsetsu fukugō e no kubun to akusento [Decomposition into syllable complexes and the accenting of Japanese borrowed words]. *Journal of the Phonetic Society of Japan* 6:1. 67-78.
Sato, Y. (2008). Defining the prosodic word in Japanese. *MGU Journal of Liberal Arts*

Studies 2:2. 13-20.

Schadeberg, T.C. (1973). Kinga: A restricted tone system. *Studies in African Linguistics* 4:1.

Scheer, T. (2004). *A lateral theory of phonology. Vol 1: What is CVCV, and why should it be?* Berlin: Mouton de Gruyter.

Scheer, T. (2008). Why the prosodic hierarchy is a diacritic and why the interface must be direct. In: J.M. Hartmann, V. Hegedüs and H. van Riemsdijk (eds.), *Sounds of silence: empty elements in syntax and phonology*. Amsterdam: Elsevier. 145-192.

Scheer, T. (2010). What OT is, and what it is not. Review of de Lacy (ed.) (2007), The Cambridge handbook of phonology. *Journal of Linguistics* 46. 193-218.

Scheer, T. (2011). *A guide to morphosyntax-phonology interface theories: How extra-phonological information is treated in phonology since Trubetzkoy's grenzsignale*. Berlin/New York: Walter de Gruyter.

Schiering, R., B. Bickel and K.A. Hildrebrandt (2010). The prosodic word is not universal, but emergent. *Journal of Linguistic*s 46. 657-709.

Selkirk, E.O. (1978). On prosodic structure and its relation to syntactic structure. Paper presented at the Conference on Mental Representation in Phonology. Published in *Nordic Prosody II*, ed. by T. Fredheim (1981). Trondheim: TAPIR. 111-140.

Selkirk, E. (1996). The prosodic structure of function words. In: J.L. Morgan and K. Demuth (eds.), *Signal to syntax*. Mahwah, NJ: Lawrence Erlbaum. 187-213.

Selkirk, E. & K. Tateishi (1988). Constraints on minor phrase formation in Japanese. In: *Papers from the 24th Annual Regional Meeting of the Chicago Linguistic Society*. Chicago: CLS. 316-336.

Selkirk, E. (2009). On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyū* 136. 35-73.

Selkirk, E. (2011). The syntax-phonology interface. In: J. Goldsmith, J. Riggle and A. Yu (eds.), *The handbook of phonological theory, 2nd edition*. Oxford: Blackwell Publishing.

Sells, P. (1995). Korean and Japanese morphology from a lexical perspective. *Linguistic Inquiry* 26:2. 277-325.

Shibata, T. (1994). Gairaigo ni okeru akusentokaku no ichi [The location of accent kernels in loanwords]. In: K. Satō (ed.), *Kokugo ronkyū 4 gendaigo –hōgen no kenkyū*. Tokyo: Meiji Shoin. 418-388.

Shibatani, M. (1979). Review of S. Haraguchi (1977), The tone pattern of Japanese: An autosegmental theory of tonology. *Language* 55. 928-936.

Shibatani, M. (1990). *The languages of Japan*. Cambridge: Cambridge University Press.

Shimizu, M. (2006). Ehime-ken Seiyo-shi Tanohama hōgen no akusento to go'on no

kankei ni tsuite. *Journal of the Phonetic Society of Japan* 10:1. 30-38.
Shimoji, M. (2009). Foot and rhythmic structure in Irabu Ryukyuan. *Gengo Kenkyū* 135. 85-122.
Shinohara, S. (2000). Default accentuation and foot structure in Japanese: Evidence from Japanese adaptations of French words. *Journal of East Asian Linguistics* 9:1. 55-96.
Sibata, T. (1962). On'in [Phonology]. In: Kokugo Gakkai (ed.), *Hōgengaku Gaisetsu*. Tokyo: Musashino Shoin.
Smolensky, P. (1993). Harmony, markedness, and phonological activity. Handout of keynote address, Rutgers Optimality Workshop 1. [ROA-87]
Smolensky, P. (2006) Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature domain markedness. In P. Smolensky and G. Legendre (eds.), *The harmonic mind: From neural computation to optimality theoretic grammar, vol. 2: Linguistic and philosophical implications* 27-160. Cambridge, MA: MIT Press.
Steriade, D. (2000). Paradigm uniformity and the phonetics-phonology boundary. In: M. Broe and J. Pierrehumbert (eds.), *Papers in laboratory phonology V: Acquisition and the lexicon*. Cambridge: Cambridge University Press. 313-334.
Sugahara, M. (2003). Downtrends and post-focus intonation in Tokyo Japanese. University of Massachusetts, Amherst dissertation.
Sugito, M. (1969). Dōtai sokutei ni yoru nihongo akusento no kaimei [An investigation of Japanese accent based on the measurement of movements]. *Gengo Kenkyū* 55. 14-39. Reprinted in M. Sugito (1982). 49-75.
Sugito, M. (1981). Akusento ni yoru koe no sage to bo'in no onchō to no kanren [The relation between accentual voice lowering and vowel articulation]. *Gengo Kenkyū* 79. 1-32. Reprinted in M. Sugito (1982). 248-269.
Sugito, M. (1982). *Nihongo akusento no kenkyū* [Studies on Japanese accent]. Tokyo: Sanseido.
Sugito, M. (1995). *Osaka Tokyo akusento jiten* [Osaka Tokyo accent dictionary] (CD-Rom). Tokyo: Maruzen.
Sugito, M. (2003). Timing relationships between prosodic and segmental control in Osaka Japanese word accent. *Phonetica* 60. 1-16.
Sugito, M. (2005). The effect of speech rate on devoiced accented vowels in Osaka Japanese. In: J.M. van de Weijer and T. Nishihara (eds.), *Voicing in Japanese*. Berlin: Mouton de Gruyter. 247-260.
Suzuki, H. (1995). Minimal words in Japanese. *CLS* 31. 448-463.
Tajima, K. (1998). Speech rhythm in English and Japanese: Experiments in speech

cycling. Indiana University doctoral dissertation.

Tajima, K. and R.F. Port (2004). Speech rhythm in English and Japanese. In: J. Local, R. Ogden and R. Temple (eds.), *Papers in laboratory phonology VI*. Cambridge: Cambridge University Press. 322-339.

Tamori, I. (1980). Coocurrence restrictions on onomatopoeic adverbs and particles. *Papers in Japanese Linguistics* 7. 151-172.

Tanaka, S. (伸一) (1992). Accentuation and prosodic constituenthood in Japanese. *Proceedings of the 5th Annual Meeting of the Tokyo Linguistic Forum*. 195-216.

Tanaka, S. (伸一) (2000). The avoidance of rising pitch within a syllable: evidence for the superheavy syllable and the independency hierarchy in Japanese. In: M. Amano et al. (eds.), *Synchronic and diachronic studies on language: A festschrift for Dr. Hirozo Nakano*. Nagoya: Nagoya Daigaku Daigakuin Bunka Kenkyūka Eigogaku Kenkyūshitsu. 521-531.

Tanaka, S. (伸一) (2001). The emergence of the 'unaccented': possible patterns and variations in Japanese compound accentuation. In: J. van de Weijer and T. Nishihara (eds.), *Issues in Japanese phonology and morphology*. Berlin/New York: Mouton de Gruyter. 159-192.

Tanaka, S. (伸一) (2003). Japanese grammar in the general theory of prominence: its conceptual basis, diachronic change, and acquisition. In: Honma, Takeru, Masao Okazaki, Toshiyuki Tabata, Shin-ichi Tanaka (eds.), *A new century of phonological theory: A festschrift for Professor Shosuke Haraguchi on the occasion of his 60th birthday*. Tokyo: Kaitakusha. 102-120.

Tanaka, S. (伸一) (2012). Mō hitotsu no Nihongo no futto: Gairaigo akusento kisoku no kigen to '-3' no wake [Another Japanese foot: the origin of the loanword accent rule and the reason behind '-3'. Jiritsu chōwateki shiten kara mita on'in ruikei no moderu. Japanese Society for the Promotion of Science Research Report. 97-110.

Tanaka, S. (伸一) (2013). Review of Labrune, Laurence (2012) The Phonology of Japanese, Oxford University Press, 320p. *Journal of the Phonetic Society of Japan* 17:1. 70-80.

Tanaka, S. (真一) (2008). *Rizumu, akusento no 'yure' to on'in, keitai kōzō* [Rhythmic and accentual variation and phonological and morphological structure]. Tokyo: Kurosio.

Tang, K.E. (2008). The phonology and phonetics of consonant-tone Interaction. Universtiy of California, Los Angeles doctoral dissertation.

Tashiro, Kōji (1953). *Hyōjungo no akusento kyōhon* [Accent textbook for Standard Japanese]. Osaka: Sogensha.

Tateishi, K. (1989). Theoretical implications of the Japanese musicians' language. In: E.

J. Fee and K. Hunt (eds.), *Proceedings of the Eighth West Coast Conference on Formal Linguistics.* Stanford: Stanford Linguistic Association.

Terada, Y. (1970). *Enshū hōgen no akusento* [Accent in the Enshu dialects]. Iwata: Bisaidō Shoten.

Teranishi, R. (1980). Two-mora-cluster as a rhythm unit in spoken Japanese sentence or verse. Abstract of a talk published in *Journal of the Acoustical Society of America* 67, Supplement 1:S40.

Terao, Y. (1992). Bunsanshutsu kenkyū kara ita onsetsu-mōra. *Nihongo no mōra to onsetsu kōzō ni kansuru sōgōteki kenkyū* (1). Research Report, University of Tsukuba. 62-75.

Trommer, J. (2011). Phonological Aspects of Western Nilotic Mutation Morphology. University of Leipizig habilitation thesis.

Trommer, J. (2012). Introduction. In: J. Trommer (ed.), *The morphology and phonology of exponence*. Oxford: Oxford University Press. 1-7.

Trubetzkoy, N.S. (1969[1939]). *Principles of phonology*. Translated by C.A.M. Baltaxe. Berkeley/Los Angeles: University of California Press.

Tsujimura, N. & S. Davis (2011). A construction approach to innovative verbs in Japanese. *Cognitive Linguistics* 22. 799-825.

Ussishkin, A. (2000). The emergence of fixed prosody. University of California, Santa Cruz doctoral dissertation.

Uwano, Z. (1977). Nihongo no akusento [Japanese accent]. In: *Iwanami Kōza Nihongo 5: On'in*. Tokyo: Iwanami shoten. 281-321.

Uwano, Z. (1981). On the accentual system of the Matsue dialect. *Nihon Kaiiki Kenkyūsho Hōkoku* 13 . 109-136.

Uwano, Z. (1984a). Chihō akusento no kenkyū no tame ni [For the sake of the study of dialect accent]. *Kokubungaku Kaisyaku to Kanshō* 49:7. 47-64.

Uwano, Z. (1984b). Niigata-ken Murakami-shi hōgen no akusento. Sugito, Miyoko (ed.), *Kindaichi Haruhiko hakushi koki kinen ronbunshū 2: Gengogaku-hen*. Tōkyō: Sanseidō. 347-390.

Uwano, Z. (1996). Kanazawa hōgen no kōbu 3haku fukugōgo meishi no akusento kisoku [Accent Rules of Compound Nouns Ending in 3-mora Elements in the Kanazawa Dialect]. *Gengogakurin 1995-1996.* 16. 337-355. Tokyo: Sanseido.

Uwano, Zendo (1997a). Kanazawa hōgen no kōbu 2haku fukugōgo meishi no akusento kisoku [Accent rules of compound nouns ending in 2-mora elements in the Kanazawa dialect]. *Tōkyō Daigaku Gengogaku Ronshū* 16. 3-127.

Uwano, Z. (1997b). Fukugō-meishi kara mita nihongo sho-hōgen no akusento [The

accent of Japanese dialects as seen from nominal compounds]. In: M. Sugito (ed.), *Nihongo onsei 2: Akusento, intonēshon, rizumu to pōzu*. Tokyo: Sanseido. 231-270.

Uwano, Z. (1999a). Fukugō-meishi kōbu-yōso akusento-gata hozon. [The preservation of accent in the second constituent of compounds]. *Gengo to bunka no shosō*. Morioka: Iwate University. 195-212.

Uwano, Z. (1999b). Classification of Japanese accent systems. In: S. Kaji, S. (ed.) *Proceedings of the Symposium Cross-linguistic Studies of Tonal Phenomena: Tonogenesis, Ttypology, and Related Topics*. Tokyo: ILCAA. 151-178.

Uwano, Z. (2003). Akusento no taikei to shikumi [Accent systems and mechanisms]. In: Uwano, Zendo (eds.), *Asakura Nihongo kōza 3: Onsei, on'in.* Tokyo: Asakura Shoten.

Uwano, Z. (2009). Tsūjiteki ni shika setsumei dekinai kyōji akusento genshō: kutō no jōshō to go'on to no kankei [A synchronic accentual phenomena that can only be explained diachronically: the relation between the phrase-initial rise and word sounds]. *Gekkan Gengo* 38:2. 74-81.

Uwano, Z. (2012a). 'Weak' units and accent shift. *Proceedings Fonetik 2012*. The XXVth Swedish Phonetics Conference (University of Gothenburg, May 30-June 1 2012). 9-13.

Uwano, Z. (2012b). Three types of accent kernels in Japanese. *Lingua* 122:13. 1415-1440.

Uwano, Z. & T. Nitta (1982). Kanazawa hōgen no meishi no akusento: akusento taikei to shozoku go'i [The accent of nouns in the Kanazawa dialect: the accent system and the classification of nouns]. *Kokugo Kenkyū* 45. 1-31.

Uwano, Z. & T. Nitta (1983). Kanazawa hōgen no akusento to go'on no kankei ni tsuite [About the relation between accent and segmental sounds in the Kanazawa dialect]. *Nihonkai Bunka* 10. 1-43.

Vance, T.J. (1987). *An introduction to Japanese phonology*. Albany, N.Y.: State University of New York Press.

Vance, T.J. (1993). Are Japanese particles clitics? *Journal of the Association of Teachers of Japanese* 27:1. 3-33.

Vance, T.J. (2008). *The Sounds of Japanese*. Cambridge: Cambridge University Press.

Vance, T.J. (2013). Review of 'The Phonology of Japanese' (L. Labrune, Oxford University Press). *Lingua* 123. 168-174.

Vance, T.J. (2015). Rendaku. In: H. Kubozono (ed.), *The Handbook of Japanese Phonetics and Phonology*. Boston/Berlin: Walter de Gruyter/Mouton.

Vaux, B. (2008). Why the phonological component must be serial and rule-based. In B. Vaux and A. Nevins (eds.), *Rules, constraints, and phonological phenomena*. Oxford: Oxford University Press. 20-61.

Voorhoeve, J. (1973). Safwa as a restricted tone system. *Studies in African Linguistics*

4:1.
Wada, M. (1958). Fukuzatsu na akusento taikei no kaishaku [The interpretation of a complicated accent system]. *Kokugogaku* 32. 72-90.
Warner, N. and T. Arai (2001). Japanese mora timing: A review. *Phonetica* 58:1. 1-25.
Warner, N., T. Otake and T. Arai (2010). Intonational structure as a word-boundary cue in Tokyo Japanese. *Language and Speech* 53:1 107-131.
Weijer, J. van de (1999). Analogical change in Optimality Theory. In: The Phonological Society of Japan (ed.), *Phonological Studies* 2. Tokyo: Kaitakusha. 145-152.
Weijer, J. van de and M. Sloos (2013). Acquiring markedness constraints: The case of French. In S. Aalberse and A. Auer (eds.), Linguistics in the Netherlands 2013. Amsterdam: John Benjamins. 188-200.
Yamada, E. (1990) Stress Assignment in Tokyo Japanese (1) and (2). *Fukuoka Daigaku Jinbun Ronsō* 21 (1575-604), 22 (97-154).
Yamada, E. (1992). Parameter setting for abstract stress in Tokyo Japanese. *Proceedings of the 1992 International Conference on Spoken Language Processing*. Alberta: University of Alberta. 1267-1270.
Yamaguchi, K. (2010a). The difference in accentuation between the present tense and the past tense of verbs: A comparison between Tokyo Japanese and Shizuoka Japanese. Handout of a talk delivered at the 5th Phonology Festa, Biwako Ogoto Onsen, Ōtsu City, Shiga-Ken, February 18-19, 2010.
Yamaguchi, K. (2010b). The differences in accentuation between the present and the past tenses of verbs in Japanese. *Journal of the Phonetic Society of Japan* 14-3. 1-10.
Yamaguchi, Y. (1984). Akusento taikei ga shashō shita mono: Shizuoka-ken Maisaka-chō hoogen wo rei to shite [Abstractions in an accent system: the case of the Maisaka dialect]. *Gendai hōgengaku no kadai 2: Kijutsu kenkyū hen*. Tokyo: Meiji Shoin. 309-332.
Yip, M. (1995). Isolated uses of prosodic categories. In: J.S. Cole and C.W. Kissberth (eds.), *Perspectives in phonology*. Chicago: University of Chicago Press. 293-311.
Yip, M. (2001). The complex interaction of tone and prominence. In: M. Kim & U. Strauss (eds.), *Proceedings of NELS* 31. Amherst, MA: GLSA. 531-545.
Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
Yokotani, T. (1997). Accent shift beyond the foot boundary: evidence from Tokyo Japanese compound nouns [in Japanese; English title as given]. *Journal of the Phonetic Society of Japan* 1:1. 54-62.
Yoshida, K. (2011). The phonetics and phonology of shiki-accent systems in mid-western Japanese dialects. Indiana University doctoral dissertation.
Zamma, H. (1992). The accentuation system of Japanese inflection. *Tsukuba English*

Studies 11. 117-148.
Zoll, C. (2003). Optimal tone mapping. *Linguistic Inquiry* 34:2. 225-268.
Zubizarreta, M.L. (1982). The formal interaction of harmony and accent: the tone pattern of Japanese. In: H. van der Hulst and Norval Smith (eds.), *The structure of phonological representations (Part II)*. Dordrecht: Foris Publications. 159-212.