

CONTRASTIVE REPRESENTATIONS IN NON-SEGMENTAL PHONOLOGY

by

Christopher Edward Spahr

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy  
Graduate Department of Linguistics  
University of Toronto

© Copyright 2016 by Christopher Edward Spahr

# Abstract

Contrastive representations in non-segmental phonology

Christopher Edward Spahr

Doctor of Philosophy

Graduate Department of Linguistics

University of Toronto

2016

This thesis develops and tests a unified model of word-level prosodic contrasts. Traditionally, word prosody has been analyzed within disparate models (such as autosegmental theory for tone, metrical theory for stress, and CV, X-slot, or moraic theory for length), meaning that it has not been possible to make clear predictions about how many different prosodic features can be employed in a single language.

I present a minimal architecture for word prosodic representations based on a single set of formal elements. A tier of segmental root nodes, or X-slots, bears the binary contrastive features that divide the segmental inventory and represents quantity contrasts through two-to-one linking, while a tier of prosodic root nodes, or “ $\pi$ -nodes”, bears the binary features dividing the autosegmental inventory. Features on  $\pi$ -nodes are used in tone languages with more than one tonal autosegment, but in privative tone languages, the  $\pi$ -node itself reflects the phonetic realization of the marked member of the tone opposition. The same featureless  $\pi$ -node is used as an autosegmental marker of accented positions in lexical stress systems, where its language-specific phonetic realization is that of stress: some combination of increased pitch, duration, and intensity.

The predictive power of this model is that it restricts systems to a maximum of two independent word prosodic contrasts, since each requires its own tier of root nodes. The  $\pi$ -tier can represent either tone or accent separately from length on the X-tier, but this leaves no means to represent a third contrast. In certain systems, surface stress may be represented covertly as length on the X-tier with tone represented on the  $\pi$ -tier, but no mechanism is available to host a third contrast, since the X-tier is already used for stress.

## Acknowledgements

I started in the Linguistics PhD program at the University of Toronto after experiencing during my MA an overwhelmingly positive and supportive environment. Whenever someone would ask me “will you be here for a minute?” so that I could watch their things while they briefly left the room, I would cheerfully quip “I’m gonna be here for years!”. It’s hard to believe that those years (~4.5 of them) have gone by, and that my time in grad school is over—the feeling is about as bittersweet as I can imagine. For a long time I’ve dreaded writing these acknowledgements, knowing that there would be people I that I would forget to thank. Whoever you are, please accept my apologies for this.

First, I owe a great debt of gratitude to Elan Dresher, my supervisor, and to the other members of my committee: Daniel Currie Hall, Peter Jurgec, Yoonjung Kang, Keren Rice, and Draga Zec. Elan has provided a tremendous amount of feedback, advice, and appropriately-placed skepticism over the years, without which I wouldn’t be the linguist I am today. Keren and Peter also oversaw this thesis from start to finish, giving numerous helpful comments on drafts along the way, and I thank them for that. Furthermore, the case studies in Chapter 4 also benefited from correspondence with Gabriela Caballero and José Ignacio Hualde. Finally, I’d like to thank Elizabeth Cowper for helping me learn how to be an inquisitive scholar early on in my grad school career, and for capping it off by hosting an awesome defense party.

I refer the reader to the acknowledgements section of Spahr 2011 for those I appreciate most from the very beginning of my post-graduate linguistics life, but I’ll give special shoutout to the people who have remained most in my life since MA year: Marisa Brook, Radu Craioveanu, and Ross Godfrey, who continued on to the PhD program, and James Byrnes and Michelle Stella, with whom I remain good local friends. Over the years I have enjoyed costudenthood with many others that I’d like to acknowledge for being part of my time here; an alphabetized non-exhaustive list follows: LeAnn Brown, Joanna Chociej, Emily Clare, Sarah Clarke, Richard Compton, Ailís Cournane, Derek Denis, Julie Doner, Clarissa Forbes, Yu-Leng Lin, Dan Milway, Alex Motut, Kenji Oda, Will Oxford, Michelle St-Amour, Julia Su, Eugenia Suh, and Tomohiro Yokoyama. Thanks to everyone for countless parties, pub nights, and lounge conversations, and a countable but delightfully large number of conference trips. It’s been an amazing ride.

Finally, I would like to thank those closest to me: my parents for a lifetime of supportiveness and Mahak Jain for years of patience. And everyone, of course, for all their love and kindness.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Word prosodic features . . . . .	2
1.2	Contrast in phonological representations . . . . .	4
1.3	The proposal in brief . . . . .	9
1.4	Organization of the thesis . . . . .	10
<b>2</b>	<b>Background</b>	<b>11</b>
2.1	Tone . . . . .	11
2.1.1	Autosegmental phonology . . . . .	12
2.1.2	Tonal feature geometries . . . . .	15
2.2	Stress . . . . .	22
2.2.1	Stress and accent . . . . .	23
2.2.2	The metrical enterprise . . . . .	26
2.2.3	Accent as an autosegment . . . . .	33
2.3	Length . . . . .	40
2.3.1	Representing length . . . . .	41
2.4	Summary . . . . .	48
<b>3</b>	<b>Towards a minimal model of prosody</b>	<b>49</b>
3.1	The minimal elements of prosodic contrasts . . . . .	49
3.1.1	Contrast and the tonal node . . . . .	49
3.1.2	A unified prosodic autosegment . . . . .	56
3.1.3	Length in a minimal model . . . . .	61
3.2	Restrictions and predictions . . . . .	64
3.2.1	Feature structure is feature ordering . . . . .	64
3.2.2	Restrictions on several-to-one linking . . . . .	69
3.2.3	On the “maximum of two” restriction . . . . .	72
3.2.4	Stress as a feature . . . . .	74

3.2.5	Stress as doubled $\pi$ -nodes . . . . .	75
3.3	Summary . . . . .	78
<b>4</b>	<b>Tonal accent and the <math>\pi</math>-node</b>	<b>80</b>
4.1	Introduction . . . . .	80
4.2	Goizueta Basque . . . . .	81
4.2.1	The accent system . . . . .	81
4.2.2	Alternations . . . . .	85
4.2.3	Morphemic analysis . . . . .	88
4.2.4	Marginal phenomena . . . . .	92
4.2.5	Summary . . . . .	96
4.3	Chogueta Rarámuri . . . . .	96
4.3.1	Stress patterns . . . . .	97
4.3.2	Tone patterns . . . . .	100
4.3.3	A representational account . . . . .	105
4.3.4	Morphemic analysis . . . . .	108
4.3.5	In-place tone alternations . . . . .	116
4.3.6	Exceptional stress alternations . . . . .	122
4.3.7	Summary . . . . .	125
4.4	Commonalities of tonal accent languages . . . . .	126
<b>5</b>	<b>Mixed systems on two tiers</b>	<b>128</b>
5.1	Introduction . . . . .	128
5.2	Papiamentu . . . . .	129
5.2.1	Contrastive tone and stress . . . . .	130
5.2.2	Theoretical accounts . . . . .	132
5.2.3	Papiamentu prosody in a unified model . . . . .	136
5.2.4	Morphemic analysis . . . . .	145
5.2.5	Summary . . . . .	150
5.3	Estonian . . . . .	151
5.3.1	The quantity system . . . . .	151
5.3.2	Representing three-way quantity . . . . .	155
5.3.3	Three-way quantity in a minimal model . . . . .	162
5.3.4	Prosodic alternations and empty root nodes . . . . .	167
5.3.5	Prosodic degemination . . . . .	174
5.3.6	Prosodic epenthesis . . . . .	177
5.3.7	Case allomorphy . . . . .	180

5.3.8	Optional allomorphy . . . . .	186
5.3.9	Q2–Q3–Q2 alternations . . . . .	188
5.3.10	The short illative . . . . .	193
5.3.11	Summary . . . . .	196
5.4	The diversity of privative $\pi$ . . . . .	197
<b>6</b>	<b>Conclusions</b>	<b>199</b>
6.1	Towards a minimalist typology of prosodic systems . . . . .	201
6.2	Future work . . . . .	205
	<b>References</b>	<b>207</b>

# Chapter 1

## Introduction

Yet a certain degree of vagueness seems to characterize most discussions of prosodic features. They seem more elusive than segmental features, and their incorporation into a linguistic system sometimes seems to strain the limits of an otherwise coherent framework.

---

Lehiste (1970: 1)

Prosodic features—principally tone, stress, and length—have long been known to serve a contrastive role in natural language phonology, and because they behave differently from segmental features, they have been treated separately in theoretical work. However, it is also understood that prosodic features themselves show distinct behaviours from each other. Consequently, whereas it has been possible to arrive at relatively coherent feature-based models of segmental contrasts (with disagreement over the details; see for example discussion in Mielke 2008; Blaho 2008; Dresher 2009), disjoint models have been proposed to account for different prosodic phenomena. Despite the fact that connections and similarities exist between them, no unified model exists for word-level prosodic contrasts, and as a result, it is not possible to make clear predictions about how many prosodic features can be independently contrastive within a single system. The purpose of this thesis is to develop and test a model that fills this gap.

Prosodic representations are a particularly ripe area for theoretical unification because of the

similarities among existing approaches. Most models of prosodic contrasts incorporate some notion of interactions or dominance between different representational tiers, implemented variously as tonal autosegments, durational autosegments, or structural relationships. Drawing on insights from existing work on the structure of word prosody and the organization of phonological contrasts, I will put forth a model in which lexical representations consist of two tiers of root nodes—one bearing segmental contrastive features and one bearing prosodic contrastive features. Segmental root nodes double as the timing units used in quantity systems, while prosodic root nodes serve as both tonal autosegments and as autosegmental markers of accented positions in lexical stress languages.

This thesis lays out a coherent architecture for the representation of prosodic contrast under a single set of assumptions, rather than taking a separate theoretical approach to each prosodic feature. Crucially, the model I propose is restrictive in the kinds of contrastive systems it predicts: it limits the number of independent prosodic contrasts in a single language to a maximum of two. I first show how this model allows for integrated and straightforward representations of the contrasts in simple prosodic systems (those with only one of tone, length, and stress). I then demonstrate that, with the right level of abstraction with respect to the surface phonetics realized in various formal configurations, it allows for elegant and morphologically compositional accounts of the apparently non-concatenative morphophonological alternations observed in a number of mixed prosodic systems (those with more than one contrastive prosodic feature). This account of word prosody puts the majority of the explanatory burden in morphophonology on phonological representations and the phonological rules that manipulate them, and moves that burden away from processes that are triggered by morphological categories or features.

## 1.1 Word prosodic features

Prosodic features, also referred to as “suprasegmentals” (by, for example, Lehiste 1970; Leben 1973), are traditionally understood to be distinctive properties larger than segments. They are



understood to exist above or on a different level than segments, and can span across several segments. Fox (2000: 5–6) notes that they must be defined in terms of relationships—they are syntagmatic rather than paradigmatic. They are sometimes thought of as secondary to segments, but yet they are in a sense primary: “they can be seen as properties of the *source* rather than of the supralaryngeal filter” (Fox 2000: 4, emphasis in original).

Their independence from segmental features, and thus the motivation for giving them different formal properties, is evidenced by particular kinds of behaviour not seen as readily in segmental phonology. For example, a key notion in the development of autosegmental phonology (Goldsmith 1976) is tonal “stability”. In many languages, the deletion of a segment often does not entail the deletion of a tone associated with that segment. Likewise, morphemes may consist solely of “floating” tones, having no segmental content of their own. Similar behaviour is seen with length. In compensatory lengthening, for example, the deletion of one segment does not entail the deletion of its duration, and results in an adjacent segment becoming phonologically long. Furthermore, morphemes may be said to consist of only floating length and no segmental component. These same properties are seen once again in stress accent languages, where the deletion of a stressed position will result in stress being realized elsewhere, and certain morphemes may only be expressed as a change in the position of stress.

A note is in order regarding the scope of the phenomena under investigation. The term “prosody” can refer not only to word-level or lexical contrasts, but to phrase-level and utterance-level manifestations of the same phonetic correlates. For example, languages may use intonation or stress to convey pragmatic information or as a reflection of syntactic structure, but not possess lexical tone or stress accent. The present study is concerned with prosody that is contrastive at the word level: those properties that are distinctive in the lexical phonology but are not features of individual vowels and consonants. The particulars of these features and their formal treatment will be addressed in turn.

## 1.2 Contrast in phonological representations

This thesis takes a contrast-centric approach to the study of phonology, sitting within a body of research that assumes that the organization of phonological representations is informed, in a principled way, by the number of contrasts in a system. Although the details of the role played by contrast vary widely (for a survey, see Dresher et al. 1994; Dresher 2009), the common thread is that phonological elements such as features are only assigned to distinct members of the phonological inventory. Since the number of contrastive features needed to distinguish the members of an inventory is smaller than the number of descriptors needed to fully specify the phonetic realizations of the members of that inventory, it follows that phonological representations are to some extent underspecified, and must be phonetically enhanced by some separate module of the grammar. I state this contrastivist conceptualization of representations as in (1):

- (1) *Contrastivism* (Dresher et al. 1994; Dresher 2009)

Phonological representations contain only the formal information used to distinguish contrastive categories

The main reason for this approach is methodological. By taking as an analytical starting point the minimum amount of formal information—that employed to represent the contrasts in a language—it is easy to falsify a representational model: a proposal that cannot reflect the contrasts attested in a given language without introducing additional machinery is not viable (that is, models that undergenerate are ruled out). The significance of the elements of contrastive representations (such as features) to phonological computation is an important consideration in contrastivist research. For example, one position taken in contrastivist work is what Hall (2007) calls the “Contrastivist Hypothesis”:

- (2) *The Contrastivist Hypothesis* (Hall 2007, paraphrased)

The phonology refers only to contrastive features

In other words, phonological processes may not make reference to information which is not formally specified. I adopt (2) as a working hypothesis in this study, extending it to all contrastive

elements, not only features. The other piece of the puzzle in contrastivist research concerns the way that contrastive elements are assigned. Among models proposed for segmental contrasts is Modified Contrastive Specification, also known as the Contrastive Hierarchy theory (Dresher et al. 1994; Dresher 2009):

(3) *The Contrastive Hierarchy* (Dresher 2009)

Feature specifications of contrastive phonological categories are assigned according to language-specific contrastive hierarchies

Contrastive hierarchies are built by successively dividing an inventory of contrastive categories (that is, phonemes) through a series of binary cuts, each of which refers to a single phonological feature. Crucially, two languages with identical phonetic inventories may show different phonological inventories, but the number of possible unique feature specifications is constrained based on inventory size. Consider for example a three-member inventory of bilabial consonants:<sup>1</sup>

(4) /p, b, m/

Under a traditional phonological analysis along the lines of Chomsky and Halle (1968), this inventory will be fully specified for the relevant features [ $\pm$ voice] and [ $\pm$ nasal], as in (5):

	/p/	/b/	/m/
(5) [voice]	-	+	+
[nasal]	-	-	+

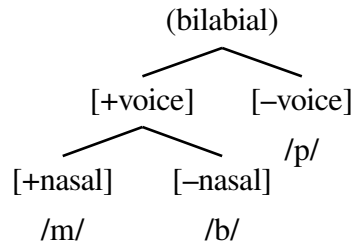
The contrastive hierarchical approach determines specifications based on language-specific feature ordering, which can be shown with branching feature trees. For the sub-inventory in (5), a language may begin with either the feature [ $\pm$ voice] or [ $\pm$ nasal]. Choosing [ $\pm$ voice] first yields the contrastive hierarchy in (6). The feature [ $\pm$ voice] first divides the inventory into its voiced and voiceless members, assigning the relevant contrastive feature. The next ordered feature then applies to divide any remaining phonemes that have not been exhaustively specified—terminal

---

<sup>1</sup>The points made here follow exemplification used by Dresher (passim).

nodes of the tree that do not end in a single segment. The two phonemes /m/ and /b/ have been specified as [+voice], but they are not distinguished from each other. This is accomplished by going on to divide the inventory based on [ $\pm$ nasal], the next feature in the ordering. In addition to their contrastive specification for [+voice], /b/ and /m/ here receive contrastive specifications for [-nasal] and [+nasal] respectively. Unlike /m/ and /b/, /p/ occupies a terminal node after it has been specified for [-voice], and thus it does not receive any further features; it is therefore contrastively underspecified for the feature [ $\pm$ nasal]:

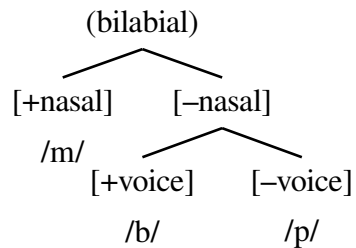
(6) [ $\pm$ voice] > [ $\pm$ nasal]



The feature specifications in the hierarchy in (6) correspond to the table in (7):

	/p/	/b/	/m/
(7) [voice]	-	+	+
[nasal]		-	+

Another language with the same surface three-member inventory of bilabial consonants could order the feature [ $\pm$ nasal] first, yielding the hierarchy in (8). In this case, the inventory will first be divided into an opposition of [+nasal] members and [-nasal] members. This leaves /m/ exhaustively specified with only the feature [+nasal]. The other two members of the inventory, /b/ and /p/, are then distinguished from each other by the feature [ $\pm$ voice] under the [-nasal] branch of the tree.

(8)  $[\pm\text{nasal}] > [\pm\text{voice}]$ 

The feature specifications in (8) are equivalent to those in (9):

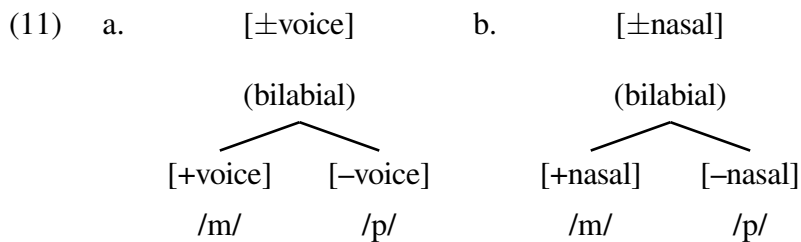
	/p/	/b/	/m/
(9) [voice]	-	+	
[nasal]	-	-	+

The key distinction between the tree in (6), with  $[\pm\text{voice}]$  ordered first, and the tree in (8), with  $[\pm\text{nasal}]$  ordered first, is that in the former, the class of  $[\text{+voice}]$  consonants includes the nasal /m/, whereas in the latter, /m/ receives only the feature  $[\text{+nasal}]$  and does not participate in the voicing opposition. That is to say, languages with different contrastive hierarchies will have different feature specifications, which reflect language-specific phonological patterning. The analyst can determine a language's feature ordering by using a corollary of the Contrastivist Hypothesis: if only contrastive features are phonologically active, then a feature that is phonologically active (evidenced, for example, by patterning together in phonological processes) must be contrastively specified.

Contrastive hierarchical specification also limits the number of contrastive features that can be active depending on inventory size. When dividing a three-member sub-inventory, as seen above, two features are necessary, with one member underspecified for one of the features. Consider, however, a two-member inventory of bilabial consonants containing just /p/ and /m/. Traditionally these would be fully specified for both  $[\pm\text{voice}]$  and  $[\pm\text{nasal}]$ :

	/p/	/m/
(10) [voice]	-	+
[nasal]	-	+

In a contrastive hierarchical model of specification, the full specifications in (10) are not possible. If the sub-inventory is distinguished first by the feature  $[\pm\text{voice}]$ , as in (11a), then /p/ will receive  $[-\text{voice}]$  and /m/ will receive  $[\text{+voice}]$ . However, since the pair is now exhaustively specified, it is not possible to further divide the inventory with another feature. If  $[\pm\text{nasal}]$  is used first, as in (11b), then the pair will be an opposition of only  $[\pm\text{nasal}]$  and have no specification for  $[\pm\text{voice}]$ .



Crucial to the contrastivist approach is that representations are not determined by phonetics, but by formal restrictions and phonological patterning of distinctive categories. The universal aspect of the phonological component of the grammar is thus not the features themselves, nor is it the phonetics that they specify; it is rather the ability to distinguish contrasts according to a particular set of principles:

(12) *Universals of phonological contrast* (Dresher 2014)

Phonological universals are manifested in the component of the grammar that organizes contrastive elements, not in the substantive content of those elements

Although the domain of phonology proper is the nature of formal representations, phonetic implementation must be handled at some point in the language faculty. Hall (2011) presents a view of the phonetic enhancement of hierarchically-assigned contrastive features, showing that it is possible to have a principled relationship between contrast and phonetics without the phonology referring to non-contrastive features. I view phonetic realizations similarly in this study, treating them as post-phonological exponents of formal categories and putting the predictive power of the model in the structure of contrastive representations.

### 1.3 The proposal in brief

While contrast has long played a central role in phonological theory, recent work has focused almost exclusively on segmental—paradigmatic—contrasts. In the present study, I use the contrastivist insights to study prosodic—syntagmatic—contrasts. I examine contrastive prosodic features without assuming that representations are phonetically deterministic. I focus on phonological activity and on developing a model of representations that restricts the distribution of prosodic contrasts.

I propose a model in which prosodic representations are made up of associations between two tiers of root nodes: segmental root nodes, or X-slots, which bear segmental contrastive features, and prosodic root nodes, or  $\pi$ -nodes, which bear prosodic contrastive features. Doubling of X-slots can be used to represent phonological length, and  $\pi$ -nodes can be used to represent contrastive tone or contrastive length. Because only two tiers of root nodes are available, this model predicts that no language may employ more than two independent prosodic contrasts. Crucially, there is no direct connection between the surface phonetic realization of a prosodic contrast and the formal representation for that contrast.

This model of word prosody can provide equivalent analyses to standard models for languages with only one prosodic feature (tone, stress, or length alone), and thus sheds little light on the “simple” cases in isolation. The innovation of the model comes from the restrictions that it places on the co-occurrence of prosodic features, and for this reason the appropriate testing ground is mixed prosodic systems, that is, languages with more than one prosodic contrast. I demonstrate the model with detailed case studies of four languages of this type: Goizueta Basque, Choguita Rarámuri, Papiamentu, and Estonian, providing for each a discussion of formal representations and a morphologically compositional analysis of prosodic alternations.

## 1.4 Organization of the thesis

This thesis is organized as follows. Chapter 2 provides background on existing approaches to the representation of prosodic features. I consider in turn tone, stress, and length, giving critical consideration to the role played by contrast in each.

Chapter 3 lays out the core assumptions of the proposal with reference to existing theoretical approaches to prosody. I discuss the relevant aspects of these theories and develop a unified model for the representation of word prosodic contrasts. I then address some formal issues arising from the model put forth and show how it restricts the number of independent prosodic features that can be present in a single language.

Chapter 4 presents case studies of the prosodic systems of Goizueta Basque and Choguita Rarámuri. In these languages, both stress and tone are contrastive and can alternate independently of one another, but the position of tone is entirely dependent on the position of stress. I show that the model argued for here straightforwardly predicts such systems and makes it possible to elegantly analyze the observed alternations.

Chapter 5 presents case studies of two typologically different languages: Papiamentu, which is said to have independent stress and tone contrasts, and Estonian, which has a surface three-way length contrast. I show that the present model provides the tools needed to formulate analyses of complex alternations without needing to resort to morphophonological rules or ad hoc representational mechanisms.

Chapter 6 concludes. I provide a summary of the thesis, explicitly lay out a property-based typology of word prosodic systems based on the model of representations developed in the preceding chapters, and discuss directions for future work.



# Chapter 2

## Background

This chapter reviews the principles of and motivations for the canonical interpretations of prosodic features. I address in turn word-level tone, stress, and length, describing the kinds of data considered as evidence for models of prosodic representation and taking a critical view of the disparity between the leading approaches to different phenomena.

### 2.1 Tone

Tone, broadly speaking, is the use of pitch in a linguistic system. Hyman (2006: 229) defines a tone language as follows:

- (1) A language with tone is one in which an indication of pitch enters into the lexical realization of at least some morphemes.

Tone as a phonological entity requires a more formal understanding than the definition in (1) would imply, however. Pitch is a property of most utterances, and pitch can be correlated with other linguistic contrasts, such as consonant voicing (Hombert et al. 1979; Kingston 2011). In formal phonology, “tone” refers to the specification and manipulation of pitch independently from segments. It is a contrastive property not of the segment, but of the tone bearing unit (TBU), which can vary in size from a segment to a mora to a syllable, depending on the language and the

analysis. Most importantly, a single tone can be shared between several TBUs, or a single TBU can bear several tones. For this reason, the study of tone as a phenomenon is concerned largely with the properties of tones themselves and the way they interact with TBUs.

### 2.1.1 Autosegmental phonology

Without a doubt, the most influential study on tone in the generative literature is Goldsmith's (1976) MIT doctoral dissertation *Autosegmental phonology*.<sup>1</sup> Prior to autosegmental phonology, suprasegmentals like tone were either considered to be a distinct kind of non-segmental property spanning several segments or, especially in the generative literature, simple features of the segments themselves. Goldsmith proposed that properties like tone are in fact segments in their own right, but that they exist on a separate tier from vowels and consonants; because of this independent segmentation, they were termed "autosegments". Evidence for autosegmental representations was primarily drawn from a number of phenomena which proved difficult to analyze under existing assumptions about features, based largely on data from Green and Igwe's (1963) description of Igbo (Niger–Congo).

If tone were represented with a pair of binary features [ $\pm$ Highpitch] and [ $\pm$ Lowpitch] inside of segmental feature matrices, then high and low toned vowels would be specified as in (2):

$$(2) \quad \begin{array}{l} \text{a.} \\ \left[ \begin{array}{l} + \text{syllabic} \\ - \text{high} \\ - \text{round} \\ \dots \\ + \text{Highpitch} \\ - \text{Lowpitch} \end{array} \right] \\ \text{/á/} \end{array} \quad \begin{array}{l} \text{b.} \\ \left[ \begin{array}{l} + \text{syllabic} \\ - \text{high} \\ - \text{round} \\ \dots \\ - \text{Highpitch} \\ + \text{Lowpitch} \end{array} \right] \\ \text{/à/} \end{array}$$

For contours that behave as though they are a concatenation of two level tones, long contour tones could then be thought of as sequences of two short vowels with different tone features. /áà/ is thus quite literally a sequence of /á/ and /à/:

<sup>1</sup>For some general overviews of work in autosegmental phonology, see e.g., Lieber (1987); Snider (1999); Goldsmith (1990).

$$(3) \quad \left[ \begin{array}{l} + \text{syllabic} \\ - \text{high} \\ - \text{round} \\ \dots \\ + \text{Highpitch} \\ - \text{Lowpitch} \end{array} \right] \left[ \begin{array}{l} + \text{syllabic} \\ - \text{high} \\ - \text{round} \\ \dots \\ - \text{Highpitch} \\ + \text{Lowpitch} \end{array} \right]$$

/áà/

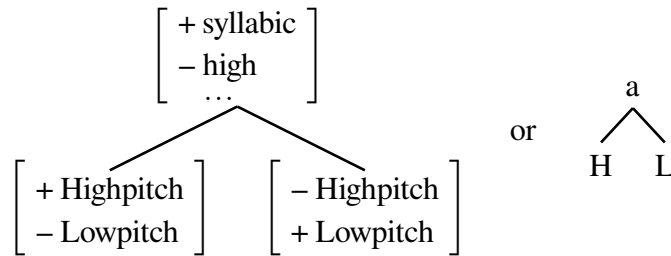
Goldsmith (1976) points out that the representation in (3) cannot be used for short contour tones, since short vowels will necessarily consist of a single segment. If a single segment's feature matrix contains two sets of values for tone features, as in (4), then there is no obvious way to distinguish between a falling and a rising tone, because the features have no inherent ordering. Furthermore, standard assumptions about feature specification give no way to arrive at features with two simultaneous values.

(4) *Untenable representation for a short contour tone* (Goldsmith 1976: 38, adapted)

$$\left[ \begin{array}{l} + \text{syllabic} \\ - \text{high} \\ - \text{round} \\ \dots \\ + \text{Highpitch} \\ - \text{Lowpitch} \\ - \text{Highpitch} \\ + \text{Lowpitch} \end{array} \right]$$

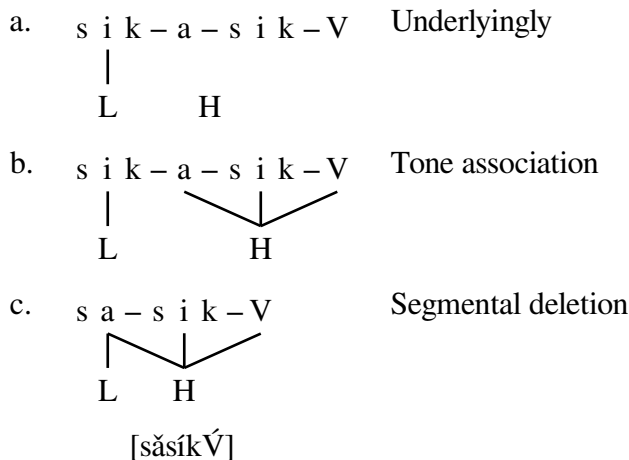
By defining tones as features of segments (i.e., autosegments) that exist on a separate tier as in (5), however, it is possible to intuitively capture the notion of a single segment being specified for two sets of tonal features. Autosegmental representations also inherently draw a distinction between rising and falling contours, because the tier of tonal features is segmented analogously to vowels and consonants; if the high tone takes linear precedence to the low tone, then there is a falling contour (from high to low), whereas the opposite order leads to a rising contour (from low to high). Segments on the two tiers are related to each other by “association lines”, which are not allowed to cross. This ensures that linear segmentation is not interrupted on either tier.

(5) *Autosegmental representation of a rising tone* (Goldsmith 1976: 42, adapted)



Another advantage of the autosegmental approach to tonal features is that it predicts the property of “stability”. When a vowel is deleted in a tone language, its tone specification does not usually disappear, as would be expected if tones were on equal footing with the rest of the segment’s features. Rather, it is typical for the tone that was originally associated with the deleted vowel to remain and be realized on an adjacent vowel. An example of this is seen in (6), adapted from Goldsmith (1976). The H tone in (6a), which is not associated with any tone bearing units, links to all free tone bearing units in (6b). In (6c), a deletion process applies on the segmental tier, removing the first sequence /ik/. The tone bearing unit that /L/ was linked to, /i/, has been deleted, but the tonal autosegment is not deleted along with it. Instead, /L/ links to /a/, which is the only tone bearing unit to which it can link without crossing association lines. /a/ is now associated to two sets of tonal features, /L/ and /H/, and is consequently realized with a rising contour tone.

(6) *Tone stability* (Goldsmith 1976: 60, adapted and simplified)



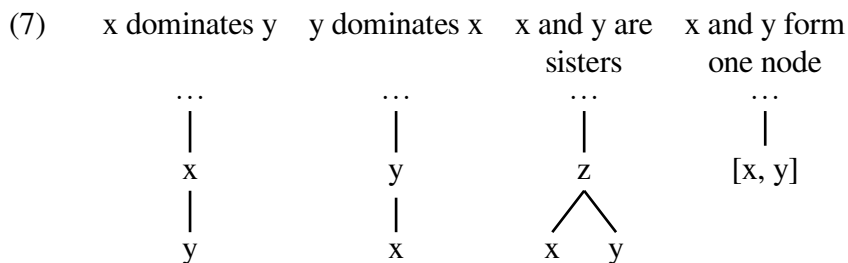
Goldsmith points out that while it is generally *possible* to account for these facts with a complicated series of segmental tone feature copying rules that must be stipulated to apply in all cases where vowel deletion would occur, this solution fails to capture the generalization about tone features being preserved: why copy and save tone features but not other segmental features? This is particularly pertinent considering that rules in many tone languages that copy whole sets of segmental features generally copy everything *except* for the tone features (Goldsmith 1976: 57). If autosegmental representation of tone is assumed, then tonal stability is expected, because deletion of a segment on one tier (a vowel) does not affect segments associated with it on another tier (a tone). Likewise, copying of features on one tier (segmental features) can occur without regard for features on another tier (tonal features).

Autosegmental representations also afford an elegant account of a phenomenon known as “floating tone”. Many tone languages show morphemes that seem to consist only of a tone, contributing no segmental material at all. In pre-autosegmental analyses, floating tone needed to be formalized as a segment which was specified only for its tone features, which underwent merger with other segments in order to attribute these features. Goldsmith (1976) proposes that floating tones are morphemes that exist only on the tonal tier, initially lacking association to any other segments. This allows surface morphological patterns that are spelled out with only a change in pitch to be analyzed with suffixes that contain floating tones, rather than with tone-altering morphological rules.

### **2.1.2 Tonal feature geometries**

Prior to autosegmental phonology, features were assumed to be divided into segments, but were not otherwise organized with respect to each other. The autosegmental approach assumes that segments exist on separate tiers and thus are in a structural relationship, but the features themselves were originally assumed to be unordered within the segment. This changed with the study of feature geometry, beginning with Clements (1985) and subsequently developed further (see Clements and Hume 1995 for an overview of foundational work). Feature geometry is based on

the idea that features are organized in hierarchical trees whose structures are posited based on phonological patterning. It is assumed that processes operate only on groups of features that form constituents, so that phonological rules can be written minimally to affect only a single node of the tree. Clements and Hume (1995: 267) lay out the possible structural relations in (7).

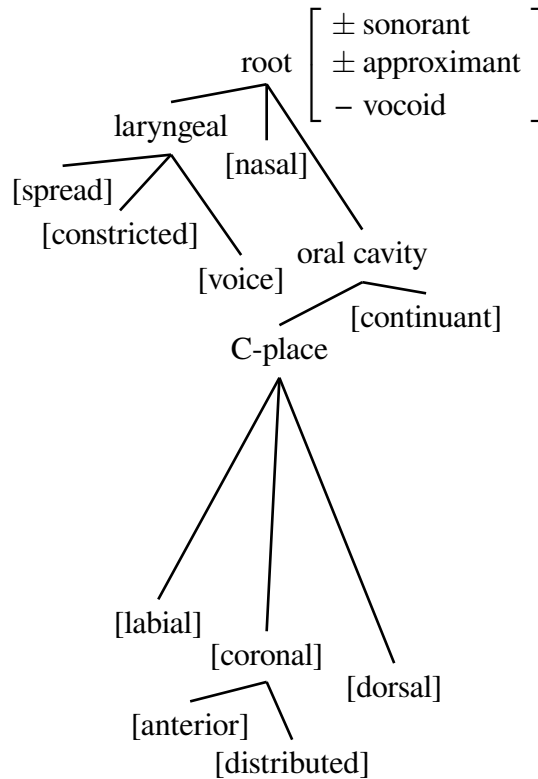


The criteria for determining these relations based on phonological activity are described as follows (Clements and Hume 1995: 267):

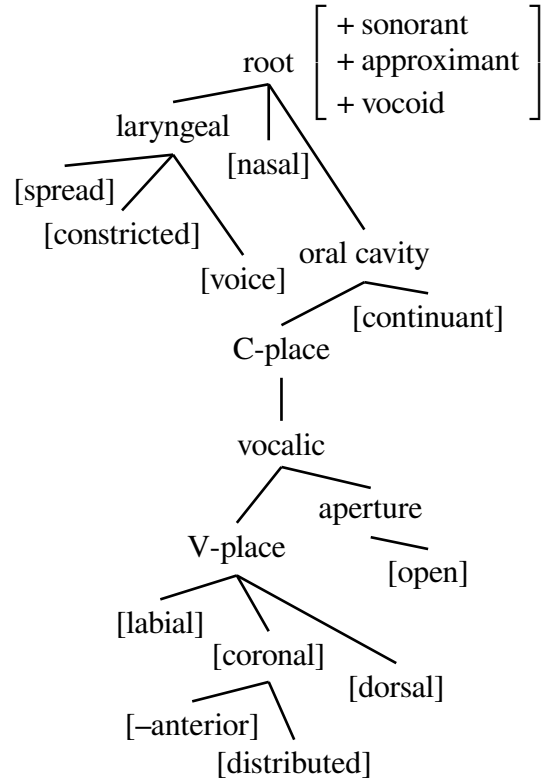
If an operation on  $x$  always affects  $y$ , but not vice versa, the first configuration is motivated. If an operation on  $y$  always affects  $x$ , but not vice versa, the second is required. If  $x$  and  $y$  can be affected independently of each other, they are each independently linked to a higher node  $z$ , as shown in the third figure. Finally, if an operation on one always affects the other, they form a single node, as shown at the right.

Many authors have examined only small portions of feature structure, but when feature geometries are studied in their entirety, the relationships evinced can be quite complex. For example, based on a number of cross-linguistic segmental phenomena, (Clements and Hume 1995: 292) arrive at the complete hierarchies in (8).

(8) a. Consonants:



b. Vocoids:



Since Clements (1985), feature geometric principles have been applied to tone features as well, leading to proposals for hierarchical structure within tonal autosegments. Autosegment-internal features are joined together at a tonal root node,<sup>2</sup> notated “o”, which bears autosegment-internal feature structure and is the point to which tonal autosegments are associated to the segmental tier. Differences in internal structure are then used to analyze different tonological activity.

Yip (1989) considers several classes of contour tones, which have similar phonetics but show distinct behaviour. This is captured by distinguishing the representations in (9) (Yip 1989: 150). Languages like Chinese, which treat contour tones as whole units, use representations like (9a); the tonal features L and H branch to form a contour within the autosegment. The root node thus bears a full contour tone and links to the tone bearing unit (here, the syllable) through a single association line. In other languages, such as many Bantu varieties, contour tones seem to be composed of

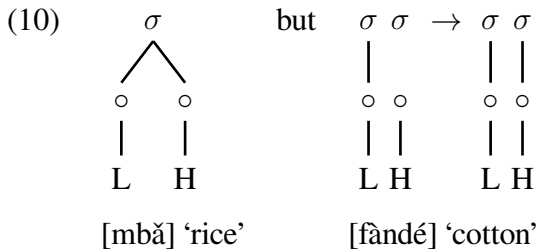
<sup>2</sup>The tonal root node (sometimes abbreviated “TRN”) has been used to varying degrees in many studies on the representation of tone, but the earliest reference to it seems to be in an unpublished manuscript by Archangeli and Pulleyblank (1986).

two separate autosegments, i.e., they are “tone clusters” as in (9b). Each autosegment has its own root node, so that autosegments have no internal branching; contours link to the tone bearing unit by two association lines, one from each root node.

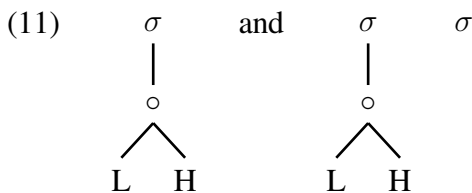
- (9) a. *Branching tone/single melodic unit*    b. *Tone cluster*



In a tone cluster language, the mapping of tonal root nodes occurs in sequence from left to right, with leftover tones being mapped to the final tone bearing unit.<sup>3</sup> Thus in Mende a sequence of /L/ and /H/ on the tonal tier will form a rising contour in a monosyllabic word but lead to consecutive low and high toned syllables in a bisyllabic word. Yip (1989: 151) illustrates this as follows:



In branching tone languages, however, where contours are not composed of corresponding level tones, tonal mapping treats contours (which share a tonal root node) as single units. A contour will consequently map as a whole, and will not split into its component parts even if another tone bearing unit is available (Yip 1989: 151):



<sup>3</sup>It is not clear that left-to-right mapping of tones to TBUs is a universal, and some languages have been argued to use right-to-left mapping instead. See Zhang (2007) for discussion of arguments for and against this view.



It should be pointed out that the distinction between these possibilities is largely based on the implicit assumption that featural representations are determined by surface phonetics (i.e., a phonetic rise in pitch entails a representation containing a sequence of the phonological elements L and H) while architectural representations (i.e., the structure of feature geometries) are determined based on phonological activity. In the contrastivist approach, these two notions are explicitly incompatible: both featural and architectural aspects of representations must be informed by phonological activity, not only by substantive properties of phonetics.

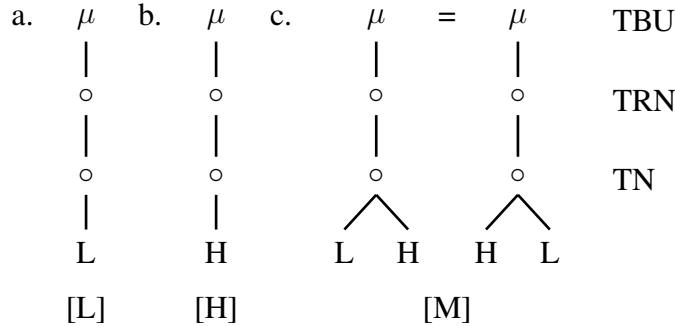
Hyman (1993) takes the study of autosegment-internal structure even further. Hyman (1993: 76) assumes that all tonal representations are composed of varying configurations of the two tone features defined in (12), which he established in earlier work (Hyman 1986).

- (12) Tone features: H = at or above a neutral reference tone height  
 L = at or below a neutral reference tone height

He expands the autosegment-internal structure proposed by Yip (1989) with the addition of an extra tier, as shown in (13) (Hyman 1993: 77, adapted). Autosegments are still organized around a tonal root node by which they connect to the tone bearing unit (here exemplified by a mora rather than a syllable), but within the autosegment, there is an additional “tonal node” available for features to link to.

- (13) Tone bearing unit (TBU):  $\mu$   
   |  
 Tonal root node (TRN):         $\circ$   
   |  
 Tonal node (TN):                $\circ$

Low and high tones are represented by the linking of L and H features to the tonal node. However, unlike branching under Yip’s (1989) tonal root node, feature linking does not have inherent linear ordering and cannot play a distinctive role in Hyman’s (1993) model. Thus a tonal node bearing both L and H as in (14c) does not reflect a contour, but is instead used to represent a level mid tone. This avoids the need for a separate feature M; mid tones are specified by combining the two available features on a single tier.

(14) *Basic level tones* (Hyman 1993: 77, adapted)

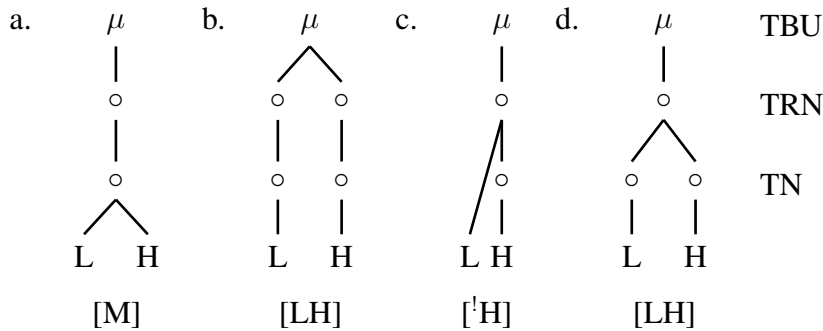
The tonal root node is used to represent what is known as tonal “register”. In addition to contrasting up to three basic tone levels (L, M, and H), many tone languages show register effects in certain configurations or intonational positions. The most common is “downstep”, where, for example in a sequence of two high tones, the second is pronounced at a lower pitch than the first, but not as low as a phonological low tone. Downstep is commonly notated with a superscript “!” preceding the downstepped tone, so that a sequence of high followed by downstepped high is written as H<sup>!</sup>H. Hyman (1993: 79) points out that although downstep is most common for high tones, it is attested cross-linguistically for all three tone levels. Register effects can also cause raising through a phenomenon known as “upstepping”. An upstepped tone, notated with a superscript “↑”, is pronounced at a higher pitch than the preceding tone at the same level, such that the pitch goes up between the two phonological high tones in the sequence H<sup>↑</sup>H.

Contrary to many other studies on the representation of tonal register, which use a separate set of register features, Hyman implements register via linking of the same H and L features directly to the tonal root node. Hyman (1993: 80) gives the representations for all possible surface tones as in (15), where features on the “T(onal)-plane” link to the tonal node and those on the “R(egister)-plane” link directly to the tonal root node. Tones with no features on the R-plane, as in (15a), (15b), and (15c), do not show register effects.

(15)	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
T-plane:	H	L	LH	H	L	LH	H	L	LH
R-plane:	∅	∅	∅	L	L	L	H	H	H
TONE:	H	L	M	<sup>!</sup> H	<sup>!</sup> L	<sup>!</sup> M	<sup>↑</sup> H	<sup>↑</sup> L	<sup>↑</sup> M

The possible structural configurations are shown in (16). (16a), with both H and L connecting to the tonal node, reflects a single mid tone. In (16b), two separate tonal root nodes connect to a single tone bearing unit, reflecting a composed contour tone (those which Yip 1989 calls “tone clusters”). (16c) is a downstepped high tone; the H linked to the tonal node specifies a high level tone, while the L linked to the tonal root node specifies downstepped register. Finally, (16d) reflects what Yip (1989) calls a “branching tone”, where a contour is treated as a single melodic unit. Branching is within the autosegment, but is represented by two tonal nodes linking to a single tonal root node, not by branching at the level of features.

(16) *Possible representations* (Hyman 1993: 81, adapted)



Studies in tonal feature geometry have made progress towards a unified architecture for tonal features, but I would argue that the greatest finding has been that surface phonetics should be given a minimal amount of consideration in determining the formal properties of representations. Yip (1989) and Hyman (1993) both require non-cluster contour tones (“branching tones”) to be complex on the featural level, despite the key observation that they are treated as single units by the phonology, i.e., they are not composed. Such an analysis would make sense if features were meant to serve as a notation for phonetics, but evidence in feature geometric research is drawn explicitly from phonological activity, not phonetics. It could also be that features are meant to serve a

dual purpose of reflecting both phonological patterns *and* surface phonetics, but the universal phonetic interpretation of tone features is undermined in feature geometric studies themselves. For example, Hyman (1993: 82, fn. 6) notes that downstepped <sup>l</sup>H is much lower in some languages than in others, to the point that it sometimes resembles M tone, but that “[t]hese variations can be accounted for if we assume that the features, especially in their register function, provide different instructions to the phonetic implementation rules in different languages.” If these phonological primes are so variable, then why assume that they are universal in their phonetic substance?

More recently, Hyman (2011) has concluded that it is not possible to devise a single set of tone features that can “do all tricks” in terms of capturing cross-linguistic tonological activity, although it is not clear what the nature of features should be in order to account for the diversity of tone systems. While there have been many proposals for tonal feature structure, few have been pursued in much detail. However, it is clear that there is motivation for tonal autosegment-internal structure. What is needed is a model of the representation of tone which has the predictive power and elegance of autosegmental phonology but the flexibility of non-universal feature structures.

## 2.2 Stress

In word-level phonology, the term “stress” has generally been used to describe a phenomenon in many languages where a certain position has a higher degree of “prominence” than other positions. When its phonetics are observable, a stressed position typically has one or more of increased pitch, duration, and intensity, as well as more carefully articulated vowel quality, depending on the language and the linguistic context (Fry 1955; Lehiste 1970). Stress has traditionally been considered to be distinct from tone in a linguistically significant way, so that generative phonologists have accounted for “tone systems” and “stress systems” within completely different frameworks—autosegmental theory for the former and foot-based metrical theory for the latter, the development of each being based principally on evidence from its own domain. This distinction is so prevalent that even some typological studies that attempt to take the entire ecosystem of prosodic systems

into account under a single group of terms (e.g., Hyman 2006, 2009) rely on the assumption that stress is *a priori* metrically based.

Regardless of the phonological interpretation of stress, it is clear that in a number of languages, the position of stress is non-predictable and can be used to express contrasts. Spanish, for example, allows contrasts such as the following (Roca 2005: 349):

- (17) a. sa.bá.na ‘savannah’  
 b. sá.ba.na ‘bed sheet’

The words in (17) constitute nouns with the same morphological structure, so it must be the case that the position of stress is stored as some kind of lexical information. However, Spanish also allows the position of stress to vary based on morphological information (Harris 1995: 869):

- (18) a. so.li.ci.tó ‘s/he solicited’  
 b. so.li.cí.to ‘I solicit’  
 c. so.lí.ci.to ‘solicitous’

For a language like Spanish, unpredictability and the presence of numerous minimal pairs leaves no doubt that stress is a phonological phenomenon, but the term “stress” has been used to describe phenomena whose phonological grounding is not so clear. In this section, I will consider what counts as data for stress and address the major proposals for capturing it phonologically. I will then argue for the treatment of lexically contrastive stress as an accentual autosegment and propose a unified formal understanding of stress and tone.

### 2.2.1 Stress and accent

The notions of “stress” and “accent” have been treated diversely in the literature on prosodic features, both terminologically (what phenomena do they refer to?) and theoretically (are they different, and if so, what are their phonological properties?).

Given the present object of study—the nature of contrastive representations and their relationship to non-contrastive phonetic realizations—the most useful conceptions of “stress” and

“accent” are those espoused by, e.g., Fox (2000) and van der Hulst (2014b). Here, the term “stress” refers to the phonetic correlates associated with prominence, generally one or more of increased pitch, duration, and intensity. “Accent”, on the other hand, refers to abstract phonological marking of some position in the phonology and/or in the lexical entries of morphemes. In some languages, the phonetic realization of accent may be that of stress (several correlates), but in other cases accent may be spelled out simply as pitch or length, or perhaps have a different phonetic correlate or no phonetic correlate at all (Hyman 2006; Dubina 2012; van der Hulst 2014b). A similar distinction is at least implicit in much of the classical literature on metrical phonology, which assumes abstract organizational rhythmic structure but recognizes that “stress” has no universal (or in some cases, even language-specific) phonetic correlates (Kager 1995b). The main difference between this and the present view concerns the nature of the abstract representations.

Just as the marking of accent does not imply the presence of stress, the presence of stress does not imply the marking of accent. In many cases, “stress” may not necessarily be reflective of word-level phonological contrasts, but rather are due to utterance-level speech rhythm or functional consequences of edge marking or syllable weight (van der Hulst 2012). Sometimes correlates may be manifested at the phrasal level but are reported to be at the lexical level because stress has only been investigated by looking at single-word utterances (Gordon 2014).

If accent is a kind of abstract positional marking, then the study of contrastive representations must be concerned with the formal nature of that abstract marking. What kind of phonological element is accent? In languages where stress is always predictable, it may not be necessary to make specific claims about the formal nature of accent (if any). However, in the case of unpredictable stress, it is necessary to make a claim about how lexical exceptions are represented.

Such specifications have been handled with a number of different mechanisms in the literature on stress, depending on the overall theoretical approach taken. In metrical phonology, which posits a kind of structural relationship used specifically for representing stress, it is possible to specify a subset of that structure underlyingly (Liberman and Prince 1977; Hayes 1995), or similarly to specify underlying metrical grid marks (Halle and Vergnaud 1987). As an alternative,

authors have devised ways of limiting or altering metrical structure, such as morpheme-specific parsing algorithms (Idsardi 1992; Halle and Idsardi 1995; Doner 2013) or morpheme-specific application/non-application of extrametricality rules (Selkirk 1984).

Analyses of lexical accent have also been stated in terms of formalisms that interact with other aspects of the phonology. For example, Goldsmith (1976) uses a special diacritic asterisk to mark accented positions, to which phrase-level intonational contours can link autosegmentally. It has also been proposed that accent is a special kind of autosegment that is used together with foot structure (Hagberg 1993, 2006), or, more simply, that stress is one of many possible phonetic spellouts of tonal autosegments (Dubina 2012).

Finally, surface stress patterns have been represented using the formal mechanisms of quantity contrasts. Bucci (2013) proposes that accent in Corato Italian is phonologically “virtual length”; i.e., although there is no vowel length distinction in the language, accent is represented with a formal vowel length contrast. The idea of analyzing accent with a quantity distinction is by no means new, however. Pater (1994) points out that such “honorary weight” was used by Chomsky and Halle (1968) to specify exceptionally stressed open penultimate syllables containing lax vowels in English; these vowels were followed by clusters of identical consonants so that the stressed syllable was closed for the purposes of stress assignment, after which a degemination rule applied, simplifying the cluster. In terms of contrast, of course, there is no difference between a language with predictable quantity-sensitive stress and a language with lexical stress represented by quantity. Both kinds of language allow non-predictable complex rhymes, and both spell them out phonetically as “stressed”, but there need not be separate phonological accent in either. The only difference is that “honorary weight” languages are those which have been described in the literature as lexical accent systems versus lexical quantity systems.

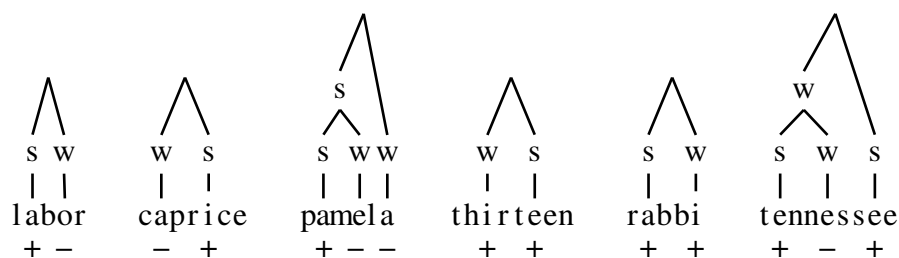
Keeping in mind the emphasis on the contrastive phonological role of accent, I will now discuss the key aspects of the predominant approach to analyzing stress.

### 2.2.2 The metrical enterprise

The family of theories known as “metrical phonology” (developed in its original incarnation by Liberman 1975, Liberman and Prince 1977, and Halle and Vergnaud 1978) has seen widespread adoption in the analysis of word stress. Metrical theory is based on the observation that in many languages, the location of primary stress is completely predictable relative to the beginning or end of the word and often co-occurs with alternating beats of secondary stress. Metrical phonology accounts for this by positing “feet”, groupings of syllables that serve to constrain patterns of stress assignment. Language-specific differences in stress patterns are accounted for by parametrizing various properties of feet.

In early metrical phonology, feet were represented by binary tree structures consisting of strong (“s”) and weak (“w”) branches, which reflected relative prominence of syllables. Liberman and Prince (1977) assumed that metrical structure was built on top of syllables specified for a [ $\pm$ stress] feature, with the restriction that a [ $-$ stress] syllable could not serve as the strong element of a foot. A metrical tree with two elements must have one strong and one weak branch, with a binary tree being built upwards for words of greater length:

(19) *Stress with metrical trees* (Liberman and Prince 1977, adapted)



The rhythmic properties of strong and weak branches were reflected in a separate level of representation known as the “metrical grid”. The metrical grid consisted of a lines of “grid marks” stacked on top of each potential stress bearer. The stress bearer with the largest number of grid marks receives primary stress, and relative prominence decreases with smaller numbers of grid marks. Various notations have been used for grid marks, including numbering based on the amount of grid marks (Liberman 1975), consecutive numbering of each grid mark (Liberman



and Prince 1977), and stacking of a single symbol such as “\*” (Halle and Vergnaud 1987) or “×” (Prince 1983; Hayes 1995). The first line of grid marks indicates stress bearing units (syllables), the second line marks relative stress at the foot level, and the third line marks the highest level of stress in the word. The metrical grids for the words in (19) would appear as follows:

(20)

×	×	×	×	×	×
×	×	×	×	×	×
×	×	×	×	×	×
×	×	×	×	×	×
labor	caprice	pamela	thirteen	rabbi	tennessee

The application of metrical trees and their spellout to metrical grids applies recursively across phrases, so that grid-based representations can be given for stretches greater than a single word. The “peaks” and “valleys” of the metrical grid give a straightforward representation of adjacent prominence among syllables, which has led to certain phenomena being accounted for by allowing the phonology to directly manipulate grid marks. A common example is the English Rhythm Rule, which shifts the highest level of prominence (i.e., swapping primary and secondary stress) in the first word of certain constructions:

(21) *The English Rhythm Rule*

thirtéén → thirtèen mén  
 Tènnessée → Tènnessèe Érníe

The corresponding metrical grids for these examples show that the Rhythm Rule results in moving a grid mark backwards from the otherwise most heavily stressed syllable in the first word:<sup>4</sup>

(22)

×	×	×	→	×	×
×	×	×		×	×
×	×	×		×	×
×	×	×		×	×
thir	teen	men		thir	teen men
×	×	×	→	×	×
×	×	×		×	×
×	×	×		×	×
×	×	×		×	×
Tennes	see	Ernie		Tennes	see Ernie

<sup>4</sup>This is considered quite fundamental in some grid-based theories of stress; Hayes (1995) handles processes such as this by positing a generalized operation of “move X”, which changes the positions of grid marks. Note that in (22), an additional line of grid marks is added to mark the highest level of prominence at the phrasal level.

The focus in further development of metrical theory moved from word-level and phrasal stress in English to the cross-linguistic typology and parametrization of foot structure, abandoning the feature [ $\pm$ stress] (Hayes 1980, 1995; Halle and Vergnaud 1987; Idsardi 1992; Halle and Idsardi 1995). Although the details have been restated and revised many times, there is a general flavour as to the type of parameters proposed in the literature, which Hayes (1995: 54) summarizes as follows:

- (23) a. **Choice of foot type**
- |                          |  |
|--------------------------|--|
| i. SIZE                  | Maximally unary/binary/ternary/unbounded                       |
| ii. QUANTITY SENSITIVITY | Heavy syllables (may/may not) occur in weak position of a foot |
| iii. LABELING            | Feet have (initial/final) prominence                           |
| iv. OBLIGATORY BRANCHING | The head of a foot (must/need not) be a heavy syllable         |
- b. **Direction of parsing** Left to right/right to left
- c. **Iterativity** Foot construction is (iterative/once only)
- d. **Location** (Creates new metrical layer/applies on existing layer)

The parameters in (23a) are essentially well-formedness restrictions on feet themselves and the units on which they can be built. (23b) and (23c) determine the side of the word from which feet are parsed and whether words contain multiple feet or only one foot. Finally, (23d) concerns the relationship between foot structure and the metrical grid.

The most widely accepted of these, even in descriptive works, have been labeling (also known as headedness), direction of parsing, and quantity sensitivity. In a non-iterative system, only a single foot is built. This foot is placed at either the left or right edge of the word, depending on direction of parsing. Stress is then assigned to the first or last syllable of each foot depending on headedness. Assuming bisyllabic feet, these three parameters make it possible to derive the four patterns schematized in (24):

(24) *Headedness and directionality with non-iterative feet*

	Left to right	Right to left
Left headed	$(\acute{\sigma}\sigma)\sigma\sigma\sigma$	$\sigma\sigma\sigma(\acute{\sigma}\sigma)$
Right headed	$(\sigma\acute{\sigma})\sigma\sigma\sigma$	$\sigma\sigma\sigma(\sigma\acute{\sigma})$

In an iterative system, feet are built continuously until no groups remain to parse. When a word contains multiple feet, it is necessary to determine which foot will contain the primary stress. Hayes (1995) does this by parametrizing an “end rule”, which promotes either the leftmost or rightmost head to the highest level of prominence. In languages with end rule left, this means that the leftmost stress is always the primary stress:

(25) *Iterative feet with end rule left*

	Left to right	Right to left
Left headed	$(\acute{\sigma}\sigma)(\grave{\sigma}\sigma)\sigma$	$\sigma(\acute{\sigma}\sigma)(\grave{\sigma}\sigma)$
Right headed	$(\sigma\acute{\sigma})(\sigma\grave{\sigma})\sigma$	$\sigma(\sigma\acute{\sigma})(\sigma\grave{\sigma})$

With end rule right, it is instead the rightmost foot whose head receives the primary stress:

(26) *Iterative feet with end rule right*

	Left to right	Right to left
Left headed	$(\grave{\sigma}\sigma)(\acute{\sigma}\sigma)\sigma$	$\sigma(\grave{\sigma}\sigma)(\acute{\sigma}\sigma)$
Right headed	$(\sigma\grave{\sigma})(\sigma\acute{\sigma})\sigma$	$\sigma(\sigma\grave{\sigma})(\sigma\acute{\sigma})$

Quantity sensitivity allows a language to interrupt foot parsing in order to ensure that a heavy syllable serves as the head of a foot. The criteria for heavy syllables are known to show a large degree of variation, both across and within languages (Gordon 2006), but heavy syllables are generally considered to be those that contain a complex rhyme (a long vowel and/or a coda consonant).

A final parameter employed in metrical phonology is extrametricality, which allows phonological material to be invisible to the foot parsing algorithm. Extrametricality is usually assumed to apply to the edges of the word domain, most often word-finally. It can affect whole syllables, or

only parts of syllables, such as a word-final consonant or mora, which causes the metrical system to treat word-final heavy syllables as light.

While metrical theory has been widely adopted in the study of stress phenomena and seems to do a good job of accounting for the data as they are typically described, it is not obvious that all of the assumptions upon which the research program has been built are compatible with the contrastivist approach to phonology. One of these assumptions, regarded as a triumph of metrical phonology, is the conflation of secondary stresses with primary stress, deriving the two from the same foot structure, with one foot being identified as that bearing primary stress. This unification certainly offers an appealing account of stress and rhythm, but contrary to what is commonly presented in the literature, on closer inspection these claims seem to overshoot their empirical basis.

As types of evidence for stress as a single entity, Hayes (1995: 9–21) lists stress-conditioned segmental processes (such as elision processes like flapping), the alignment of phrasal intonation contours relative to stressed positions, and the Rhythm Rule (as discussed above). Hayes argues that, because all of these refer to a single position, the one identified as stressed, stress should be considered a single phonological category. The problem is that all of these phenomena seem to refer to primary stress (i.e., accent), not uniformly to the entirety of foot structure; observations about primary stress cannot be *a priori* extended to secondary stress. Moving beyond English, the cross-linguistic evidence upon which metrical theory has been built is even less convincing. It is true that “stressed” positions often show privileged phonological status (de Lacy 2006; Walker 2011), but many if not all of these seem to be observations about primary stress, not of “stress” as a whole. De Lacy (2014) extensively discusses the unreliability of evidence for stress systems that has been used in the metrical literature, much of which has been taken from short impressionistic descriptions in old reference grammars rather than explicit phonological evidence or primary phonetic investigation. Hayes (1995) himself admits that this kind of evidence is not reliable.

In cases where a greater amount of metrical structure does appear to be referenced by the phonology, it has often been necessary to introduce new assumptions about how feet are formed

or how rules are formulated in order to employ coherent footing (see e.g., Dresher and Lahiri 1991; Bennett 2013). Rarely does the straightforward application of the basic tenets of metrical theory have phonological consequences outside of the context of accent, which leads to the conclusion that at least the metrical structure used to account for most secondary stress is not motivated in the lexical phonology.

More reconcilable with contrastivist methodology is the so-called “accent first”/“main stress first” approach, argued for extensively in work by Harry van der Hulst (e.g., van der Hulst 1984, 2012; Goedemans and van der Hulst 2014; van der Hulst 2014a, among others). Contrary to the traditional metrical approach, where fully-specified feet based on (often non-contrastive and impressionistically evidenced) rhythm are parsed before one foot is picked out as the bearer of primary stress, accent first assumes that abstract accent is first identified, followed by the assignment of rhythm using a non-metrical mechanism. That is, a single phonological marking of “accent” is first located, which will be realized as *primary* stress. Only then is secondary stress calculated. In more traditional terms, this means that foot parsing is top-down rather than bottom-up; the first step is to identify the most prominent position (primary stress), rather than to lay down a full set of feet. Goedemans and van der Hulst (2014) note that Hayes (1995) chooses bottom-up parsing more or less arbitrarily, even though he himself points out that for the majority of systems, it is not possible to tell the two analyses apart. Accent first simply makes the opposite choice by assuming top-down parsing.

While the accent first approach makes minimal assumptions about the degree to which metrical structure is specified in the phonology, it does not preclude the use of binary feet and/or headedness to identify accented positions. Even for languages in which the position of primary stress is by and large predictable, it is possible for accent to be present in the lexical phonology: Goedemans and van der Hulst (2014: 127) “assume that as soon as a language allows for exceptional stress locations, *all* stress locations are dependent on an accentual algorithm.” In their model, this algorithm is based on metrical principles. Metrical phonology is avoided, however, for dealing with secondary stresses which are entirely predictable and exceptionless. For van der

Hulst (2012, 2014a), this is done with only a metrical grid. Unlike Prince (1983) or Gordon (2002) however, who attempt to account for all metrical phonology using only grid marks and no foot structure, the accent first approach seeks to avoid conflation of accent and stress. Secondary stress is described with a different mechanism to keep it separated from accent, even in languages where it is analytically possible to unify the two phenomena. It should be borne in mind that, in an explicitly contrast-minded approach such as the present study, the grid mark formalism used by van der Hulst should not be interpreted as a reflection of phonological categories, but rather as a schematic statement of language-specific phonetics—the prosodic equivalent of redundant features.

In languages with accent, there is often a default pattern which can be described in metrical terms. Whenever there exists non-default accent, however, some questions remain: How are lexical exceptions to the default pattern encoded? How is accent formally represented (default or otherwise)? Most work in the metrical tradition assumes that foot structure itself is the representation of stress and accent, and that foot structure is assigned by some kind of parsing algorithm. If the specification of accent is necessarily cast in the same terms as foot structure, then lexical stress must be specified by somehow forcing a particular parsing. This has been implemented variously by specifying metrical structure underlyingly (used as early as Liberman and Prince 1977) and specifying parsing algorithms in some morpheme-specific way (Idsardi 1992; Halle and Idsardi 1995; Doner 2013). Alternatively, the effect of lexical stress has been achieved by defining extrametricality not as a language-wide parameter, but as morpheme-specific rule (Selkirk 1984) or even as an element in underlying representations (Hayes 1995), in order to coax the phonology into assigning primary stress to a non-default position.

As can be seen from a brief look at the literature on metrical phonology, there are many possible mechanisms that can be invoked in order to specify lexical stress, and it is not clear how many of them may be active in a single language. Taking this into consideration, the restrictive predictions that metrical theory is said to make about possible stress systems is compromised.

The alternative to an algorithm-centric analysis of accent is to assume that accent is repre-

sented with some kind of positional marking rather than with metrical structure. In this view, lexically accented positions would be marked directly in underlying forms and a default accent algorithm would apply only in the case that accent needs to be inserted.<sup>5</sup> Following accent assignment, in line with the accent first approach, fully predictable stresses are applied in phonetic implementation. In this conception, the representational role played by feet is marginalized, opening the door for an analysis in which accent can be represented not with a separate metrical module of the phonology, but with a formal mechanism widely motivated in phonological theory: the autosegment.

### 2.2.3 Accent as an autosegment

The idea of treating accentual systems autosegmentally is not a new one. For example, Japanese pitch accent, while spelled out on the surface primarily with various pitch contours, seems to involve the accentual marking of a single position underlyingly. Hyman (2006: 230) points out, however, that the nature of this marking has been the subject of much debate in the literature, with both “accentual” and “tonal” analyses.<sup>6</sup> An accentual account might involve lexical specification using metrical principles, but a tonal analysis of the same data by marking accented positions with a tonal autosegment (such as H).

An autosegmental approach to a system like Japanese is obvious because of its surface pitch-based nature, but so-called stress accent systems have also been analyzed autosegmentally. The first author to do this explicitly seems to have been Hagberg (1993, 2006), but Hagberg points out that Halle and Vergnaud (1987) assume this at least implicitly; building on Liberman (1975), they refer to a “separate autosegmental plane” for the representation of stressed phonemes (Hagberg 2006: 58–59). In Hagberg’s model, accent is represented as a special kind of autosegment notated “\*”. In the case of lexical accent, this autosegment can be linked underlyingly to the position

---

<sup>5</sup>The position of this default accent can be calculated using any number of principles, including a parametrized metrical system.

<sup>6</sup>Hyman cites Haraguchi (1977), McCawley (1978), Poser (1984), and Pierrehumbert and Beckman (1988), but this is by no means an exhaustive list of works on the topic; see also Byrnes (2011).

bearing primary stress, meaning that lexical accent languages are accent first.

While Hagberg does away with the structural representation of accent, he still requires foot structure. The \* autosegments are used not simply as markers of lexical accent, but for the representation of all stress, both primary and rhythmic. Binary feet are parsed as in traditional metrical theory, but feet are not inherently headed. Instead, the assignment of stress to the first or second element of a foot is due to a separate parameter which dictates the linking of autosegments within the domain of feet. The same formal autosegment is assigned for both primary stress and rhythmic stress, so these are undesirably conflated; in fact, it seems that Hagberg only uses feet in order to derive the placement of rhythmic stress, so with the exception of how it treats lexical accent, his model is largely a restating of classical metrical accounts without the use of headedness.

It is in analyzing cases of *lexical* accent that the advantages of Hagberg's account become apparent. Consider the data from Mayo (Uto-Aztecan) in (27):

(27) *Mayo stress patterns* (Hagberg 2006: 130, adapted)

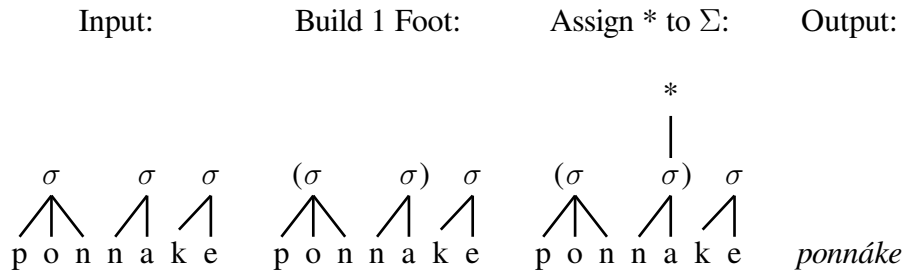
First Syllable Stress:	Second Syllable Stress:
a.      chúpnake 'will harvest (tran)' híchupnake 'will harvest (intr)' híhichupnake 'will always harvest'	b.      ponnáke 'will play (tran)' hipónnake 'will play (intr)' hihíponnake 'will always play'
c.      bwá'anake 'will eat (tran)' hí'ibwanake 'will eat (intr)' híhi'ibwanake 'will always eat'	d.      ba'áte 'irrigate (tran)' hibá'ate 'irrigate (intr)' hihíba'ate 'always irrigate'
e.      chíknake 'will sweep (tran)' hína'ikia 'count (intr)' híhichiknake 'will always sweep'	f.      wiséka 'sawing (tran)' hichíweka 'shelling (intr)' hihíwiseka 'always sawing'
g.      ná'ikia 'count (tran)' hína'ikia 'count (intr)' híhina'ikia 'always count'	h.      chiwéka 'shelling (tran)' hichíweka 'shelling (intr)' hihíchiweka 'always shelling'

Stress falls on either the first or second syllable of a word depending on the root morpheme, regardless of the number of prefixes added. Stress is thus both fixed at a certain distance from the word edge, but lexically specified for the distance from the word edge. Hagberg analyzes this as



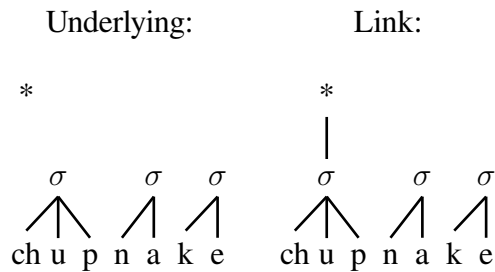
a contrast between roots with and without floating accent. Roots with second syllable stress lack underlying accent, so the input contains no \* autosegments. A single bisyllabic foot is parsed at the left edge of the word, after which a \* is assigned to the second syllable of that foot. This leads to default second syllable stress:

(28) *Derivation of second syllable stress in Mayo* (Hagberg 2006: 131)



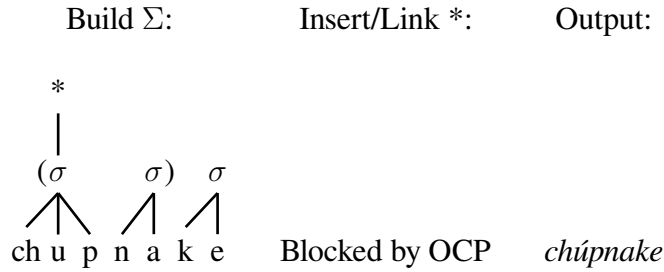
Roots with first syllable stress contain a lexical accent specification, but this takes the form of a floating \* rather than one which is underlyingly linked to any particular position. When forms are derived, \* links from left to right similarly to autosegmental association rules that have been proposed for tone languages:

(29) *First syllable accent linking* (Hagberg 2006: 132)



This is followed by the building of a foot. A \* is crucially not inserted on the second syllable because of the obligatory contour principle (OCP), which prevents identical elements from being immediately adjacent on the same tier. The OCP is widely invoked in the literature on tonal autosegments, and is directly applicable in Hagberg’s model. The result is an output with a single word-initial stress:

(30) *First syllable accent parsing* (Hagberg 2006: 132, adapted)



Hagberg also shows that autosegmental stress can be used to account for affixes that cause a stress shift. Stress in Tagalog shows a similar pattern to Mayo, except that stress is determined with respect to the right edge of the word rather than the left edge. Thus, stress appears on either the penultimate or the final syllable:<sup>7</sup>

(31) *Tagalog stress* (Hagberg 2006: 176, citing French 1988)

Penultimate stress:	Final stress:
pú:noʔ      ‘trunk of a tree’	punóʔ      ‘full’
bá:soh      ‘drinking glass’	basóh      ‘target practice’
bú:kas      ‘tomorrow’	bukás      ‘open’
tá:sah      ‘cup’	tasáh      ‘having the point sharpened’
ará:lan      ‘place for studying’	aralán      ‘apprentice’
kaʔibí:gan      ‘friend’	kaʔibigán      ‘desire, inclination, preference’

Also similarly to Mayo, the addition of most affixes (here suffixes) preserves the distance of stress from the word edge:

(32) *More Tagalog stress* (Hagberg 2006: 176, citing French 1988)

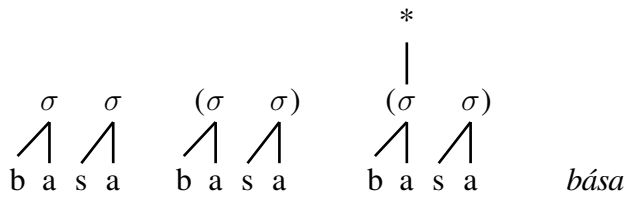
ʔá:ral      ‘study’	ará:l-an      ‘place for studying’
bá:sa      ‘read’	basá:-hin      ‘to read’
pá:sok      ‘enter’	pasú:k-an      ‘(someone) places (something) inside of X’
dugóʔ      ‘blood’	dugu-án      ‘bloody’
ʔupóʔ      ‘sit’	ʔupoʔ-án      ‘(someone) sits on X’
bilíh      ‘to buy’	bilhín      ‘to buy X’

In penultimately stressed words, a foot is built at the right edge of the word, followed by insertion of a \* at the left edge of that foot. This derives the default stress pattern:

<sup>7</sup>Hagberg indicates added vowel length in stressed open syllables, but notes that this length is predictable based on the location of stress. He thus omits it in his formal analysis.

(33) *Derivation of penultimate stress in Tagalog* (Hagberg 2006: 177)

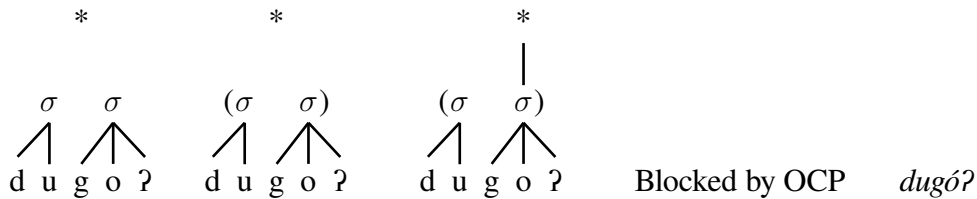
Input:      Build Foot:      Insert/Link \*:      Output:



In final-stressed words, a floating \* is present. Hagberg assumes that the direction of association of \* autosegments is parametrized, and that association applies from right to left in Tagalog. Following the building of a foot, association of \* occurs, creating final stress. Default insertion of \* at the left edge of the foot is blocked from applying due to the OCP:

(34) *Derivation of penultimate stress in Tagalog* (Hagberg 2006: 178)

Input:      Build Foot:      Link Floating \*:      Insert/Link \*:      Output:



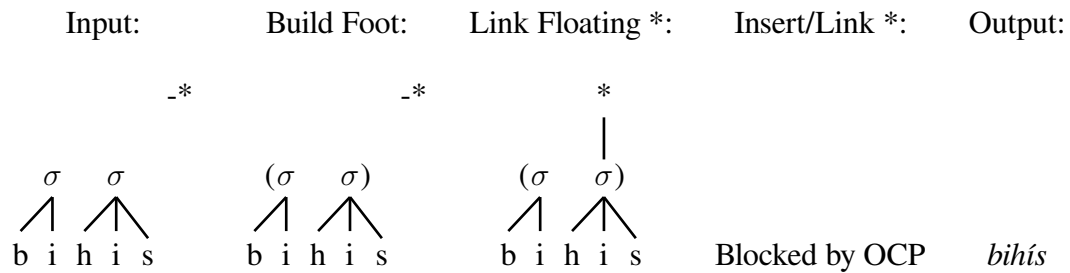
In Hagberg's model, it is possible for a morpheme to consist solely of a floating \* autosegment, and this is exactly what he claims for the Tagalog suffix that turns nouns into adjectives. On the surface, this involves shifting stress to the final syllable:

(35) *Forming adjectives by shifting stress* (Hagberg 2006: 179)

bí:his	‘way of dressing’	bihís	‘dressed up’
gú:tom	‘hunger’	gutóm	‘hungry’
bú:tas	‘hole’	butás	‘punctured’
bí:lang	‘number’	biláng	‘counted’
bá:li?	‘fracture’	balí?	‘fractured’

When a \*-only suffix is attached, linking again occurs, blocking stress insertion. The only difference between (36) and (34) above is that the underlying floating \* comes from the suffix, not from the root:<sup>8</sup>

<sup>8</sup>For clarity, I have made several changes here from the derivation given by Hagberg. First, he shows suffixation

(36) *Floating \* suffixation* (Hagberg 2006: 179, adapted)

Hagberg claims that if a final-stressed stem (i.e., one with a floating \*) is combined with the \* adjectivizer suffix, only one of the stress autosegments will be realized due to the OCP-driven \* deletion rule in (37). However, he points out that none of his sources give examples of such roots, so this claim remains untested.

(37) *Delete \** (Hagberg 2006: 179)

$$* \rightarrow \emptyset / * \_$$

Hagberg's proposal for stress as an autosegment, while insightful, is not unproblematic in the eye of the present study. First and foremost, it does not draw a distinction between contrastive stress (accent) and non-contrastive (rhythmic) stress. At least in part, this shortcoming leads Hagberg to retain feet as a structural property of representations used for identifying some non-contrastive stresses, rather than as a consequence of a default accent assignment algorithm. Furthermore, his model requires positing an additional grammatical primitive, the autosegment \*, which is used only in stress systems. This is decidedly not in the spirit of developing a unified model in which all prosodic systems are captured with a common set of representational elements. These problems aside, Hagberg's attempt to analyze stress systems using the same autosegmental principles proposed for tone systems is a welcome interpretation of the facts. Furthermore, though he still requires foot structure, he does reduce the complexity of the metrical module of the grammar by eliminating the notion of headedness.

---

as a separate step, which I have omitted for space reasons. Second, he shows foot building as applying after the rule that links floating accent, with a monosyllabic degenerate foot being built around the linked accent. Hagberg states that the relative ordering of these processes is inconsequential, and while he does not give a derivation where footing comes first, I choose to do so in order to retain consistency of presentation.

Dubina (2012) arrives at an autosegmental analysis of stress phenomena independently of Hagberg. Dubina looks at free stress (i.e., lexical accent) systems, observing that many phenomena concerning the behaviour of stress in these languages can be captured more elegantly and with fewer stipulations than are required with foot-based representations. Since Dubina is concerned with non-predictable stress, he takes a contrast-centric approach where Hagberg does not. Even more importantly, given the present goal, Dubina seeks unified representations for tone and stress systems; he proposes that an existing tonal autosegment (H tone) is used to represent stress accent, rather than positing a separate element (such as Hagberg's \*).

Dubina's proposal is based on the more liberal phonetic interpretation of phonological tone espoused by Lockwood (1983). Lockwood argues that typology of prosodic systems should not be determined based on phonetic qualities. For example, pitch is not the only way that a language can have a "tone" system. He suggests instead that phonological tone may be realized with a number of phonetic qualities beyond properties of pitch, including duration and intensity, both of which have traditionally been associated with "stress" systems. Dubina points out that the idea of accentual autosegments with varying phonetic realizations was also put forth by Revithiadou (1999), who posited two autosegments, "strong accent" and "weak accent". In metrical stress accent systems, these correspond to the strong and weak branches of a foot respectively, while in pitch accent systems, they are realized as high and low pitch. Revithiadou also uses the traditional H and L autosegments in discussing tone systems, so like Hagberg (1993, 2006), she does not provide a model that unifies the representation of stress and tone systems. They are in fact redundant, Dubina notes, calling the strong and weak accents a "multiplication of entities" and suggesting that they are "simply high and low tones in disguise" (2012: 113).

The phenomenon of contrastive stress, and the accent that underlies it, calls out for some kind of autosegmental analysis. Autosegmental stress would allow for contrastive stress languages to be understood in the same terms as tone languages, making for an overall less disparate conception of prosodic systems.

## 2.3 Length

The term “length” refers broadly to the duration of linguistic units. While many factors (including surface stress and tone) can affect duration, duration itself may be used as a phonological property. A language with a length or “quantity” contrast is one in which the durations of certain segments are phonologically significant, much like a tone language is one in which pitch is phonologically significant.

Consider Finnish, which Suomi et al. (2008: 29) describe as a “full-fledged quantity language”. In Finnish, a short–long contrast is available for vowels in all positions and for consonants intervocalically and in some consonant clusters. The contrast is extremely robust, and can arise from both the underlying forms of monomorphemic words and from the addition of inflectional morphology. Thus minimal sets like (38) abound in the language:<sup>9</sup>

(38) *Finnish length contrasts*

muta	‘mud (nominative)’
mutaa	‘mud (partitive)’
mutta	‘but’
muuta	‘other (partitive)’
muutta	‘other (abessive)’
muuttaa	‘to change’

While the same phonological contrasts are available for most varieties of Finnish, there is a great deal of variation in the phonetics. For example, the durations of segments with the same phonological length are known to differ depending on factors such as dialect and position within the word or phonological phrase (Wiik and Lehiste 1968; Lehtonen 1970; Ylitalo 2009).

Phonetic duration is in some ways analogous to fully predictable stress in that it has no consequences in the phonology and cannot be used contrastively. Rather, it is merely the phonetic substance of how a finite number of categories are produced in certain contexts. In the study of phonological contrast, however, our questions concern the representational nature of length—how does it fit in with the understanding of prosodic contrasts?

---

<sup>9</sup>In (38), as in the orthography, short segments are written with a single letter and long segments are written with doubled letters.

### 2.3.1 Representing length

Most proposals for the representation of phonological quantity are based on the observation that, like the other prosodic features (tone and stress), length is different from segmental features. Traditionally, long segments were treated in one of two ways: In a system of binary distinctive features, length can be analyzed as a contrast of a segmental feature  $[\pm\text{long}]$ , as in (39a). Alternatively, long segments could be represented very literally as doubled versions of their short counterparts, as in (39b).

(39) *Pre-autosegmental representations of contrastive length*

	Long	Short
a. Featural:	/a:/	/a/
	[+long]	[-long]
b. Doubled:	/aa/	/a/

McCarthy (1979, 1981) presented an alternative based on autosegmental principles as a way of accounting for the so-called “nonconcatenative” morphology of languages like Arabic. The kind of pattern in question, exemplified in (40), is quite pervasive in many Semitic languages; McCarthy (1981: 374) points out that this list of forms is “hardly exhaustive”.

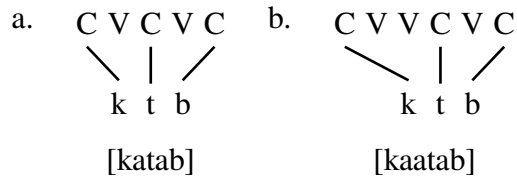
(40) *Arabic triconsonantal paradigm* (McCarthy 1981: 374)

a.	kataba	‘he wrote’
b.	kattaba	‘he caused to write’
c.	kaataba	‘he corresponded’
d.	takaatabuu	‘they kept up a correspondence’
e.	ktataba	‘he wrote, copied’
f.	kitaabun	‘book (nom.)’
g.	kuttaabun	‘Koran school (nom.)’
h.	kitaabatun	‘act of writing (nom.)’
i.	maktabun	‘office (nom.)’

Although some of the forms in (40) contain prefixes and/or suffixes, they share a common root consisting of only the consonants /ktb/. Depending on the morphological form, different consonant and vowel “melodies” can combine with different “prosodic templates” (i.e., the sequences in which the consonants and vowels appear). McCarthy proposes that the different levels

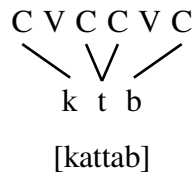
(vowels, prosodic templates, and consonants) constitute separate autosegmental tiers. In forms with all short consonants, associations between consonants and C-slots are one to one:<sup>10</sup>

(41) *Consonant association to prosodic templates* (McCarthy 1981: 288, adapted)



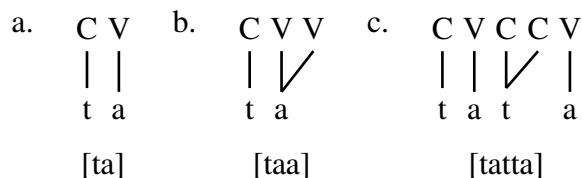
When the prosodic template contains more C-slots than there are consonants to associate with, consonants may spread across multiple C-slots. If this results in two-to-one association of C-slots to consonants as in (42), the consonant is geminated. However, since length is represented on a separate tier from the consonant itself, a long consonant is still considered to be a single segment.

(42) *Two-to-one association* (McCarthy 1981: 288, adapted)



While this approach to the CV “skeleton” was originally developed to deal with nonconcatenative morphology of the type seen in Semitic languages, the skeleton has been used to represent quantity even in languages with purely concatenative morphology. The standard assumption is that short segments are associated with a single CV-slot and long segments are associated with two CV-slots:

(43) *Representation of length with a CV tier*

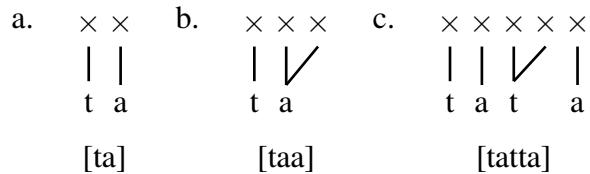


<sup>10</sup>I have simplified the representations here somewhat from those given by McCarthy (1981), leaving out the marking of morphological structure.



For McCarthy (1979, 1981), C-slots and V-slots were nodes corresponding to the feature-geometric specifications of [–syllabic] and [+syllabic], respectively. In order to refer to phonological environments containing either C *or* V (i.e., the presence of *any* segment), his model needed a feature [+seg]. In later work on skeletal representations (Kaye and Lowenstamm 1984; Levin 1985), the structure of the skeletal tier was generalized to “X-slots” representing all segments. The one-to-one and two-to-one mappings for short and long segments in such a system are analogous to those used with CV-slots, except consonants and vowels do not link to distinct elements:

(44) *Representation of length with a generalized X tier*



An important property of models based on a skeletal tier is that all realized segments are associated with some skeletal slots (one or two, depending on length). However, there is also the possibility that a segmental melody is not associated with a skeletal slot, like *z* in (45). McCarthy’s (1981) analysis allows for such “floating” melodies, and in fact blocks many-to-one association of melodies to skeletal slots by default. In other words, *z* will not generally link to *C* along with *y*; language-specific rules are required in order for such associations to occur.

(45) *Non-realization of unassociated melodies* (McCarthy 1981: 382)



This principle has been exploited to account for systems with apparent subtractive morphology. One well-known case is that of French “latent consonants”:

(46) *Fixed and latent consonants in French* (Tranel 1995: 131, adapted)

- a. honnê[t]e    un honnê[t]e citoyen  
       honnê[t]eté    un honnê[t]e homme    un honnê[t]e personne
- b. peti[∅]        un peti[∅] citoyen  
       peti[t]esse    un peti[t] homme        une peti[t]e personne

In (46a), the word-final /t/ in *honnête* ‘honest’ is what Tranel (1995) calls a “fixed” consonant; it is pronounced in all contexts. The word-final /t/ in *petit* ‘small (masc.)’ is only pronounced in certain contexts; it is a “latent” consonant. Tranel (1995) reviews a number of different approaches to account for this pattern, noting that early debate concerned whether latent consonants were inserted or deleted. However, he concludes that the preferable approach is one in which latent consonants are underlyingly floating consonant melodies which associate with empty X-slots in the relevant contexts. In word-final positions, a latent consonant does not normally have its own X-slot. Since the /t/ remains floating, it is not realized:

(47) *Latent consonant not realized*

×	×	×	×
p	œ	t	i
[pœti]			

In a context in which the latent consonant is realized, there must an X-slot present. For the realization of latent consonants across a word boundary before a vowel-initial word (a phenomenon known as *liaison*), this X-slot comes from some source; depending on the particular analysis, this X-slot can be present underlyingly on the second word or inserted by rule. At any rate, when the two words come together in a phrase, the floating consonant melody from the first word associates with the empty X-slot of the second word, causing the consonant to be realized:

(48) *Latent consonant realized*

×	×	×	×
p	œ	t	i
[pœti]			
			/
			ɔ
[pœtit ɔm]			

Although Tranel (1995) does not address it specifically, the final latent-but-realized consonant in feminine adjectives such as *petite* can be accounted for within the same model by positing a feminine suffix consisting of only an empty X-slot.<sup>11</sup> Thus for the feminine form in (49b), the X-slot of the suffix allows the final /t/ to be realized by giving it a place to link to, while the latent /t/ of the unsuffixed masculine form in (49a) (as in 47) is not realized.

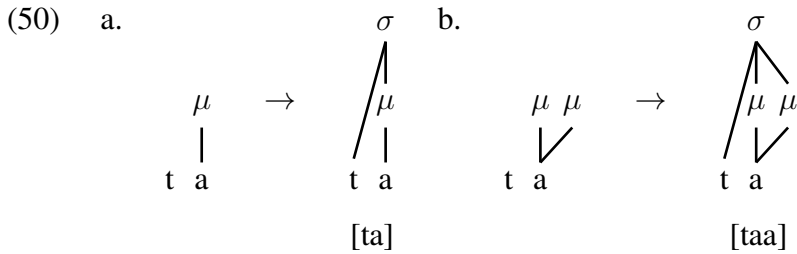
(49) a.	$\times \times \times \times$ $\mid \mid \mid \mid$ p œ t i t [pœti] ‘small (masc.)’	b.	$\times \times \times \times -\times$ $\mid \mid \mid \mid \vdots$ p œ t i t [pœtit] ‘small (fem.)’
---------	--	----	---

Analyses in this spirit have been used to account for data which seem to involve morphological deletion processes (i.e., deletion of the final segment in certain forms) in a number of languages. Examples include Southern Sierra Miwok (Sloan 1991), Estonian (Spahr 2013, further developed in Chapter 5 of this thesis), and Yine (Zimmermann 2013). Bye and Svenonius (2012) also generally advocate for the use of unspecified prosodic material and floating segmental melodies to account for apparent subtractive processes, among other phenomena.

An alternative but related approach developed in parallel to CV-slot/X-slot representations of quantity contrasts, known as moraic theory (Hyman 1985; McCarthy and Prince 1986; Hayes 1989). In this model, representations of segmental and syllabic quantity are encoded with a unit known as the “mora”. Vowel quantity is represented analogously to the one-to-one and two-to-one mapping used in skeletal models. That is, a short vowel receives one mora while a long vowel receives two moras. Moras are normally only assigned to segments in the syllable rhyme; short consonants, when syllabified as onsets, do not usually receive moras:<sup>12</sup>

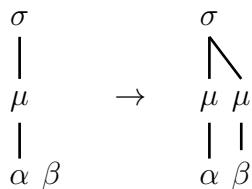
<sup>11</sup>There may of course be other complications to deal with in such an analysis, such as accounting for dialects in which the feminine suffix is accompanied by a schwa.

<sup>12</sup>The weightless status of onsets is not uncontroversial, however; cf. Topintzi (2010).

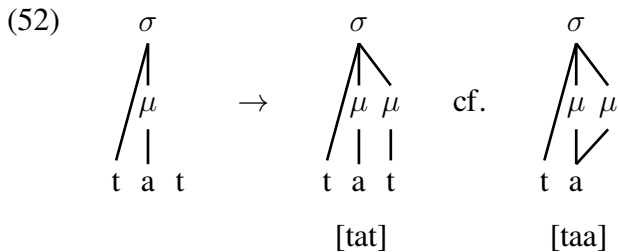


Moraic theory is based on assumptions that are integrated with those of metrical theory, in that syllables and stress assignment play a central role. When heavy syllables attract stress (in languages with quantity sensitivity), the fact that syllables with long vowels bear two moras makes them heavy. In languages where closed syllables also attract stress, moras are assigned to coda consonants by a rule called “Weight by Position”:

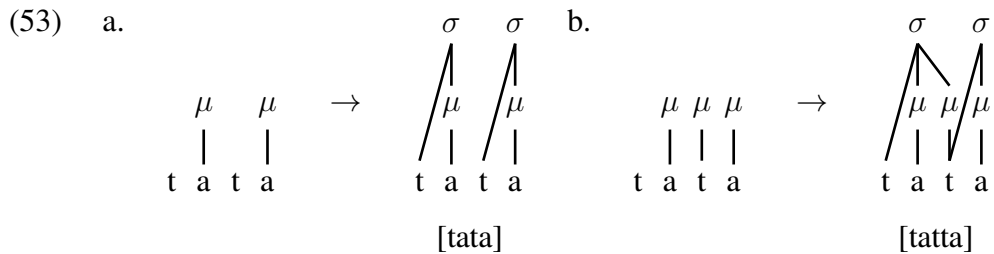
(51) *Weight by Position* (Hayes 1989: 258)



CVC syllables are consequently moraicly equivalent to CVV syllables:



Since codas are understood to be weight-bearing consonants, underlying moras can be exploited to represent geminates, as seen in (53b). The presence of a mora causes the consonant to contribute weight to the first syllable, but it is simultaneously “flopped” into the second syllable, where it serves as a weightless onset. Compare this to the word in (53a), where the medial consonant serves only as a weightless onset in the second syllable; at no point in the derivation does it bear a mora, so it is never adjoined to the first syllable.



The exact nature of the mora is a subject of some debate in the literature. Hayes (1989) conceived of the mora as a unit of weight that could be specified underlyingly in association with segments or derived (via Weight by Position) in order to calculate a property of syllables under the assumption that segmental quantity could always piggyback on weight. However, in some languages, syllabic weight and segmental quantity are difficult to view as independent properties, while in others, they seem to be independent properties, especially where geminates do not necessarily contribute to syllable weight (Tranel 1991; Broselow 1995; Odden 1997; Kraehenmann 2001). For systems like this, it is not entirely clear how long segments should be represented and how many different configurations of segment-to-mora and segment-to-syllable linking should be allowed.

The mora has also been described as a timing unit. Broselow et al. (1997) propose that cross-linguistic variation in the linking of moras to segments is phonetically meaningful on the surface: splitting a mora between two segments will result in those segments having shorter durations than segments which bear their own moras. Based on similar assumptions, Spahr (2011) suggests that certain dialects of Finnish may regularly insert a “half mora” which has consequences for both phonetic duration and phonological gemination processes.

For the most part, the differences between CV-slot/X-slot and moraic theories of length are notational—both share the insight that length is represented as an additive property of some element of quantity existing on some suprasegmental tier. The differences between the theories largely concern the identification of surface stress as a property of syllable weight, which is not independent from the representation of quantity. As such, syllable weight as a distinct lexical unit has no place in a model of quantity that is concerned purely with contrastive representations.

## 2.4 Summary

In this chapter, I have given a review of the literature on word prosodic representations and highlighted some of the key concepts in thinking about prosody, including especially the notions of autosegmental tiers and autosegment-internal feature organization. Given the goals of the present study, I have remained critical of the extent to which aspects of various proposals reflect lexical contrasts rather than serve simply as notations for generalizations about surface phonetics.

In the next chapter, I will consider the characteristics of existing models of prosodic features that are compatible with the contrastivist approach and lay out the principles of the unified model of prosodic representations that I will defend throughout the rest of this thesis.

# Chapter 3

## Towards a minimal model of prosody

This chapter lays out the minimalist model of word-level prosodic contrasts adopted in the present study. In Section 3.1, I argue for a reductionist model of prosodic representations that draws on the existing approaches in the literature. This model relies on two tiers of root nodes: one of X-slots, which bear segmental contrastive features and represent length contrasts, and one of prosodic root nodes, or “ $\pi$ -nodes”, which bear the contrastive features of prosodic autosegments. In Section 3.2, I explore some theoretical issues that come up in this approach and set some restrictions on possible representations.

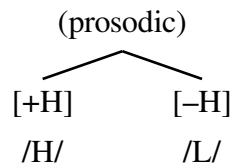
### 3.1 The minimal elements of prosodic contrasts

#### 3.1.1 Contrast and the tonal node

Let us first address contrast as it applies to the structure of tone. I take the contrastivist point of view, where phonological specification is determined from the number of oppositions present in a system. In other words, contrastive representations are based on the size of the inventory, where contrastive elements are used to divide the inventory into its constituent parts. I adopt the notion of the contrastive feature hierarchy (Dresher 2009) as an explicit model of inventory division in conjunction with standard assumptions (due to Goldsmith 1976) about tonal autosegments,

specifically, that they constitute separate groupings of features from the segmental inventory, with association lines representing connections between the two tiers.<sup>1</sup> By this I mean that the inventory of tonal autosegments receives its contrastive specifications in a feature hierarchy isolated from that which divides the segmental inventory. As with segmental hierarchies, the number of features depends on the size of the inventory. For a two-tone system, in which /H/ and /L/ both behave as separate units in the phonological inventory, this means that a single contrastive feature can be used, which we may call [ $\pm$ H(ighpitch)]:

(1) Inventory: /H, L/



The tonal contrastive features in (1) have the same properties as the segmental features assumed in other studies on contrast. First, they are language-specific, not universal (Dresher 2014, see also Mielke 2008). That is, languages do not have a universal set of features, but rather the universal ability to divide inventories using binary contrasts. Second, a feature defines the realization on one side or the other of some boundary along a single phonetic dimension; non-contrastive aspects of phonetic realization are filled in later by enhancement (Hall 2011, and references therein).<sup>2</sup> Thus, the feature [ $\pm$ H] in (1) specifies whether an autosegment attributes pitch above or below the cutoff separating the two categories of a binary tone opposition. When two (auto)segments are in a position of full contrast, their contrast is defined by the positive and negative values of a feature, and their phonetic realizations must stay distinct along the phonetic dimension defined by that feature. Note that this two-member tonal inventory could just as easily be divided by a feature [ $\pm$ L(owpitch)], with positive and negative values reversed, as in (2). What

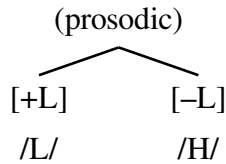
<sup>1</sup>Krekoski (2013) has used contrastive hierarchical feature specification to analyze tone categories in Chinese languages, but he does not explicitly address the autosegmental nature of tone.

<sup>2</sup>Godfrey (2012) proposes that binary tonal contrastive features may be slightly more abstract in their phonetics, for example allowing a feature “[ $\pm$ B]” where [+B] is a property of the phonetically highest tone and [-B] is a property of the phonetically lowest tone. Odden (2011) proposes a feature [extreme] with similar properties. I do not rule out features of this type here.

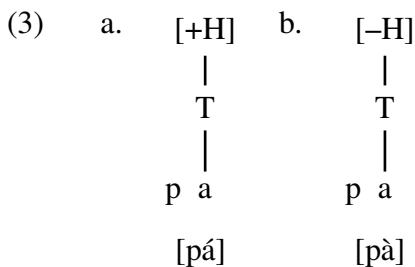


is important is that a single feature is possible in a two-member inventory—the substantive name of that feature is not important.

(2) Inventory: /H, L/

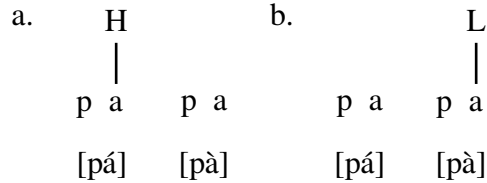


In line with common assumptions in the literature on tonal feature geometries, contrastive autosegmental features are linked to a tonal root node (notated as “T” in 3), which itself is connected to segments via association lines:



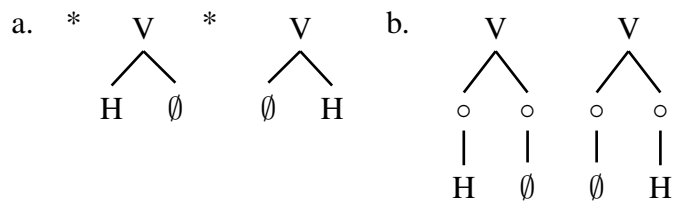
The assignment of contrastive features is only possible for an inventory that contains two or more members; if an inventory does not contain at least two categories, then it cannot be divided into more than one part. While there is no language with a *segmental* inventory consisting of only one member, many tone languages have autosegmental inventories consisting of only one member. These are known in the literature as “privative tone languages” (Hyman 2001).

In privative tone languages, only one of the tone values of a surface two-tone system is represented formally. The surface contrast may be between phonetic high and low tone, but underlyingly, one of the two is null. Thus a privative tone language will either be /H,  $\emptyset$ / (as in 4a) or /L,  $\emptyset$ / (as in 4b). In other words, the marked member of the tonal opposition can be spelled out as either high or low pitch; in this case markedness is defined as the presence of an autosegment, not the phonetic content of that autosegment.

(4) *Representation of privative tone* (as in Hyman 2001)

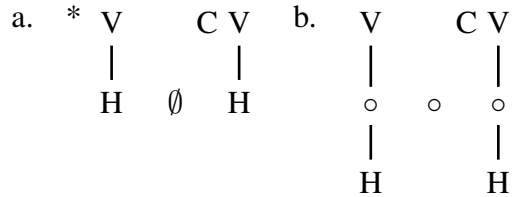
Evidence for privative tone systems comes from phonological activity, and crucially not from surface phonetics. Privative tone systems are defined largely by what they cannot do with respect to the unmarked member. Hyman (2001) uses privative high tone (/H/ versus /∅/) systems to exemplify this. First, composed HL contours are not available in privative high tone languages, because the linking of both a tonal autosegment and the lack of an autosegment to a single tone bearing unit as in (5a) fails to specify two levels of pitch; these could only be realized as high. Hyman admits that more sophisticated representations using tonal root nodes as in (5b) would allow a contour between specified /H/ and an underspecified root node, but he prefers to argue against the existence of empty tonal root nodes altogether.

## (5) If /H/ versus /∅/, then we should not get HL or LH contour tones (Hyman 2001: 239, adapted)



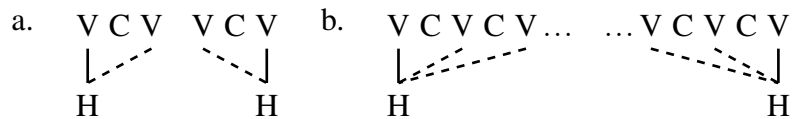
Privative tone languages are also restricted in terms of how they may employ floating tones. While in a binary tone language, either /H/ or /L/ autosegments can be specified in the underlying forms of morphemes, Hyman points out that it is not clear what this would mean for the unmarked member of a privative system; what is meant by a floating null, as in (6a)? He also points out that this could be accommodated with underspecified tonal root nodes, and asserts that the possibility in (6b) should be ruled out.

(6) If /H/ versus /∅/, then we should not get floating L tones (Hyman 2001: 240, adapted)



Finally, privative tone languages may not refer to the null member of the opposition in phonological processes such as assignment, shifting, and spreading. Whereas a binary tone system can allow e.g., spreading of both /H/ and /L/, a privative /H/ language will only allow spreading of /H/. Thus a great number of tone spreading rules are allowed for /H/, both leftward (perseverative) and rightward (anticipatory), either by one tone bearing unit (bounded) or for continuing iteratively (unbounded), as in (7), but corresponding rules spreading the null tone are not possible to any extent.

(7) *Spreading (perseverative/anticipatory)* (Hyman 2001: 241. adapted)



How do privative tone systems work if we adopt the principle of inventory-based contrastive feature specification? If the division of contrasts within the autosegmental inventory builds a separate contrastive hierarchy from that of the segmental inventory, then this means that autosegments in a privative tone language do not bear any contrastive features, since there is no inventory to divide:

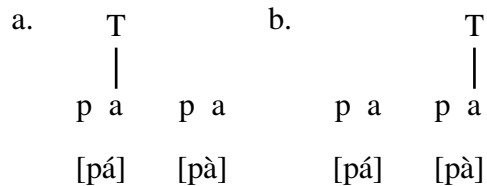
(8) Inventory: /H/ or /L/

(no prosodic feature hierarchy possible)

This does not mean that privative tone oppositions are impossible, however. If we assume that tonal autosegments obligatorily and minimally consist of a tonal root node, then the tonal root node itself can be used to represent privative tone, in the absence of tone features. The presence of a tonal root node therefore reflects formal markedness, not the surface phonetics.

Thus in a privative high tone language, the phonetically high tone is marked with a root node as in (9a), while in a privative low tone language, the phonetically low tone is marked with a root node, as in (9b). Since the single-member inventory does not allow the specification of tonal features, empty tonal root nodes like those in (9) cannot end up in opposition to tonal root nodes bearing features. This means that, in privative tone systems, representations with contours and floating tones like in (5b) and (6b) above are ruled out not by banning unspecified root nodes, as Hyman suggests, but rather by banning *specified* root nodes.

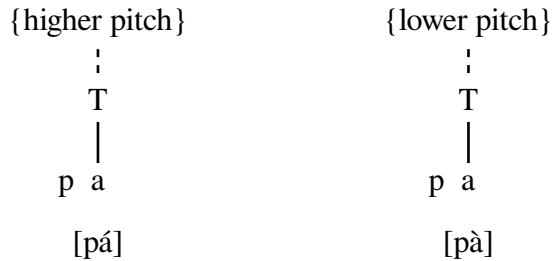
(9) *Privative tone without features*



The difference between a privative high and privative low language, then, lies in the language-specific phonetics, not in the formal phonology; privative tones always have the same lexical representation, that is, the presence of a featureless tonal root node. Thus, in the present proposal, root nodes are similar to features in some ways when it comes to the relationship with phonetics. Features are binary valued elements whose specifications supply two sides of a contrast along some phonetic dimension. For example, [+H] supplies high pitch while [-H] supplies low pitch. The removal of a given feature causes a lack of specification of the phonetics supplied by that feature. When all of the features in an inventory are stripped away leaving only the root node, a single privative element is revealed in an otherwise binary system. However, this privative element, the root node itself, still has a language-specific realization that supplies phonetics along some dimension. In a privative high tone language, the root node supplies higher pitch, as in (10a), while it supplies low pitch in a privative low tone system, as in (10b):

(10) *Assignment of tonal root node phonetics*

- a. “Privative H” language    b. “Privative L” language



This analysis is not without precedent: Rice (2000) draws similar conclusions about the representation of privative tone in Athapaskan languages. The marked member of privative tone contrasts in modern Athapaskan languages developed historically from a glottal stop or glottalized consonant. However, the phonetic realization of the privative tone differs depending on the language. In Slave, glottals developed into high tone, while the corresponding cognates in Sekani have phonetically low tone. This is schematized in (11).

(11) *Surface tones in Athapaskan*

Proto-Athapaskan	Slave	Sekani	
*ʔ	H	L	←Marked tone
*∅	L	H	

Rice proposes that the marked tones in both languages have the same formal representation, despite having different surface phonetics. Representing this markedness using minimal contrastive feature specifications, she concludes that the marked tone consists only of a tone node bearing no tonal features, while the unmarked tone contains no tone node at all. The opposition is represented as in (12). The differences in the surface phonetics come about as a result of language-specific phonetic enhancement; they are not represented formally.

## (12) a. Marked tone    b. Unmarked tone



The use of a bare tonal root node to represent privative tone contrasts follows directly from the assumption that feature specifications are based on inventory size, and if bare root nodes, as privative elements, are assumed to have their own language-specific phonetic realizations, then diachronic shifts from “privative high” to “privative low” systems are expected. These are observed in Athapaskan (as seen above), and a similar shift is known to have occurred in Bantu, changing some privative high systems to privative low systems (Hyman 2001).

The upshot is that prosodic systems with a single-member prosodic inventory use representations based on marking with a bare prosodic root node as the only autosegment; in analyzing such systems, we need only be concerned with the formal properties of this autosegment, not its phonetic realization.

### 3.1.2 A unified prosodic autosegment

It is clear from studies by Hagberg (1993, 2006), Revithiadou (1999), and Dubina (2012), as well as earlier work by Halle and Vergnaud (1987), that it is advantageous to have an autosegmental understanding of stress accent. With the goal of developing a unified model of word-level prosodic contrasts, however, the key theoretical question is how these autosegments relate to each other in our understanding of the phonological module as a whole. It is true, as Dubina (2012) points out, that proposals that posit a separate class of autosegments (like those by Hagberg and Revithiadou) involve multiplication of entities, at least for systems where accent is realized with pitch. Dubina solves this by proposing that there are no separate autosegments that differentiate free stress from tone; both are formally high tones.

This unified formal representation for tone and stress also predicts what Dubina calls the “Incompatibility of Freedoms Hypothesis”, that is, “that free stress is incompatible with unpredictable tone” (Dubina 2012: 13). In other words, the location of stress and the location of tone cannot covary in order to encode contrasts independently from one another in the same language.<sup>3</sup>

---

<sup>3</sup>Dubina (2012) does note that some languages (e.g., Papiamentu and Ma’ya) are described as falsifying the stronger view of this hypothesis, and he struggles to capture these systems within his model. I will address Papiamentu in Chapter 5.

This is highly desirable in the contrastivist research program, which aims to understand the limits of the distribution of contrasts within a formally restrictive model by considering different constraints on the co-occurrence of contrastive specifications.

On closer inspection, though, Dubina's tone-centric representations fall short on satisfying the requirements of a contrast-based analysis. One problem is that it is implicitly couched in phonetic substance, despite being highly abstract. While the prosodic autosegment employed allows a great deal of variation in terms of its phonetic realization (including realizations that do not reference pitch at all), Dubina still insists that it is formally *tone*. If the primitive available to grammars has such a wide range of interpretations, then there is no reason to consider it inherently *tonal* at all; the autosegment's name should be as abstract as its behaviour warrants.

The other problem concerns properties of the formal representation of this autosegment: it is considered to be a high tone (H), not just a tone. If a language has a single-member autosegmental inventory (i.e., one autosegment, which represents stress), then under the present assumptions, no contrastive feature specifications should be available within that inventory because there are no contrasts to divide. Recalling the discussion of the contrastive representation of tone above, this is not a problem, since in the absence of contrastive features, the root node itself is available in contrastive representations. This was seen for privative tone systems, where privative H-tone and privative L-tone both used a bare tonal root node for the marked member of the opposition. Privative systems were thus claimed to be formally identical, differing only in the language-specific realization of their tonal root nodes. Dubina's use of autosegmental stress bears striking similarity to these claims, both in terms of being formally tonal, and in terms of tones having language-specific realizations.

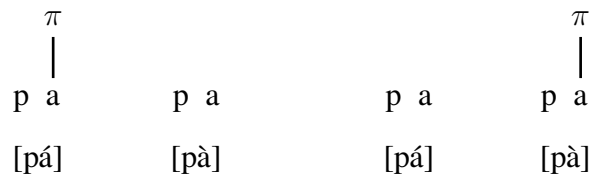
Combining Dubina's analysis with the insight into contrastive feature specifications gleaned from examining privative tone systems, it is possible to achieve a more tightly integrated formal unification of prosodic representations. If the tonal autosegments used to mark stressed positions do not bear contrastive features, then there is nothing inherently tonal about them at all. Since autosegments in privative tone systems do not bear contrastive features either, the same can be

said for them. I thus propose that phonological systems operate on a single formal autosegment: the prosodic root node, or “ $\pi$ -node”.

Just like the featureless tonal root nodes discussed above and the tonal autosegments assumed by Dubina (2012),  $\pi$ -nodes have language-specific phonetic realizations; their cross-linguistic similarities are only formal in nature. In the present model, a  $\pi$ -node will appear any time prosodic autosegments are present in the representation. Thus a “privative H” language is a featureless  $\pi$ -node language where  $\pi$ -nodes are realized with higher pitch, as in (13a), and a “privative L” language is a featureless  $\pi$ -node language where  $\pi$ -nodes are realized with lower pitch, as in (13b):

(13) *Privative tone with featureless  $\pi$ -nodes*

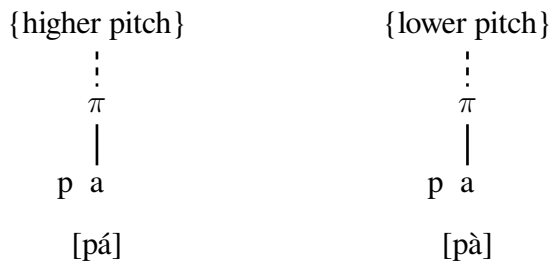
- a. “Privative H” language    b. “Privative L” language



The assignment of these language-specific realizations can be schematized as follows:

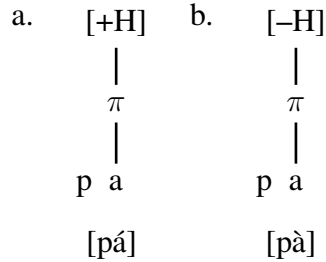
(14) *Assignment of  $\pi$ -node phonetics*

- a. “Privative H” language    b. “Privative L” language

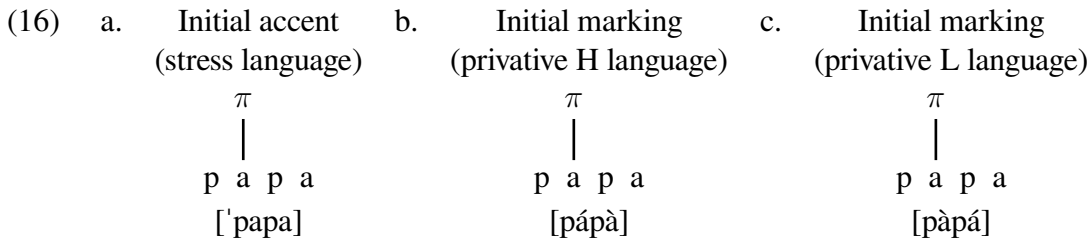


The  $\pi$ -node serves the same formal function as the tonal root node, in that it is the point to which contrastive features dividing the prosodic inventory link. In a binary tone system, for example, a single contrastive feature would be used to specify the two tone heights, as in (15). Here both sides of the phonetic dimension of pitch height are contrastively specified by the binary feature, such that the  $\pi$ -node bearing [+H] assigns higher phonetic pitch while that bearing [–H] assigns lower phonetic pitch.

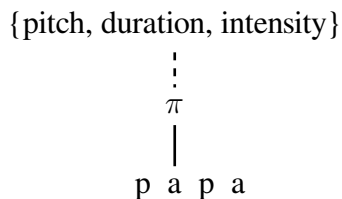


(15) *Binary tone with feature-bearing  $\pi$ -nodes*

Featureless  $\pi$ -nodes can also be used to mark accented positions in free stress languages in a way analogous to the \* autosegment proposed by Hagberg (1993, 2006) and the H tone proposed by Dubina (2012). There is no formal representational distinction between  $\pi$ -nodes in different kinds of systems, meaning that stress-accentual  $\pi$ -nodes can be operated on by the same autosegmental principles as tonal  $\pi$ -nodes. Consequently, accented positions in stress languages look identical to marked positions in privative tone languages:



In comparison to privative H and privative L languages, then, free stress languages are characterized by having  $\pi$ -nodes with yet another language-specific phonetic realization, which is often one or more of increased pitch, duration, and intensity:

(17) *Non-formal phonetics of stress*

An advantage of using formally identical representations for accent and privative tone is that it eliminates the need to arbitrarily choose between competing analyses. Hyman (2006), looking at the typology of word prosody, proposes that the most important typological distinction to be

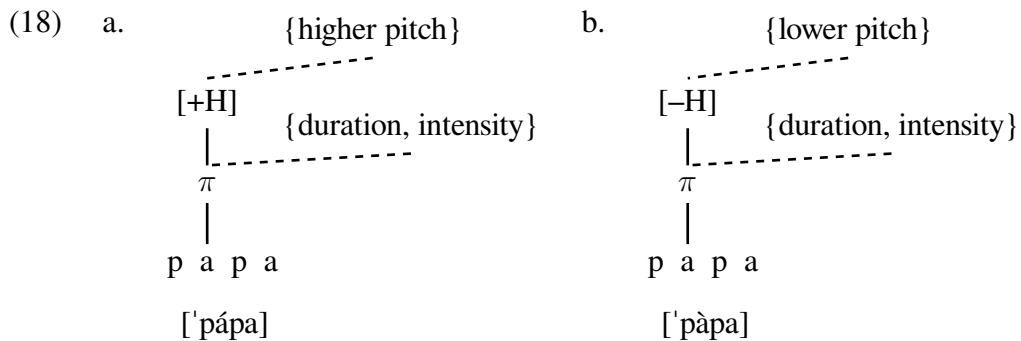
made should distinguish languages with “stress accent”, which require one and only one syllable per word to be metrically marked with the highest level of prominence, from languages without this requirement. Prototypical stress accent is thus said to be “culminative” and “obligatory”, with other prosodic systems being broadly considered tonal. However, referring also to Gussenhoven (2006), Hyman notes that there exist “pivotal” systems, which appear to be privative tone languages because prominence is marked with higher pitch, but in which this tonal marking is culminative/obligatory; pivotal systems can be analyzed equally well in metrical or tonal terms. In the present model, this analytical ambiguity does not exist; pivotal systems are simply privative  $\pi$ -node languages where the phonetic realization of a  $\pi$ -node is defined in terms of pitch and independent restrictions ensure that one and only one  $\pi$ -node may appear in each word. Metrical considerations are inconsequential to these representations.

With regards to the  $\pi$ -node model of autosegmental representations, I have discussed two kinds of systems: those in which phonetic realization is a language-specific property of the  $\pi$ -node (stress accent systems and privative tone systems) and those in which phonetic realization is specified by contrastive features linked to the  $\pi$ -node (binary tone systems). However, a third type of system is predicted in which  $\pi$ -nodes have a language-specific phonetic realization but still host contrastive features, which themselves specify separate phonetic properties. I argue that such systems are widely attested in a diverse set of families, most commonly being described as languages in which tone is only contrastive in accented positions.<sup>4</sup> For these systems, accent is represented by the placement of the  $\pi$ -node and spelled out with some phonetic realization (say, duration and intensity) while the contrastive feature specifications (such as  $[\pm H]$ ) are spelled out with separate phonetic realizations (some indication of pitch). It follows automatically that accented positions syntagmatically show the phonetic properties of the  $\pi$ -node and may paradigmatically show any of the phonetic properties of the contrastive feature specifications, but that unaccented positions may show none of these properties. Thus, tone is a separate but not independent contrast; formally

---

<sup>4</sup>Systems of this type are often been called “restricted tone systems” (by e.g., Hyman 2006) or by the sporadically-defined term “pitch accent”. Lockwood (1983) calls this phenomenon “tonal accent”, a class in which he includes Swedish, Punjabi, and North Pame. I will present data from from Goizueta Basque and Choguita Rarámuri to show how such systems are handled in the  $\pi$ -node model in Chapter 4.

speaking, it is dependent on accent in a very literal sense:



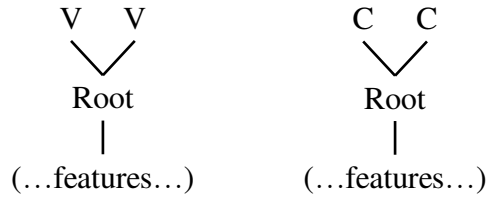
Having established the preliminaries of a unified account of the autosegmental representations of both tone and accent, we need only consider how prosodic autosegments interact with segments. This is a matter inherently tied in with phonological quantity, to which I turn next.

### 3.1.3 Length in a minimal model

Although proposals for the representation of length in the literature differ in their details and basic assumptions, most have in common several-to-one linking between two tiers: one encoding segments and one encoding their quantity. Since such a division of labour accounts for a number of phonological phenomena, it is important to incorporate this insight into the present model while emphasizing the goal of keeping representations based on contrast.

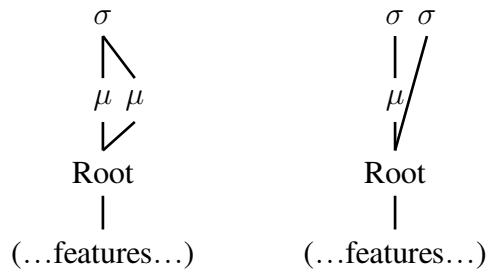
Selkirk (1990) considers the representation of length within the CV-slot/X-slot and moraic models in light of work on the feature-geometric internal structure of segments. For McCarthy (1979, 1981), the segmental melodies that linked to CV-slots in prosodic templates consisted of unordered feature bundles of the type employed by Chomsky and Halle (1968). If these feature bundles had hierarchical structure, then skeletal slots would be separate elements that link to the entire feature bundle via the root node. For long segments, this means a single feature-geometric tree associated with two skeletal positions:

- (19) a. Long vowel      b. Long consonant



Likewise, in moraic theory, length is represented through several-to-one linking to a single segmental root node. The difference from CV-slot or X-slot theory is that this linking occurs with higher levels of prosodic structure (moras and/or syllables):

- (20) a. Long vowel      b. Long consonant



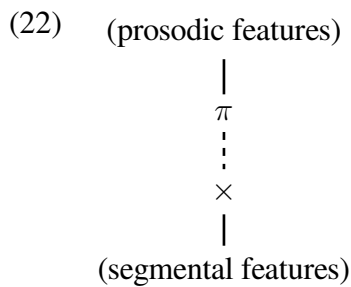
Selkirk (1990) proposes that quantity is represented not by separate timing units such as skeletal slots or moras, but by feature geometric root nodes themselves. A short segment consists of the feature structure linked to a segmental root node as usual, but in a long segment, the feature structure branches to be split across two root nodes:

- (21) a. Short segment      b. Long segment



I adopt the core of Selkirk's proposal here. Quantity is represented with one-to-one (for short) and several-to-one (for long) linking of segmental root nodes to contrastive segmental features. I refer to segmental root nodes as "X-slots" and notate them "×", but they should be understood to refer to segmental root nodes, not to an entity separate from the feature structure itself. Putting this together with the prosodic root node ( $\pi$ -node) proposed above, we arrive at the architecture

in (22). Word-level contrastive representations consist of two tiers of root nodes each bearing contrastive binary features specified within separate inventories; features on X-slots represent members of the segmental inventory, while features on  $\pi$ -nodes represent members of the autosegmental inventory.



In the model in (22), phenomena that have traditionally been described as “prosodic” are those that involve different configurations of linking between root nodes and other elements (either other root nodes or features). The two tiers of root nodes are formally parallel, and because they can both play roles in “prosodic” contrasts, there are certain surface phenomena that it may be possible to represent on either tier.

For example, while many lexical stress systems are represented with linking of a bare  $\pi$ -node (which serves as an accentual autosegment), it is also possible for unpredictable surface stress to be derived from a quantity contrast represented on the X-tier. In traditional metrical terms, this would be a quantity-sensitive stress language, but one with a covert quantity distinction (“honorary weight”). Unpredictable stress could formally come from association with an autosegment or a length contrast, but the number of contrasts available in the two systems is the same. This possibility provides a restrictive explanation for limited exceptions to Dubina’s (2012) “Incompatibility of Freedoms Hypothesis”: a language may in fact show independent tone and stress contrasts if tone is represented with  $\pi$ -nodes and stress is represented as covert length with X-slots. Crucially, however, a language that does so cannot additionally show a length contrast independently from tone and stress. In other words, it is predicted that a single language can show no more than two independent prosodic contrasts.

## 3.2 Restrictions and predictions

Having established the contrastive elements at play in the present model, I will now make explicit the restrictions on formal configurations of these elements. In general these concern: (i) restricting allowable feature specifications according to the contrastive structure of the prosodic inventory, and (ii) requiring the only point of interface between the prosodic and segmental inventories to be in associations made between the two tiers of root nodes (i.e., X-slots and  $\pi$ -nodes).

### 3.2.1 Feature structure is feature ordering

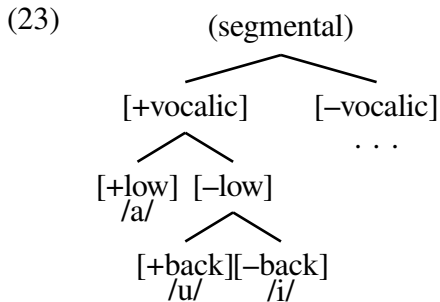
Hierarchical organization is a property of a number of theories of feature specification, including feature geometries (Clements 1985, and many others) and constraint rankings in Optimality Theory (Prince and Smolensky 2004) and related models. Dresher (2009) instantiates hierarchical organization as a way to derive contrastive underspecification in a principled manner; the inventory is divided in a series of cuts one feature at a time. The order in which features are used to make these cuts is determined on a language-specific basis, meaning that different languages can (under)specify different features, even if they have identical surface inventories. The hierarchical properties of the resulting feature specifications are most apparent when notated as branching feature trees, which can themselves be called “contrastive hierarchies”.

While contrastive hierarchies have been insightful in studies of contrastive (under)specification, it has not been entirely clear what role (if any) is played by the hierarchy itself. Is it merely an acquisition device for determining underspecification, with the synchronic representations consisting of only unordered feature bundles, or does the hierarchical nature of feature ordering carry over into phonological representations?

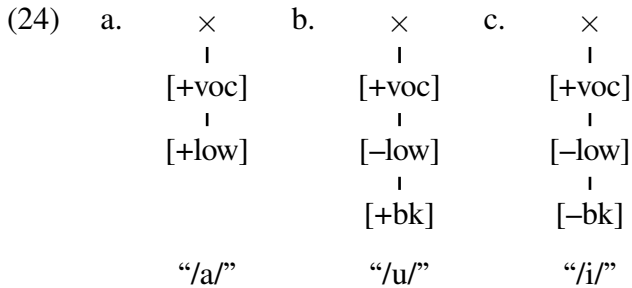
In past work, I have argued on the basis of neutralization processes that the contrastive hierarchy is in fact a phonologically real object which constrains the application of synchronic processes (Spahr 2014a). Specifically, I have claimed that neutralization (that is, the loss of a contrast) involves the deletion of a binary contrastive feature so that a neutralized segment has no specifi-

cation at all for the feature that reflects the lost contrast. More importantly, this “archiphonemic” feature deletion must apply to the feature that is ordered last, rather than arbitrarily deleting any of the specified features—deletion occurs from the bottom of the feature tree. This prohibition is only possible if the hierarchy itself, not only the features it specifies, can be referenced by phonological processes. The result is that licit phonological representations must consist of feature specifications corresponding to nodes of the contrastive hierarchy, where non-terminal nodes reflect neutralized segments (archiphonemes); a feature not at the bottom of the hierarchy cannot be deleted without first deleting all of the features below it.

Let us now consider what this looks like when contrastive representations are broken down with respect to root nodes. Take as a starting point the abbreviated contrastive hierarchy in (23), which reflects the segments of a language with the vowel phonemes /a, i, u/. For expository purposes, I assume that vowels are separated from consonants by the feature [ $\pm$ vocalic]; the portions of the tree under [ $-$ vocalic] (which reflect the consonantal sub-inventory) are not shown.

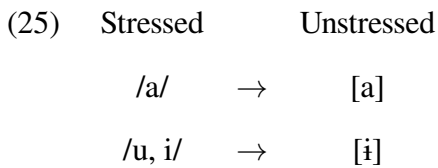


In the conception of segmental representations that I espouse here, the contrastive hierarchy tree serves as a shorthand for the representations of all phonemes in the inventory. I explicitly assume that the root node of the contrastive hierarchy corresponds directly to the root node which serves as a linking point for contrastive features in multi-tiered representations, so here “(segmental)” means that these are features which must link to an X-slot. The order in which features link starts at the tree’s root node and moves down, forming a “chain” one feature at a time until the terminal node is reached. The hierarchy in (23) thus expands to the inventory in (24):



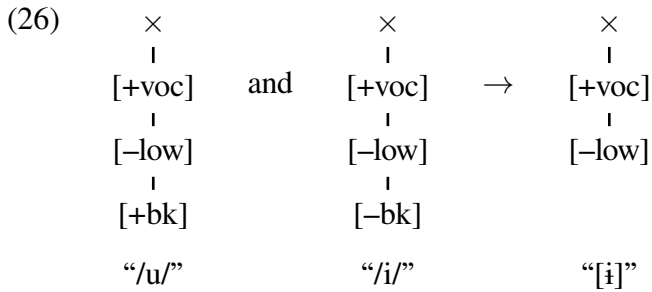
Note that while features are ordered, they are not fully feature geometrical. That is, although a contrastive hierarchy resembles a feature geometry on the page, and can be invoked to account for similar kinds of data, there is no branching in the actual featural representations that it entails. Rather, the hierarchy reflects a set of one-dimensional feature chains. This is distinct from the feature geometric approach, where the featural organization of an individual segment *is* multi-dimensional, being able to contain a series of branches.

Several things fall out from these properties of contrastive representations. The first concerns feature deleting processes, which I have claimed only apply to features at the bottom of the contrastive hierarchy. Let us imagine that a language with the inventory in (23) has a phonological vowel reduction process whereby the contrast between /u/ and /i/ is lost in unstressed position, resulting in a single segment pronounced [ɨ]:

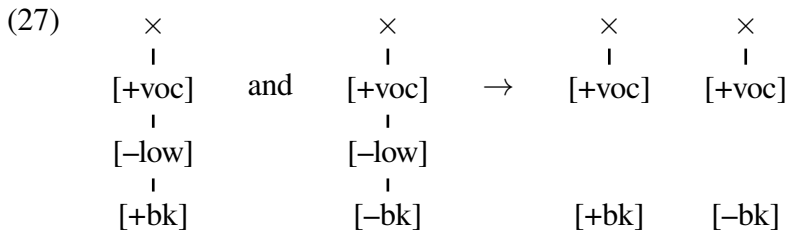


With respect to the contrastive hierarchy, this constitutes a loss of backness which can be formalized as the deletion of all values of the feature [±back]. The terminal nodes /u/ and /i/ in stressed position thus alternate with a non-terminal node pronounced [ɨ] when they appear in unstressed position. Looking at the representations in (26), we can see that this involves simply trimming features off of the ends of chains, making them shorter (i.e., reflective of a smaller number of contrasts):

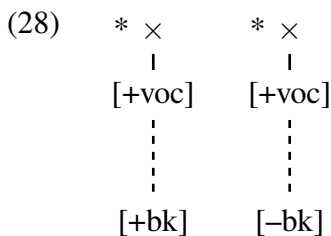




Now imagine a process that deletes the feature  $[\pm\text{low}]$ , which is non-terminal for the segments /u/ and /i/. If representations consisted simply of unordered feature bundles, there would be no obvious way to rule out such a process. With feature ordering as in (27), however, we see why such a process is problematic: the chain is severed and the features [+back] and [-back] are left unconnected to segmental root nodes.

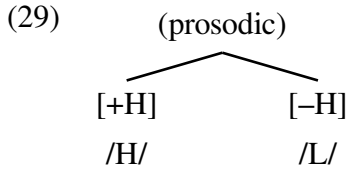


There is no way for  $[\pm\text{back}]$  features to reconnect with a root node by relinking to the (now terminal) [+vocalic] feature, because the resulting feature specifications ( $[\text{+vocalic}, \text{+back}]$  and  $[\text{+vocalic}, \text{-back}]$ ) do not correspond to nodes of the contrastive hierarchy in (23):

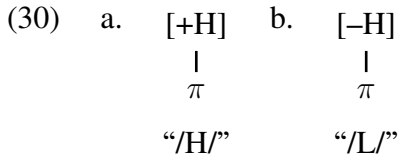


The concept of hierarchically determined contrastive feature chains linked to a root node carries over to prosodic features, as well. However, because prosodic features link to a separate kind of root node ( $\pi$ -nodes rather than X-slots), their hierarchy is formed in isolation from that of segmental features. Suppose that the language under discussion has an inventory of two contrastive

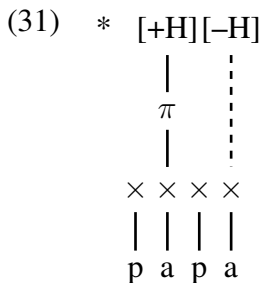
tonemes /H/ and /L/. These would be divided by a single binary contrastive feature, giving the following prosodic contrastive hierarchy:



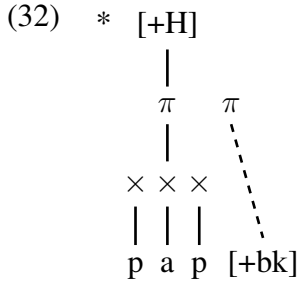
Possible representations correspond to nodes of this hierarchy, again linked one at a time from the root node outward, into chains connected to  $\pi$ -nodes. The hierarchy in (29) thus expands to the inventory in (30):



Since prosodic and segmental contrastive features belong to different inventories, it is not possible for the prosodic feature  $[\pm H]$  to link directly to an X-slot, as in (31); a chain consisting of  $[\pm H]$  linked to an X-slot corresponds to no node of either hierarchy. Note that here, as in other representations, the symbols “p” and “a” are notational shorthand for ordered feature chains corresponding to the terminal phonemes /p/ and /a/ in the segmental inventory.

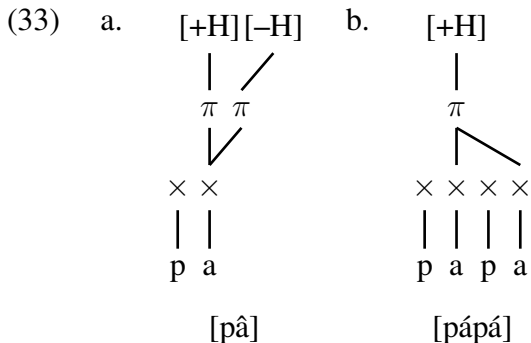


This restriction on cross-linking is paralleled on both tiers, so that just as prosodic features cannot link directly to segmental root nodes, segmental features cannot link directly to prosodic root nodes:



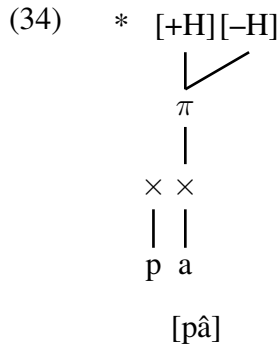
### 3.2.2 Restrictions on several-to-one linking

I assume that links between the two tiers of root nodes are the only point of interface in which both several-to-one and one-to-several associations are permitted. These configurations are used to represent contour tones, for example, as in (33a) and tone spreading (as in 33b):

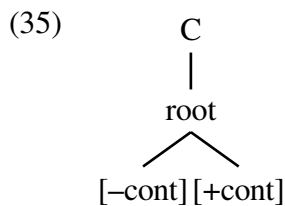


Branching between root nodes and features is more restricted, however. The relationship between root nodes and features may not be branching. In other words, a single root node may host no more than one feature chain. This means that composed contour tones must be represented by linking of two different  $\pi$ -nodes as in (33a), never by linking of two parallel features to a single  $\pi$ -node as in (34):<sup>5</sup>

<sup>5</sup>This restriction raises questions about the representation of short diphthongs, which in the present model could have two different phonological interpretations depending on the language: (i) they could be separate members of the segmental inventory, not being composed of two separate vowel melodies (and thus analogous to non-composed affricates or contours, as discussed below), or (ii) they could be composed of two X-slots, with their length being represented on the  $\pi$ -tier. The latter interpretation applies to long versus overlong diphthongs in Estonian, as analyzed in Chapter 5.

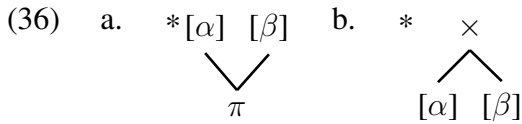


Yip (1989) argues for a distinction between contour tones like (33a) and those like (34) on the basis of phonological behaviour; the former are used in cases where contours show independence of their component parts (“tone clusters”), while the latter are used where contours map and/or spread as whole units. She likens this to the use of branching segmental roots to distinguish prenasalized segments and affricates from consonant clusters, as proposed by Clements (1985) and Sagey (1986). Affricates are said to branch for their value of  $[\pm\text{continuant}]$  but pattern as a single segment:

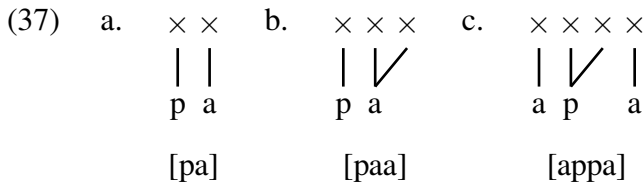


Lombardi (1990) has argued against representations like (35), given that inherent linear ordering of feature specifications implies that ordering can be contrastive for affricates. Lombardi claims that no such contrasts are attested. Following this, such representations are forbidden in the present model. Affricates (as well as non-cluster contour tones), insofar as they constitute single-unit contrasts, must belong to an inventory divided by binary features, and be represented with non-branching contrastive feature chains. Single (auto)segments are not analyzed with branching structures precisely *because* they cannot be decomposed linearly. Krekoski (2013) takes a similar position on the status of tones in Chinese languages, arguing that tone classes (both level and contour) must be treated as separate categories differentiated with contrastive features, not decomposed on the basis of their surface phonetics.

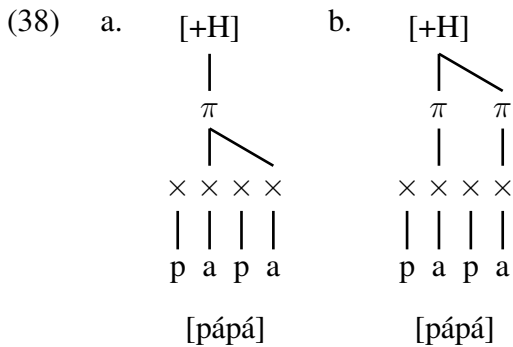
The formal restriction on branching root nodes can be schematized as follows:



The other configuration to consider is the sharing of a single feature chain across several root nodes. In the case of quantity distinctions, it is clear that two-to-one linking of root nodes to segmental features is desirable where segmental root nodes (X-slots) serve directly as timing units, as argued by Selkirk (1990). We thus retain conventional representations of quantity contrasts like the following:



It is less clear whether the same configurations are available for  $\pi$ -nodes, and what, if at all, they would represent (though see §3.2.5 below for one possibility). In the case of tone spreading, for example, two possible representations would be available:



While there is no way to rule out (38b) without stipulating that prosodic root nodes are different from segmental root nodes in their feature-bearing abilities, it is not immediately obvious what kind of phonological activity would distinguish the configurations in (38). At the very least, the spreading of features across several  $\pi$ -nodes would only be available in a binary (or greater) tone system, as no features would be assigned in a privative system. I leave this issue open.

### 3.2.3 On the “maximum of two” restriction

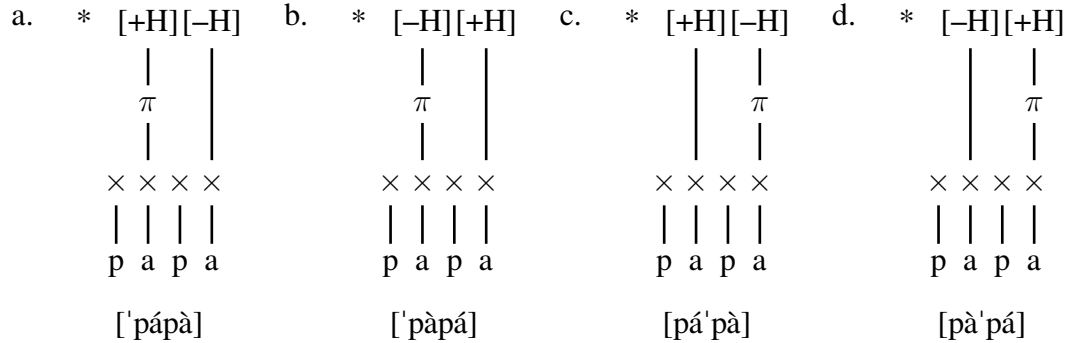
Recall that one of the empirical predictions of the present model of word prosody is that no system may show three different prosodic contrasts (say, tone, length, *and* stress) with full syntagmatic independence. This means that sets of contrasts like that in (39) may not exist in a single language.

(39) *A set of contrasts predicted to be impossible*

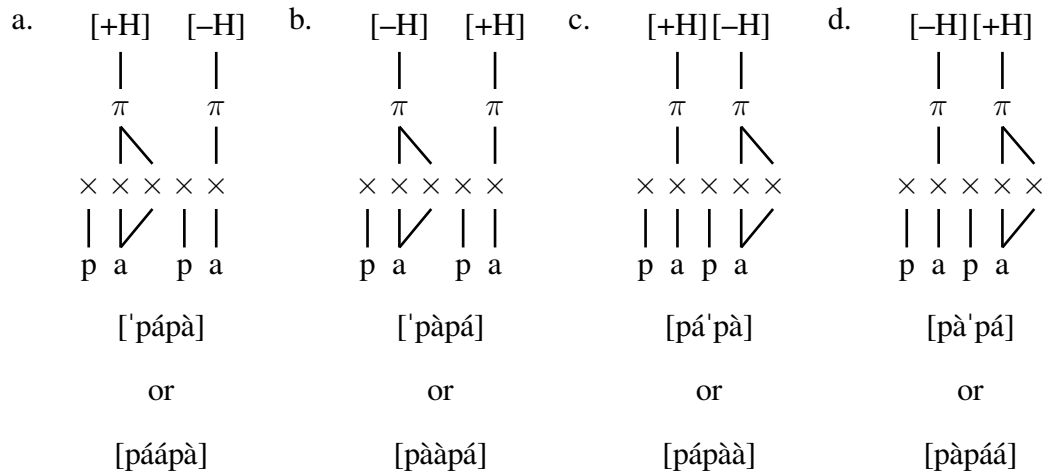
	'V-V	'VV-V	'V-VV	'VV-VV	V-'V	VV-'V	V-'VV	VV-'VV
HH	'pápá	'páápá	'pápáá	'páápáá	pá'pá	páá'pá	pá'páá	páá'páá
HL	'pápà	'páápà	'pápàà	'páápàà	pá'pà	páá'pà	pá'pàà	páá'pàà
LH	'pàpá	'pàápá	'pàpáá	'pàápáá	pà'pá	pàà'pá	pà'páá	pàà'páá
LL	'pàpà	'pàápà	'pàpàà	'pàápàà	pà'pà	pàà'pà	pà'pàà	pàà'pàà

This restriction can be broken down into smaller pieces. Consider first tone and stress. We will see in Chapter 4 that in languages like Goizueta Basque and Choguita Rarámuri, tone and stress are both contrastive in the sense that they can be manipulated separately from each other. However, in these languages, the position of tone and the position of stress are not independent from each other; descriptively, it can be said that tone is only realized in stressed syllables. This follows from the assumption that the root node of the tonal features (the  $\pi$ -node itself) is an autosegment representing stress.

Because contrastive features must be hosted by a root node in order to constitute licit lexical representations, the placement of tone features independently from a  $\pi$ -node as in (40) is prevented categorically. If each tonal feature *were* linked to a  $\pi$ -node, on the other hand, then there would be no way to contrastively distinguish which  $\pi$ -node is the “stress” in lexical representations. This captures Dubina’s (2012) “Incompatibility of Freedoms Hypothesis”.

(40) *Ill-formed representations for co-existing “tone” and “stress”*

The X-tier is, of course, available for the representation of prosodic contrasts, so it is possible to use  $\pi$ -nodes and tonal features to represent “tone”, with the independent syntagmatic contrast of “stress” being represented as a formal length contrast (that is, “honorary weight”) as in (41). However, since this means using the same representation for length and stress, a language using such representations may only have one of contrastive length and contrastive stress:

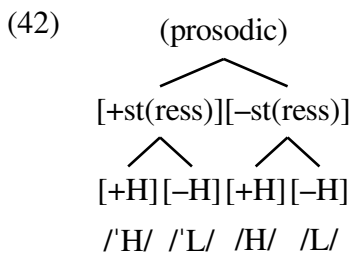
(41) *Stress and/or length on the X-tier*

It should be pointed out that a stress language and a quantity language (regardless of whether there exists a tone contrast) may not actually be distinguishable from each other formally. In a “quantity-sensitive” system where stress is always predictable, non-default stress occurs on heavy syllables, i.e., groups of more than one X-slot. When weight is contrastive and stress (as in independent entity) is not, the two can be thought of as one and the same: a single property is specified

at the lexical level, and it corresponds on the surface to both additional duration and a phonetic property of “stress”.

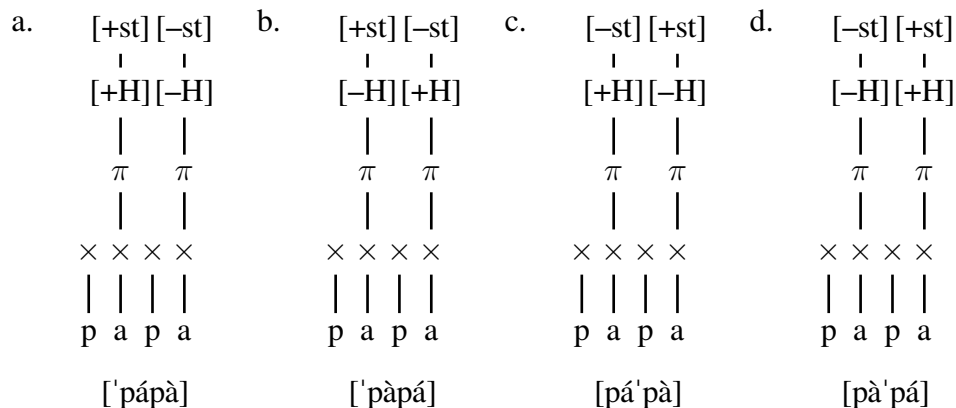
### 3.2.4 Stress as a feature

The model does predict some possibilities that may allow systems to show what, on the surface, appear to be three independent prosodic contrasts. For example, if “stress” is a binary feature alongside “tone”, rather than being represented by the  $\pi$ -node itself, then stress and tone could be manipulated independently of each other:



This would allow representations like those in (43), along with which a length contrast could exist on the X-tier.

(43) “Stress” and “tone” as features



It may be possible to avoid this problem if only root nodes can be restricted to being culminative/obligatory. If this is done, the placement of  $\pi$ -nodes, but not features, can be restricted within the word. That is, words can be restricted to a single placement of a  $\pi$ -node, but cannot



be limited to a single placement of one value of a given feature, so that there is no way to require a word to have one and only one [+stress] feature.

Another possible factor restricting systems like that in (43) may have to do with the nature of language-specific features. Perhaps features (the number of which is unbounded) may only contrastively specify a phonetic property along a single dimension (e.g., pitch), whereas  $\pi$ -nodes may specify along several (e.g., intensity and duration). This would mean that a canonical lexical stress system, where stress has several phonetic correlates, can only be represented using  $\pi$ -nodes (or perhaps X-slots), but not by features. However, the nature of phonetic implementation at this level of detail is beyond the scope of the present study, so an in-depth investigation of these issues will have to be revisited in future work.

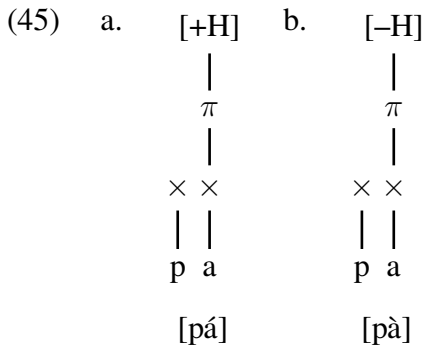
### 3.2.5 Stress as doubled $\pi$ -nodes

The present model structures lexical representations around two tiers of root nodes which host separate sets of contrastive binary features: the X-tier hosts segmental features while the  $\pi$ -tier hosts prosodic features. Contrasts can be represented using  $\pi$ -nodes in a few different ways. First, the presence versus absence of a  $\pi$ -node can be used to represent the syntagmatic contrast seen in privative tone and free stress systems:

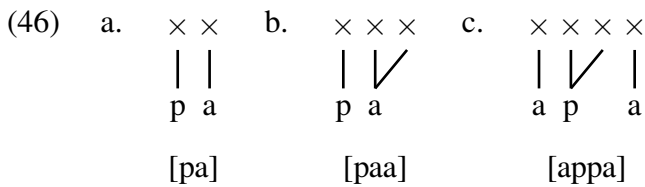
- (44) a.  $\pi$  b.
- |       |                            |
|-------|----------------------------|
|       |                            |
| × ×   | × ×                        |
|       |                            |
| p a   | p a                        |
| [pá]  | [pà] (privative high tone) |
| or    | or                         |
| [ˈpa] | [pa] (free stress)         |

Paradigmatic prosodic contrasts are represented with contrastive features on  $\pi$ -nodes, giving a binary tone system, or, when the  $\pi$ -nodes themselves simultaneously enter into a syntagmatic

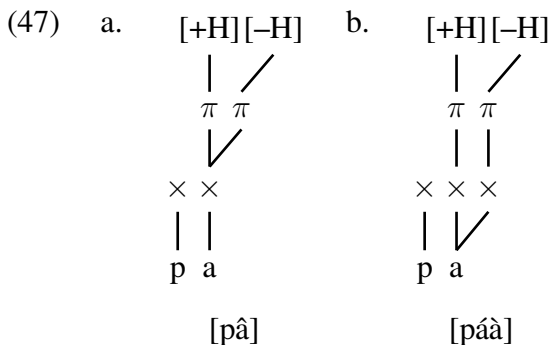
contrast, a language where tone (represented with binary features) is only contrastive in stressed positions (represented by the presence of a  $\pi$ -node):



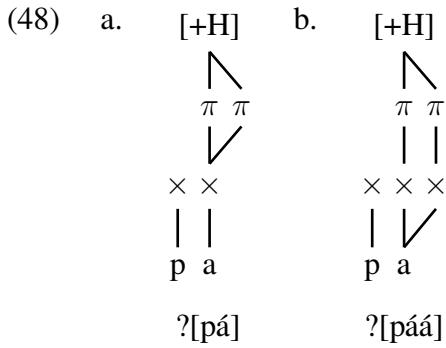
Segmental root nodes (X-slots) can also be used to represent a prosodic contrast (in a syntagmatic sort of way) by doubling them to represent long segments:



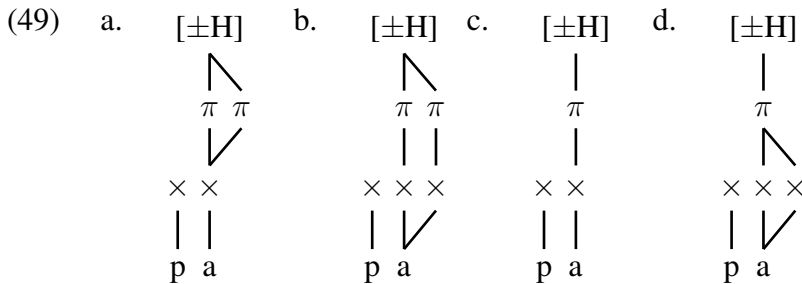
Compositional contour tones are represented by the linking of two  $\pi$ -nodes bearing separate features. For long contour tones, this gives representations like in (47b), where one  $\pi$ -node links to each X-slot. In languages that allow short contour tones, however, this means that two  $\pi$ -nodes are linked to a single X-slot, as in (47a):



One possibility that we have not yet considered, however, is whether it is possible to use the double linking of  $\pi$ -nodes to X-slots in a contrastive manner, as a sort of short contour tone without the contour:

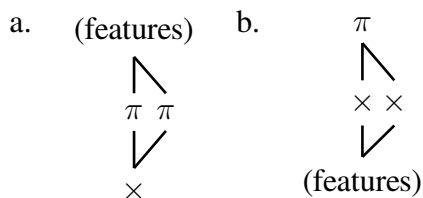


If  $\pi$ -nodes have language-specific phonetic realizations and their doubling can have some kind of additive effect (analogous to the additive effect of doubled X-slots for representing length contrasts), then there is no obvious way to rule out systems which co-vary doubled  $\pi$ -nodes,  $\pi$ -node features, and doubled X-slots, yielding a system with three prosodic contrasts, as in (49). Doubled  $\pi$ -nodes could perhaps represent “stress” in such a system, but I give only a set of formal contrasts, not phonetic realizations:



The representation in (48a/49a) seems unlikely, but if the two tiers of root nodes are considered to be more or less parallel in the configurations that they are allowed to appear in, then there is no difference in the number and type of association lines seen in (49a) and (49d); both contain a single feature chain linked to two root nodes, with those two root nodes being linked to a single root node on the other tier:

(50) *Parallel configurations*



Although the present model does predict systems where there is a contrast between single and doubled  $\pi$ -nodes linked to a single X-slot (at least without stipulating additional restrictions), this contrast cannot be entirely independent from the others. For a language with short contour tones, all short contour tones would necessarily be “stressed”, since all short contours would necessarily bear two  $\pi$ -nodes (one for each feature in the contour). Furthermore, the contrastive use of doubled  $\pi$ -nodes separately from tone would only be available in non-privative tone systems, since there must be tonal  $\pi$ -node features present in order to exploit three separate contrasts in all positions.

### 3.3 Summary

In this chapter, I have argued that tone and lexical stress are both represented by a single element, the  $\pi$ -node, which is formally equivalent to a tonal root node. In binary or greater tone systems,  $\pi$ -nodes bear contrastive feature chains derived from the prosodic contrastive hierarchy, while in privative tone systems and stress accent systems, the  $\pi$ -node is used as a single privative autosegment without bearing any features. The phonetic realization of the  $\pi$ -node is determined language-specifically, and it need not be given a substantive name like “tone” or “stress”.

$\pi$ -nodes link to a tier of segmental root nodes (X-slots) with essentially the same formal properties as the skeletal positions/moras in standard models of quantity. Because it is possible for a surface stress contrast to be derived from covert quantity (“honorary weight”), stress can be represented on the X-tier instead of with  $\pi$ -nodes. This means that it is possible for a language to simultaneously show a tone contrast and a stress contrast, but in doing so must sacrifice the ability to have a separate quantity contrast. Thus, the proposed model predicts that no more than two independent prosodic contrasts may coexist in a single system.

The model I have laid out here provides equivalent mechanisms for capturing the activity covered by existing models of prosodic contrasts, and it does so with an integrated architecture in which different prosodic features are not represented with disparate devices. In the following two

chapters, I will show that in addition to this advantage, it also allows for elegant and morphologically compositional accounts of the prosodic activity seen in mixed prosodic systems: languages with more than one prosodic contrast. It is by looking at such cases that the claims about prosodic contrast co-occurrence and interaction can be tested.

# Chapter 4

## Tonal accent and the $\pi$ -node

### 4.1 Introduction

In the previous chapter, I argued for the autosegmental treatment of accent on the grounds that it allows for a unified account of word-prosodic contrasts based on a single set of phonological primitives. In addition to its minimalist appeal, however, this model straightforwardly predicts the class of tonal accent languages—those with both tone and stress on the surface, but where the realization of tone is restricted to the position of stress. Such systems exist because  $\pi$ -nodes perform two roles: they serve as the autosegments marking accented positions (having stress-like phonetic realizations), and at the same time they serve as the root nodes to which tonal contrastive features link. Thus, the dependency between tone (the features) and accent (the  $\pi$ -node that bears them) is built into the model of representations.

In this chapter, I develop analyses of the prosodic systems of two languages of this type: Goizueta Basque and Choguita Rarámuri. These case studies demonstrate the formal dependencies between tonal features and accentual  $\pi$ -nodes in mixed prosodic systems, which become clear when considering the individual processes at play in such systems. At the same time, however, the alternations seen in languages of this type support the degree of independence provided by separating prosodic representations onto a root node–accentual tier and a featural tier.

## 4.2 Goizueta Basque

The variety of Basque spoken in the town of Goizueta (in Navarre, Spain) has been investigated recently by Hualde et al. (2008), Hualde et al. (2010), and Hualde (2012), among others. The language is notable because it is said to have both lexical stress and lexical tone. While the population of Goizueta is only about 900, the local variety is used regularly by speakers of all ages (Hualde et al. 2008).

In this section, I will examine the language's prosodic contrasts and develop a system of rules and representations that accounts for alternations involving both contrasts using the model proposed in the present study. I will argue that in Goizueta Basque, the  $\pi$ -nodes are used to autosegmentally represent the syntagmatic stress contrast, while prosodic features linking to those  $\pi$ -nodes represent the tonal contrast. This captures the fact that tone is always realized in the same position as stress, since the tonal contrast is hosted by the stress contrast. I then demonstrate that, using the elements of these representations, it is possible to formulate an elegant and morphologically compositional analysis of several prosodic alternations.

### 4.2.1 The accent system

Hualde et al. (2008: 1) state that “both stress and tone are lexically contrastive” in Goizueta Basque. According to their description, the language shows a lexical stress contrast, where stress can fall on either the initial or post-initial syllable. The stressed syllable is then specified for one of two tonal melodies, which they call “Accent 1” (notated with an acute accent) and “Accent 2” (notated with a grave accent). These are lexically contrastive in all words of two or more syllables, even in uninflected forms, as in (1).<sup>1</sup> Because neither the location of stress nor the tone realized in that location are predictable and they independently covary, both are contrastive.<sup>2</sup>

---

<sup>1</sup>In this section, the following abbreviations are used: abs. = absolute, agent. = agentive, com. = comitative, dat. = dative, erg. = ergative, excess. = excessive, pl. = plural, sg. = singular.

<sup>2</sup>The examples in (1) are representative of the lexicon as a whole. According to Hualde (2012: 1342–1344), Accent 1 words with initial stress are all bisyllables, and most Accent 2 words with initial stress are borrowings. I assume here that these are simply facts about the lexicon, and not distributional restrictions on prosodic contrasts.

(1) *Stress and accent contrasts in uninflected words* (Hualde et al. 2008: 3, adapted)

	<b>Postinitial</b>		
<b>Accent 1</b>	ardí ‘sheep’	mendí ‘mountain’	gizón ‘man’
	alába ‘daughter’	ittúrri ‘fountain’	abérats ‘rich’
	basérritar ‘farmer’	emákume ‘woman’	txistúlari ‘flutist’
	<b>Initial</b>		
	úme ‘child’	áma ‘mother’	séme ‘son’
	<b>Postinitial</b>		
<b>Accent 2</b>	intxàur ‘walnut’	purè ‘puree’	Fermìn ‘a name’
	belàrri ‘ear’	tipùla ‘onion’	eskòla ‘school’
	<b>Initial</b>		
	bàso ‘glass’	mòro ‘Moor’	lìbru ‘book’
	mèdiku ‘doctor’	fàbrika ‘factory’	lèngusu ‘cousin’

The phonetic correlates of stress are increased duration along with increased intensity relative to unstressed syllables—pitch is not used as a correlate of stress (Hualde et al. 2008). The tonal contrast, realized within the stressed syllable, is conveyed with differences in the level and contour of pitch. Accent 1 has a high rising “circumflex” contour, while Accent 2 has a falling contour starting from mid or low pitch (Hualde et al. 2008, 2010).

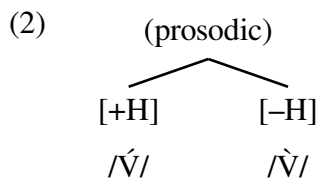
Although Hualde et al. (2008) refer to stress and tone as “independent” contrasts, they recognize that the location of tone is entirely dependent on stress; although tone can be manipulated separately, the location of tone is always predictable from the location of stress. In other words, Goizueta Basque does not violate Dubina’s (2012) Incompatibility of Freedoms Hypothesis since the location of stress and the location of tone do not independently covary. The notation used by Hualde et al. (2008, adopted here) captures this fact because for each word, only the location of tone needs to be marked; separate marking of stress is superfluous.

The model developed in the present study straightforwardly predicts prosodic systems like the

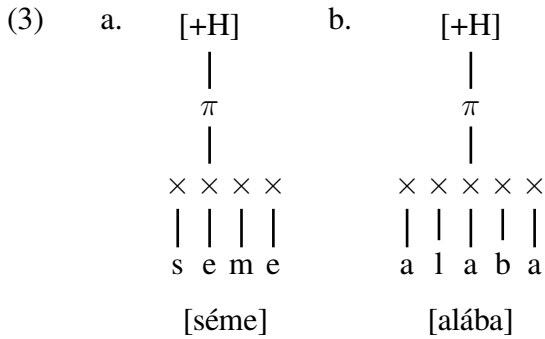


one in Goizueta Basque. We have seen that  $\pi$ -nodes are phonological primitives that can have language-specific phonetic realizations. In privative tone systems, segments linked to  $\pi$ -nodes are realized with higher (or lower) pitch than surrounding segments, while in stress accent languages,  $\pi$ -nodes are realized with the phonetic correlates of “stress”. In non-privative tone languages, on the other hand,  $\pi$ -nodes serve as tonal root nodes, hosting the contrastive binary features which divide the prosodic inventory. Goizueta Basque is a kind of hybrid system that uses  $\pi$ -nodes with independent phonetic realizations to represent stress, where at the same time those  $\pi$ -nodes bear binary contrastive features to represent “tone”. The  $\pi$ -nodes here are accentual in that they abstractly mark only a single position in a word, that is,  $\pi$ -nodes are culminative and obligatory. Although the  $\pi$ -tier resembles that of a language with free stress-accent, the use of features linked to  $\pi$ -nodes resembles a non-privative tone language, because each  $\pi$ -node must bear contrastive features specified in the prosodic contrastive hierarchy.

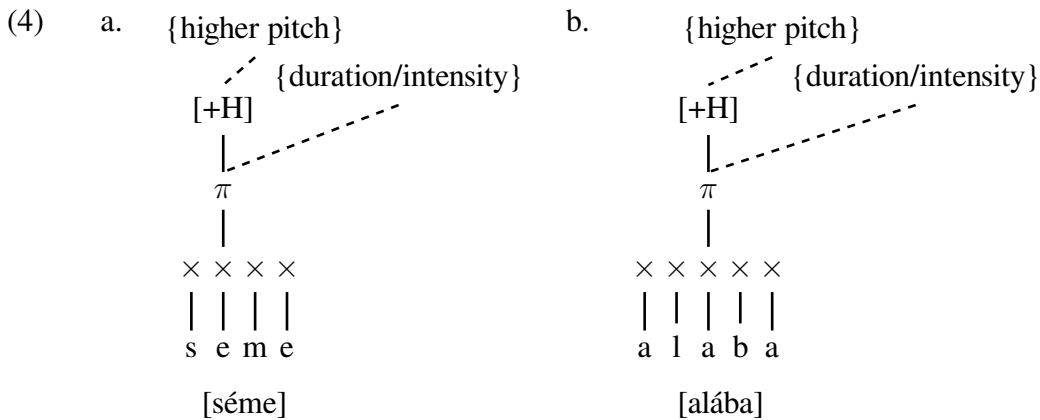
In Goizueta Basque, the prosodic inventory consists of two members (Accent 1 and Accent 2), which need to be divided by one contrastive feature in the prosodic contrastive hierarchy. The hierarchy will have the shape in (2), where Accent 1 is notated  $/\acute{V}/$  and Accent 2 is notated  $/\grave{V}/$ . I choose to divide the inventory based on a feature called  $[\pm H]$ , as a mnemonic for the fact that Accent 1 reaches a higher pitch than Accent 2.



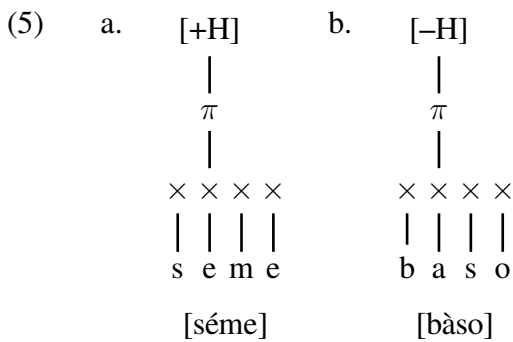
These features are linked to  $\pi$ -nodes, so that a word with Accent 1 contains a single  $\pi$ -node bearing the feature  $[+H]$ . The difference between an Accent 1 word with initial stress, as in (3a), and one with post-initial stress, as in (3b), is simply the position to which the  $\pi$ -node links:



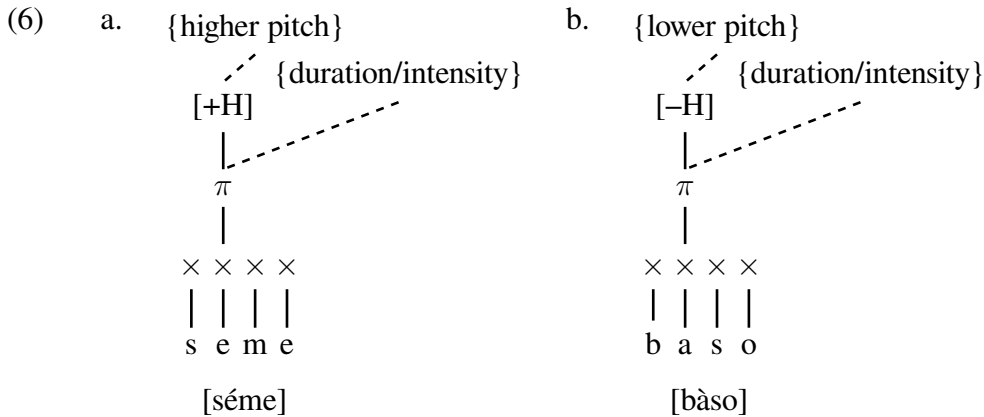
In both words, the phonetic correlates of stress (duration and intensity) are assigned by the  $\pi$ -node itself and the phonetic realization of pitch is assigned by the contrastive feature on that  $\pi$ -node. Thus both words paradigmatically receive the same tonal properties but syntagmatically receive stress in a different location:



When words differ only in their tonal contrast,  $\pi$ -nodes link to the same location, but the specification of contrastive features on those  $\pi$ -nodes differs:  $[+H]$  in (5a) and  $[-H]$  in (5b).



Duration and intensity (reflective of stress) are assigned by the  $\pi$ -node to the same location, but the tonal feature  $[\pm H]$  assigns different pitch realizations to the stressed position:



## 4.2.2 Alternations

In addition to being lexically contrastive, tones can undergo morphophonological alternations. As seen in (7), nominals with Accent 1 base forms retain the position of stress and value of tone across all singular forms, regardless of the segmental suffixes they take. In the plural forms, segmental suffixes are still added, but in addition to this, Accent 1 is replaced by Accent 2. This alternation is purely tonal, in that it does not change the position of tone—stress stays on the same (lexically specified) syllable, but the lexical tone changes.<sup>3</sup>

<sup>3</sup>A brief note on morphology is in order. The base forms given for the examples here, such as *gizón* ‘man’ in (7), are unsuffixed bare roots. In the (Standard) Basque noun phrase, only the last word in a noun phrase containing a single noun receives inflectional marking (Hualde and de Urbina 2003: 171). Thus, bare forms are used when the noun appears non-phrase-finally. The definite article /-a/ appears before case suffixes in inflected definite forms and triggers deletion of a root-final vowel, if any. The morphophonological behaviour of this suffix varies in different dialects (Hualde and de Urbina 2003: 175–176), but no sources are explicit about the morphemic or morphophonemic status of /-a/ in the Goizueta variety. Since more clear data is unavailable and, at any rate, this morpheme has no effect on the realization of prosody, I will consider it part of the case suffix for the purposes of discussion here, except when sources notate it otherwise.

(7) *Singular/plural accent alternations* (Hualde et al. 2008: 4, adapted)

	<b>Singular</b>	<b>Plural</b>
gizón ‘man’		
abs.	gizóna	gizònak
erg.	gizónak	gizònak
dat.	gizóinari	gizòinari
com.	gizóinakin	gizòinakin
mendí ‘mountain’		
abs.	mendía	mendìk
erg.	mendìk	mendìk
dat.	mendìri	mendìri
com.	mendìkin	mendìkin
séme ‘son’		
abs.	sémea	sèmek
erg.	sèmek	sèmek
dat.	sèmeri	sèmeri
com.	sèmekin	sèmekin

Note that the tonal nature of the distinction is particularly evident in cases other than the absolutive. In the absolutive, the singular and plural forms differ in terms of both their segmental suffixes (/–a/ versus /–(a)k/) and the tone realized in the stressed syllable (Accent 1 versus Accent 2). In all other cases, however, the singular and plural forms are segmentally identical, being distinguished solely on the basis of tone.

The pattern of plural forms acquiring Accent 2 carries over into monosyllabic stems as well. While monosyllabic words (unsuffixed monosyllabic stems) do not show tonal contrasts (Hualde et al. 2008: 3, fn. 3), longer words derived from the same roots do show the contrast, with Accent 1 occurring in the singular and Accent 2 occurring in the plural:

(8) *Derived tonal contrast in monosyllabic stems* (Hualde et al. 2008: 3, fn. 3, adapted)

lúrr-a ‘the land’    lùrr-ak ‘the lands’

Words that are lexically specified for Accent 2 do not undergo any tonal alternations in their plural forms, and take the same segmental suffixes as those with lexical Accent 1. This means that, apart from the absolutive case (which marks number segmentally), all singular and plural forms are homophonous:

(9) *Words without singular/plural contrast* (Hualde et al. 2008: 5, adapted)

	bàso ‘glass’	belàrri ‘ear’
	Sg./Pl.	Sg./Pl.
abs.	bàsoa/bàsok	belàrri/belàrrik
erg.	bàsok	belàrrik
dat.	bàsori	belàrri
com.	bàsokin	belàrrikin

Furthermore, words with lexical Accent 1 do not undergo the alternation to Accent 2 when their stems contain more than three syllables, as in (10). Hualde et al. (2008: 5, fn. 6) offer no explanation for this, but they do note that in some related varieties, singular/plural alternations are even more restricted, occurring only with stems of one or two syllables.

(10) *No alternation in words of more than three syllables* (Hualde et al. 2008: 5, adapted)

	emákume ‘woman’
	Sg./Pl.
abs.	emákumea/emákumek
erg.	emákumek
dat.	emákumeri
com.	emákumekin

Hualde et al. (2008) schematize the generalizations about tonal alternations with a single morphophonological rule:

(11) *Plural Accent Rule* (Hualde et al. 2008: 5)Accent 1  $\rightarrow$  Accent 2 in plural

(Condition: if stem has three or fewer syllables)

The rule stated in (11) is intended to be purely descriptive of the inflectional role that tone plays, not to make a theoretical point. Tonal alternations in Goizueta Basque are not limited to plural marking—they also occur when several other segmental suffixes are added, including the superlative and excessive forms of adjectives, and several derivational suffixes. As with the singular–plural alternation, these only occur when the stem is lexically Accent 1. Examples are given in (12).

(12) *Accent 1  $\rightarrow$  Accent 2 in derivational morphology* (Hualde et al. 2008: 5, adapted)

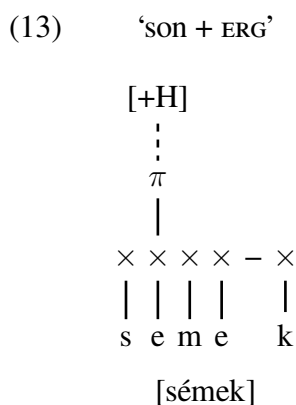
béltza	‘black (abs. sg.)’	bèltzana	‘the blackest one’	bèltzagi	‘too black’
zuría	‘white, (abs. sg.)’	zurina	‘the whitest one’	zurigi	‘too white’
urdína	‘blue, (abs. sg.)’	urdinana	‘the bluest one’	urdinagi	‘too blue’
beldúr	‘fear’	beldürtia	‘the fearful one’		
herría	‘the village’	herritarra	‘the villager’		
hiría	‘the town’	hiritarra	‘the citizen’		

Having established a system of representations for Goizueta Basque word prosody and describing the nature of morphophonological alternations bearing on these prosodic contrasts, I will now demonstrate how it is possible to use this system to account for the observed alternations in a compositional morphological analysis.

### 4.2.3 Morphemic analysis

As established above, stress is positionally marked by linking a  $\pi$ -node to the vowel that bears stress, and I assume that this linking is present underlyingly on roots. Tone, however, represented

by the binary contrastive feature  $[\pm H]$ , does not need to be linked underlyingly, as it will necessarily be realized through linking to the only available  $\pi$ -node. Affixal morphology, such as the ergative suffix, is concatenated linearly.<sup>4</sup> Thus *sémek* ‘son (erg. sg.)’ is derived as in (13), where the root consists of a  $\pi$ -node linked to the first vowel and a floating  $[+H]$  feature representing tone, which links to the  $\pi$ -node by convention. The ergative suffix, consisting only of an X-slot and the segmental features for /k/, is attached to the end of the word, and is realized as a single segment:

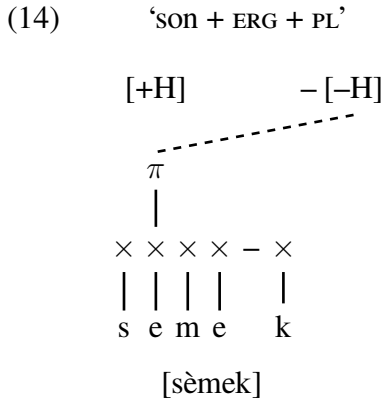


The plural form *sèmek* ‘son (erg. pl.)’ is segmentally identical to the singular, and stress appears in the same position. The only difference is in the realization of tone within the stressed syllable—the plural has Accent 2 instead of Accent 1. An elegant analysis of this alternation is achieved if we assume that it is derived as in (14). The same root morpheme, with a linked  $\pi$ -node and floating  $[+H]$  feature, as well as the same segmental /-k/ ergative suffix, are used, but additionally there is a plural suffix which consists only of a floating  $[-H]$  feature. This plural suffix is concatenated linearly like the other morphemes.

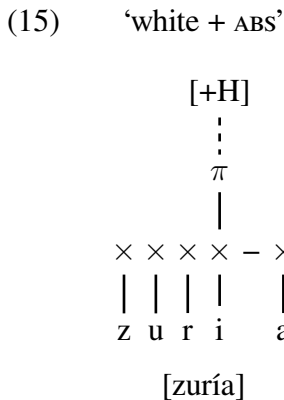
I assume that the association convention linking features to  $\pi$ -nodes occurs from right-to-left in Goizueta Basque. This causes the rightmost floating feature—the one that came from the suffix rather than from the root—to link to the  $\pi$ -node. The unassociated feature is then subsequently deleted or simply left uninterpreted by the phonology. Since the feature linked to the  $\pi$ -node is

<sup>4</sup>I assume here and throughout this study that, in the absence of evidence to the contrary,  $\pi$ -nodes and contrastive prosodic feature chains are not underlyingly linked to other tiers in the absence of evidence to the contrary. The contrastive elements on a given tier are, however, linearized with respect to each other.

now [-H] rather than [+H], the plural is realized with Accent 2 instead of Accent 1, despite being segmentally identical to the singular:



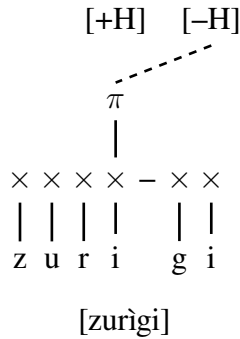
Other morphemes that trigger the Accent 1  $\rightarrow$  Accent 2 alternation function similarly. For the form *zuría* 'white (abs. sg.)', the root contains a  $\pi$ -node linked to the second syllable and a floating [+H] feature, while the suffix is simply segmental /-a/. The root's floating feature links to the  $\pi$ -node, and the surface form is derived:



In the excessive form *zurìgi* 'too white', the root still contains a floating [+H] feature (since it is the same morpheme), but the suffix (which triggers Accent 1  $\rightarrow$  Accent 2) contains both a segmental component /-gi/ and a prosodic component consisting of a floating [-H]. Unlike the plural forms discussed above, this floating [-H] is part of the first suffix; it does not come from a separate suffix. As in (14), the rightmost floating feature ([-H]) links to the  $\pi$ -node due to the right-to-left association convention, and the derived form is realized with Accent 2 instead of Accent 1:

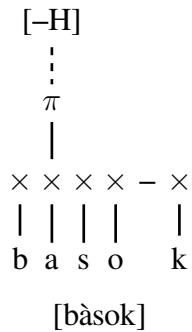


(16) ‘white + EXCESS’

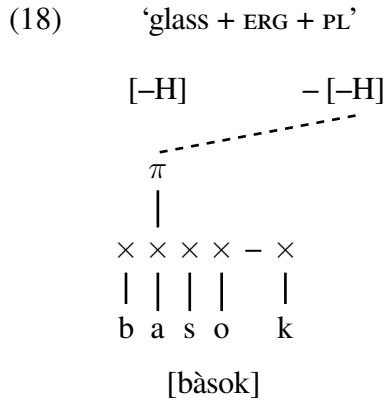


Roots that have underlying Accent 2 differ from those with Accent 1 only in their floating prosodic features. Their stress is still marked by the linking of a  $\pi$ -node in their underlying forms, and they take the same set of inflectional suffixes. Thus *bàsok* ‘glass (erg. sg.)’ take a segmental /-k/ ergative suffix, and is realized with Accent 2 when its only floating feature, the root’s lexical [-H], links to its  $\pi$ -node:

(17) ‘glass + ERG’



In the form *bàsok* ‘glass (erg. pl.)’, the plural suffix consisting of only a floating [-H] feature is also added to the word. The right-to-left association convention then links the rightmost floating feature (the second [-H], which comes from the plural suffix) to the  $\pi$ -node. This results in the word being realized with Accent 2, which is homophonous with its singular form:



Hualde et al. (2008: 5, fn. 5) note that for some speakers, plurals like (18) may not be homophonous with their singular counterparts like (17). In such cases, the plural forms are realized with even lower pitch than the singular forms. Hualde (p.c.) reports that this claim is doubtful, and that it was only reported by one native speaker. Unfortunately no further data on the matter is available. While the forms in (17) and (18) are expected to have identical phonetic realizations, the present analysis derives them from different contrastive representations, so the claim that plurals are realized with lowered Accent 2 can be accommodated if shown to be true. This would be done by modifying the analysis so that a floating [-H] feature causes a linked [-H] feature to its right to be realized with lower pitch, instead of simply being deleted or unrealized. The analysis presented by Hualde et al. (2008), on the other hand, offers no obvious way to accommodate this possibility, as they assume that the form in (18) simply receives vacuous application of their rule requiring Accent 2 in the plural.

#### 4.2.4 Marginal phenomena

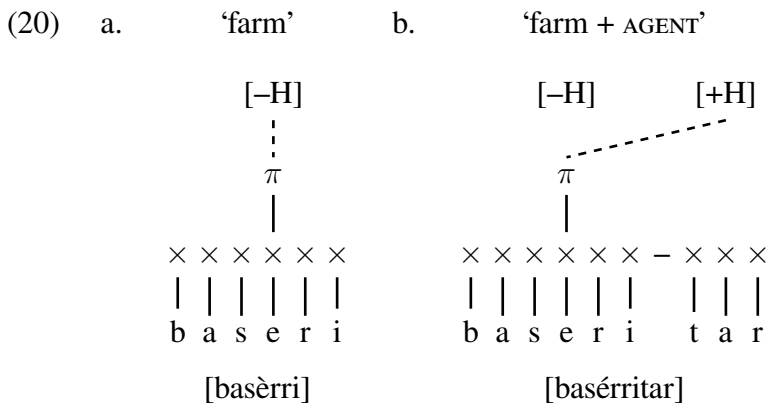
Hualde’s (2012) description is in line with that given by Hualde et al. (2008), but he briefly mentions some additional prosodic alternations in Goizueta Basque. It is clear that alternations deriving Accent 2 from an underlying Accent 1 are commonly triggered by both inflectional and derivational morphemes, while words with underlying Accent 2 generally retain it across their paradigms. Hualde (2012: 1344, fn. 11) states, however, that “[e]xceptionally we find a change

in the opposite direction in derived words”, giving the following example:<sup>5</sup>

(19) *Derived Accent 2*  $\rightarrow$  *Accent 1 alternation* (Hualde 2012: 1344, fn. 11, adapted)

basèrri ‘farm’ basérritar ‘farmer’

No indication is given as to whether the alternation is triggered by this particular morpheme, by the root, or in all contexts involving derivational morphology, and other data given do not make this possible to determine. However, the present understanding of Goizueta Basque word prosody does predict such an Accent 2  $\rightarrow$  Accent 1 alternation to be possible without introducing any additional processes or representational mechanisms. Assuming that the pattern in (19) is productive and triggered by the agentive derivational suffix, the data can be accounted for as in (20). The root underlyingly contains a  $\pi$ -node linked to its second vowel and a floating  $[-H]$  feature. When the root is unsuffixed, as in (20a), the floating feature links to the root and the word is realized with post-initial stress and Accent 2. In the derived form in (20b), an agentive suffix is concatenated to the root. Its segmental shape is  $/-tar/$ , but it also contains a floating  $[+H]$  feature. This rightmost floating feature attaches to the  $\pi$ -node due to the right-to-left association convention, and the derived form is realized in Accent 1.<sup>6</sup>



Hualde (2012: 1344) also mentions an alternation involving a stress shift: “In some instances, the derivational suffix [that causes a change in tone] also causes a change in the position of the

<sup>5</sup>Hualde (2012: 1344) does show that *basèrri* ‘farm’ is etymologically a compound of *bas-* ‘forest’ and *(h)erri* ‘town’, but I will treat it as a monomorphemic root, as this issue has no bearing on the discussion here.

<sup>6</sup>The sequence  $\langle rr \rangle$  is an orthographic convention for a trilled  $/r/$ , in contrast with tapped  $/r/$ , written  $\langle r \rangle$  (Hualde and de Urbina 2003: 29). I thus treat it as a single consonant (linked to a single X-slot) in the contrastive representations.

stress, generally to the syllable immediately preceding the suffix”. The examples given which fit this description are shown in (21). Note that these both involve bisyllabic roots shifting their stress from initial to post-initial position, and coincide with the derivational alternation from Accent 1 to Accent 2.

(21) *Derived stress shifts* (Hualde 2012: 1345, adapted)

lótsa	‘shame’	lotsà-ti-a	‘the ashamed one’
Léitza	‘name of a town’	leitza-rr-a	‘the Leitzan’

Hualde does not make it clear if the “some instances” in which stress shift occurs are determined based on particular roots or particular suffixes, or if stress and tone are simply present in the lexical entries of whole words (without any productive alternation). Because of this it is difficult to offer any definitive analysis, but depending on the actual pattern, the present model provides a means of capturing patterns like those in (21). To demonstrate this, let us assume that it is the case that these particular derivational suffixes trigger stress to shift from the initial to the post-initial syllable in bisyllabic stems, in addition to triggering the Accent 1  $\rightarrow$  Accent 2 alternation. The form *lotsà-ti-a* ‘the ashamed one’, for example, can be accounted for by positing the underlying form of the suffix meaning ‘characterized by’ in (22). The Accent 1  $\rightarrow$  Accent 2 alternation is caused by the floating  $[-H]$  feature, as with the morphemes discussed above, but the suffix also contains a floating  $\pi$ -node, which can be exploited to derive the stress shift.

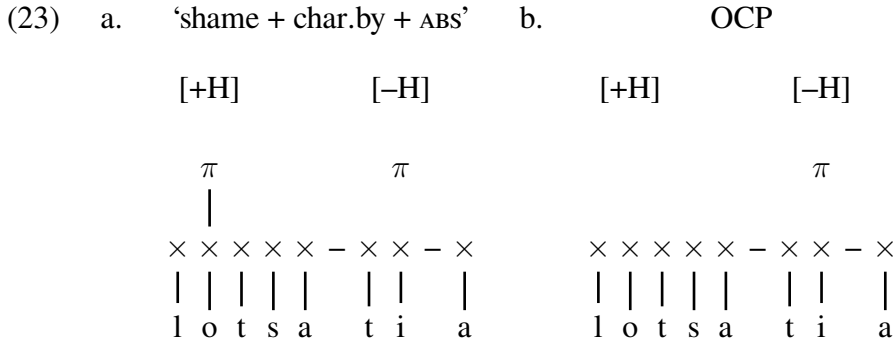
(22)

[−H]
$\pi$
× ×
t i

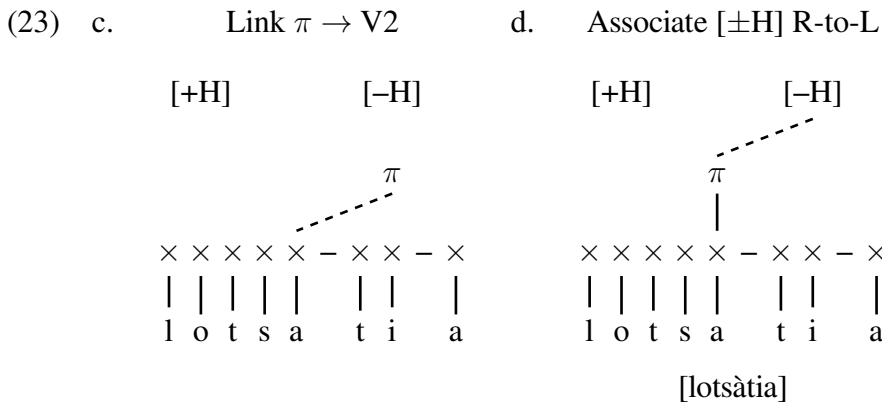
‘characterized by’

The suffix in (22) is attached to the root ‘shame’ (along with the purely segmental absolutive suffix /-a/) as in (23a). This form contains two  $\pi$ -nodes, and because stress accent (represented

by  $\pi$ -nodes) is culminative/obligatory in Goizueta Basque, one of the  $\pi$ -nodes must be removed as a repair. This is done by application of the obligatory contour principle (OCP) in (23b), which (in this language) preserves the rightmost  $\pi$ -node. This removes the root's lexical accent.



Following this, the word contains no marking for stress-accent. I assume that this is repaired by a linking process which associates floating  $\pi$ -nodes link to the post-initial vocalic X-slot, as in (23c). Finally, in (23d), the rightmost floating feature ([-H]) associates by convention to the root's  $\pi$ -node. This set of processes derives a word with post-initial Accent 2 from a stem with underlying initial Accent 1.



While these more marginal alternations in Goizueta Basque word prosody, particularly stress shift, are not described very thoroughly in the literature, it is clear that they are present to some degree. The present model clearly predicts that such processes are possible and provides mechanisms to account for them based on standard assumptions about autosegmental phonology.

### 4.2.5 Summary

I have shown that the word prosodic system of Goizueta Basque, characterized by (partially) independent stress and tone contrasts, is neither exotic nor difficult to explain within the present understanding of non-segmental contrastive representations. The analysis given above separates the contrastive elements that specify the phonetic realizations of stress and tone but accounts for the formal dependency of tone on the position of stress. Most importantly, it is supported by phonological activity, because it allows for an elegant compositional analysis of morphophonological alternations involving prosodic contrasts.

## 4.3 Choguita Rarámuri

Choguita Rarámuri is a variety of Rarámuri (Uto-Aztecan, also known as Tarahumara) spoken in Choguita, in the state of Chihuahua, Mexico. 234 people live in Choguita, and the majority of them are native Rarámuri speakers (Caballero 2008: 4–7). Caballero (2008) provides an extensive description and analysis of many issues in the phonology and morphology of the language, and Caballero (2011) and Caballero and Carroll (to appear) look more closely at the word prosody system. The data discussed in this section come primarily from Caballero and Carroll's (to appear) description of the language.<sup>7</sup>

Choguita Rarámuri is particularly relevant to the present study because of the complexity of its prosodic system. Typologically, the language is similar to Goizueta Basque in that stress is positionally contrastive within the word and tone is contrastive only within that stressed position. However, the language has three contrastive tones (versus Goizueta Basque's two), and shows a number of alternations in both tone and stress. For this reason, Choguita Rarámuri provides a more rigorous testing ground for the present model of word prosodic contrasts.

In this section, I will first present the key facts of the prosodic system. I will then propose an analysis of Choguita Rarámuri word prosody in which stress is represented by  $\pi$ -nodes and tone

---

<sup>7</sup>Because I cite a pre-publication version of Caballero and Carroll (to appear), I do not give page numbers.

is represented by binary contrastive features hosted by those  $\pi$ -nodes. I will demonstrate how, assuming that the language can employ both floating  $\pi$ -nodes and floating prosodic features, the alternations observed in the interaction between a number of classes of roots and suffixes can be accounted for in a straightforward compositional morphological analysis. This requires only the concatenation of phonological material and the application of repair rules, with no morphophonological rules referring to specific morphemes.

### 4.3.1 Stress patterns

Choguita Rarámuri uses both a stress contrast and a tone contrast, and although the two contrasts show a dependency relationship (tone is contrastive only in stressed positions), stress and tone can be the targets of independent alternations. I will begin by describing the patterns of stress assignment, without giving consideration to the realization of the tonal contrast.

Stress in Choguita Rarámuri may fall on any of the first three syllables of the word (Caballero 2011; Caballero and Carroll to appear). The position of stress is determined lexically, and many pairs are distinguished by the position of stress, as illustrated in (24).<sup>8</sup> The primary phonetic correlate of stressed position is increased duration, along with augmentation of onset consonants (through aspiration) and phonetic (or in some cases, phonological) vowel reduction in unstressed positions (Caballero 2008, 2011). There is no evidence of secondary stress.

(24) *Pairs differing in stress* (Caballero and Carroll to appear)

- |    |        |            |        |                |
|----|--------|------------|--------|----------------|
| a. | 'sawa  | 'smell'    | sa'wa  | 'leaf'         |
| b. | 'kotʃi | 'pig'      | ko'tʃi | 'dog'          |
| c. | 'kori  | 'visit'    | ko'ri  | 'chili pepper' |
| d. | 'misa  | 'mass'     | mi'sa  | 'crush'        |
| e. | 'nowi  | 'have son' | no'wi  | 'maggot'       |

<sup>8</sup>In the transcriptions given here, the alveolo-palatal affricate has been changed to /tʃ/, with no tie bar, contrary to the notation used by Caballero and Carroll (to appear). This has no bearing on the present discussion of prosodic contrasts. Glosses in this section use the following abbreviations: cond. = conditional fut. = future, imp. = imperative, imperf. = imperfective, pl. = plural, pst = past, sg. = singular, tr. = transitive.

The location of stress is not simply fixed according to a lexical specification, however; stress alternations occur depending on phonological and morphological factors. In Caballero and Carroll's (to appear) description, the location of stress can be determined based on the morphological classes of the root and the suffix. Roots can be either "unstressed" or "stressed", and suffixes can be either "stress-neutral" or "stress-shifting". Stressed roots are lexically accented in a particular position, and stress is realized in this position on the root across all paradigms. Unstressed roots, on the other hand, are assigned stress depending on the class of the suffix. Bare stems and forms with stress-neutral suffixes receive default stress on the second syllable of the root, or the only syllable of the root if it is monosyllabic; stress-neutral suffixes cannot bear stress.

When unstressed roots are followed by stress-shifting suffixes, stress is assigned to the third syllable of the *word*. For a bisyllabic root, this means that stress will fall on the suffix (which forms the word's third syllable). For a trisyllabic root, stress falls instead on the final syllable of the root. Thus it is not the case that a stress-shifting suffix is necessarily a stress-bearing suffix. In bisyllabic words, stress falls on the second (final) syllable. These patterns are illustrated in (25). For stressed roots, lexically accented positions are indicated by underlining the relevant vowel.

(25) *Morphologically conditioned stress alternations* (Caballero and Carroll to appear)

	Root	Bare stem (prs)	Stress-neutral (pst)	Stress-shifting (cond)
Unstressed	ru 'say'	'ru	'ru-li	ru-'sa
	tʃapi 'grab'	tʃa'pi	tʃa'pi-li	tʃapi-'sa
	ruruwa 'throw liquid'	ru'ruwa	ru'ruwa-li	ruru'wa-sa
Stressed	<u>su</u> 'sew'	'su	'su-li	'su-sa
	pu <u>t</u> ʃi 'blow'	'putʃi	'putʃi-li	'putʃi-sa
	se <u>m</u> e 'play violin'	se'me	se'me-li	se'me-sa
	hu <u>m</u> isi 'run away (pl.)'	'humisi	'humisi-li	'humisi-sa
	asi <u>s</u> i 'get up'	a'sisi	a'sisi-li	a'sisi-sa
	bi <u>n</u> ihi 'accuse'	bini'hi	bini'hi-li	bini'hi-sa



The role played by stress-shifting versus stress-neutral suffixes is limited by the morphology: only the class of the suffix closest to the root is relevant to stress assignment. This is illustrated in (26) by the unstressed root /suku/ ‘scratch’. In (26a) and (26b), the first suffix is stress-shifting and the class of the second suffix (shifting or neutral) is irrelevant—stress is realized on the third syllable of the word in both cases. In (26c) and (26d), the first suffix is stress-neutral, and the class of the second suffix is irrelevant—both have stress realized on the second syllable of the root, the default position.

(26) *Multiple suffixation* (Caballero and Carroll to appear)

- |    |                                 |                         |
|----|---------------------------------|-------------------------|
| a. | suku-'nale-sa                   | V + shifting + shifting |
|    | scratch-desid-cond              |                         |
|    | ‘If s/he would want to scratch’ |                         |
| b. | suku-'nale-ki                   | V + shifting + neutral  |
|    | scratch-desid-pst.1             |                         |
|    | ‘I wanted to scratch’           |                         |
| c. | su'ku-si-ma                     | V + neutral + shifting  |
|    | scratch-mot-fut.sg              |                         |
|    | ‘S/he will go along scratching’ |                         |
| d. | su'ku-ri-li                     | V + neutral + neutral   |
|    | scratch-cause-pst               |                         |
|    | ‘S/he made it scratch’          |                         |

Several things are clear from the stress data discussed thus far: (i) Choguita Rarámuri has lexical accent, because the position of stress can be lexically marked, (ii) the position of stress can be placed by a default algorithm, and (iii) the placement of stress is sensitive to morphological boundaries, since default stress can only be realized within the root and only the suffix closest to the root can affect the placement of stress.

I summarize the assignment of stress as the algorithm in (27), where “accent” refers to positional marking of lexical stress (in stressed roots), and “stress” refers to the actual surface real-

ization of increased duration.

(27) *Choguita Rarámuri stress assignment algorithm*

1. If root is accented:
  - (a) Assign stress to accented position
  - (b) End algorithm
2. If root is adjacent to stress-shifting suffix:
  - (a) Assign stress to third (else rightmost) syllable of **word**
  - (b) End algorithm
3. Else:
  - (a) Assign stress to second (else rightmost) syllable of **root**
  - (b) End algorithm

### 4.3.2 Tone patterns

I will now consider the data concerning tonal contrasts. Stressed positions in Choguita Rarámuri show three different contrastive tones, which Caballero and Carroll (to appear) refer to as “HL”, “M”, and “L”. Note that HL refers to a single phonological category, not a tonal cluster of H and L. Phonetically, HL involves a fall in pitch, while M and L “have nearly constant pitch over the course of the vowel” (Caballero and Carroll to appear). Caballero and Carroll schematically notate M tone in all unstressed positions.

Tone, like stress, is lexically contrastive, as illustrated by the sets in (28). Stressed positions, where tone is contrastive, are marked with '. The stressed vowel is marked with an acute accent for the HL category, a grave accent for the L category, and no diacritic for the M category.

	First-syllable stress			Second-syllable stress			Third-syllable stress		
3- $\sigma$	'HL.M.M	* <b><i>L.M.M</i></b>	* <b><i>M.M.M</i></b>	M.'HL.M	M.'L.M	M.'M.M	M.M.'HL	M.M.'L	M.M.'M
2- $\sigma$	'HL.M	'L.M	'M.M	M.'HL	M.'L	M.'M	—	—	—
1- $\sigma$	'HL	'L	'M	—	—	—	—	—	—

Table 4.1: Attested and unattested tonal melodies by root type and stress position (Caballero and Carroll to appear)

(28) *Tonal minimal pairs* (Caballero and Carroll to appear)

- a. [HL] 'tó 'bury' [L] 'tò 'take'
- b. [HL] 'mé 'win' [L] 'mè 'mescal'
- c. [M] 'pa 'throw' [L] 'pà 'bring'
- d. [M.HL] ko'lí 'chili pepper' [M.L] ko'li 'spatial root'
- e. [M.HL] i'sí 'urinate' [M.L] i'sì 'do'
- f. [M.M] ni?'wi 'to be lightning' [M.L] ni?'wì 'have a wedding'
- g. [HL.M] 'nótʃa 'pretentious' [L.M] 'nòtʃa 'hard working'

In trisyllables, stress can appear in any position, and with any of the three tone categories:

(29) *Prosodic contrasts in trisyllables* (Caballero and Carroll to appear)

- a. [M.M.HL] napa'bú 'get together'
- b. [M.M.L] siki'rè 'cut'
- c. [M.L.M] na?'sòwa 'stir'
- d. [M.HL.M] re?'étʃa 'play with food'
- e. [HL.M.M] 'húmisi 'run away (pl.)'

In fact, virtually all possible combinations of stress positions and tones are attested in Choguita Rarámuri roots, as illustrated in Table 4.1. The only exceptions are trisyllabic roots with initial stress and L or M tone (written with bold italic), which Caballero and Carroll (to appear) consider an accidental gap. With stressed roots (those in which accent is lexically specified), tone generally

remains fixed along with the position of stress. This is the case whether a stress-neutral or stress-shifting suffix is attached:

(30) *Fixed prosody on stressed roots* (Caballero and Carroll to appear)

	<i>Bare root</i>	<i>Stress-neutral suffix (pst)</i>	<i>Stress-shifting suffix (fut.sg)</i>
a.	'tó 'bury'	'tó-li	'tó-ma
b.	na?'pó 'weed'	na?'pó-li	na?'pó-ma
c.	'pa 'throw'	'pa-li	'pa-ma
d.	mu'ru 'bundle'	mu'ru-li	mu'ru-ma
e.	'nè 'use'	'nè-li	'nè-ma
f.	ni'wì 'make a wedding'	ni'wì-li	ni'wì-ma

For unstressed roots, the tones assigned may differ as the surface position of stress alternates. Tone patterns for unstressed roots are illustrated in (31). When an unstressed root appears alone or with a stress-neutral suffix, it receives default stress on the second (or only) syllable of the root. Although the position of stress is default, the tone realized in stressed position is not; unstressed roots are lexically specified for tone.<sup>9</sup> When a stress-shifting suffix is applied, however, the pattern is slightly more complicated. Shifted stress always falls on the third (or final) syllable of the word. With roots of fewer than three syllables, stress falls on the suffix, as in (31d-f). The tone realized on these suffixes is a lexical property of the suffixes. If shifted stress falls on the root, as seen with the three-syllable roots in (31a-c), the stressed syllable receives HL tone. This occurs regardless of the lexical tones associated with the unstressed root or the stress-shifting suffixes (Caballero and Carroll to appear):

<sup>9</sup>Caballero and Carroll (to appear) note that the lexical tone associated with an unstressed root is always M or L; they know of no unstressed roots with lexical HL tone.

(31) *Tones on unstressed roots* (Caballero and Carroll to appear)

	<i>Bare root</i>	<i>Stress-neutral suffix (pst)</i>	<i>Stress-shifting suffix</i>
a.	ro'sowa 'cough'	ro'sowa-li	roso'wá-ma (fut.sg)
b.	na?'sòwa 'stir'	na?'sòwa-li	na?so'wá-ma (fut.sg)
c.	ra'ʔitʃa 'speak'	ra'ʔitʃa-li	raʔi'tʃá-ma (fut.sg)
d.	re'wa 'see'	re'wa-li	riwi-'sá (cond)
e.	'tò 'take'	'tò-li	to-'ká (imp.sg)
f.	'tò 'take'	'tò-li	to-'sì (imp.pl)

While the tone alternations in (31) always coincide with a shift in the position of stress, there are also cases in Choguita Rarámuri where tone alternates without the position of stress changing. As noted above, stressed roots generally retain the position of stress and the same tone throughout all morphological contexts. However, some stressed roots with L tone are lexical exceptions to this. When these exceptional roots are combined with stress-shifting suffixes, the location of stress remains the same, but a change in tone takes place from L to HL:

(32) *Tone alternating roots* (Caballero and Carroll to appear)

	<i>Bare root</i>	<i>Stress-neutral (pst)</i>	<i>Stress-shifting (fut.sg)</i>
a.	'pà 'bring'	'pà-li	'pá-ma
b.	'ʔà 'give'	'ʔà-li	'ʔá-ma
c.	na'wà 'arrive'	na'wà-li	na'wá-ma
d.	a'tʃà 'seat (tr.)'	a'tʃà-li	a'tʃá-ma
e.	bi?'wà 'clean (tr.)'	bi?'wà-li	bi?'wá-ma
f.	ni'wà 'make'	ni'wà-li	ni'wá-ma

Choguita Rarámuri also has in-place tone alternations that affect unstressed roots. Consider the data in (33). Bare roots with lexical M tone receive default stress on the second syllable of the root, which realizes the lexical tone. As expected, the root remains unchanged from its bare form when the stress-neutral past tense suffix /-li/ is added. The imperfective suffix /-i/, on the other

hand, does not leave the root unaffected; while it does not change the default position of stress (it is stress-neutral), it causes the tone realized there to alternate to L.

When the stress-shifting conditional suffix /-sa/ is applied to the same roots, stress shifts to the third syllable of the word, falling on the suffix. The tone realized is then HL, the lexical tone of the suffix. In this way, it is possible for forms derived from the same root to contain all three contrastive tones.

(33) *Imperfective tone alternations* (Caballero and Carroll to appear)

	<i>Bare root</i>	<i>Stress-neutral</i>		<i>Stress-shifting</i>
		Past	Imperfective	Conditional
a.	a'wi 'dance'	a'wi-li	a'wì-i	awi-'sá
b.	ri'wa 'find'	ri'wa-li	ri'wà-i	riwa-'sá
c.	ra'ra 'buy'	ra'ra-li	ra'rà-i	rara-'sá
d.	tʃu'ku 'crawl'	tʃu'ku-li	tʃu'kù-i	tʃuku-'sá

Finally, it is also possible for morphological marking on unstressed roots to be expressed by tone alternation alone. In Choguita Rarámuri, the imperative singular can be indicated in two ways. The first is a stress-neutral segmental suffix /-ka/, which leaves default stress and lexical tone on the root. The second, however, involves no segmental material at all. When this imperative allomorph is combined with stressed HL tone roots, it is only the tone of the stressed position that changes, being realized as L tone. This is exemplified in (34).

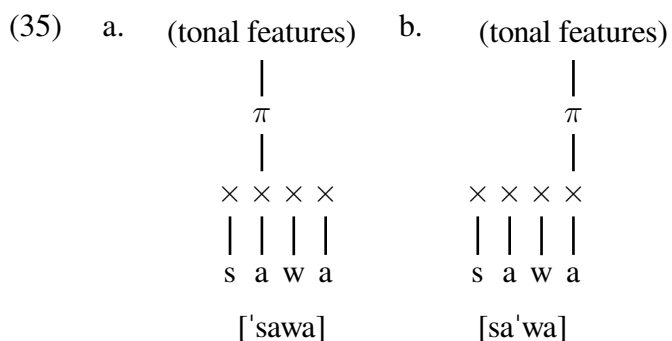
(34) *Suffix and L tone imperative marking* (Caballero and Carroll to appear)

	<i>Bare stem</i>	<i>Imperative suffix</i>	<i>Imperative L tone</i>
a.	ma'tó	ma'tó-ka	ma'tò
	'to carry on the shoulders'		
b.	suʔu'ní	suʔu'ní-sa	suʔu'nì
	'to finish doing something'		
c.	ti'só	ti'só-ka	ti'sò
	'to walk with a cane'		

### 4.3.3 A representational account

I will now consider how the prosodic system of Choguita Rarámuri can be represented in the present model.<sup>10</sup> The distribution of stress and tone in Choguita Rarámuri is quite similar to that seen in Goizueta Basque, as discussed in the preceding section; stress is contrastive for its position, while one of several contrastive tones is realized at the position of stress. In other words, a paradigmatic tone contrast is realized on top of a syntagmatic stress contrast. Consequently, the representations I propose for Choguita Rarámuri prosody are similar to those given for Goizueta Basque.

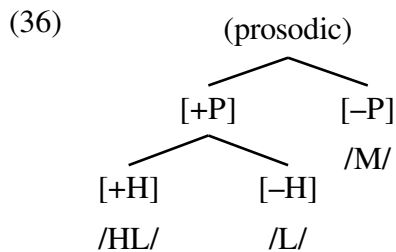
The stress contrast is represented by the linking of a  $\pi$ -node to that position, and the tone realized there is represented by one or more contrastive prosodic features hosted by the  $\pi$ -node. Abstracting away, for the moment, from the processes that determine the placement of  $\pi$ -nodes, the structure of output representations is illustrated in (35). A word with initial stress has a  $\pi$ -node linked to the X-slot of the initial vowel, a word with second-syllable stress has a  $\pi$ -node linked to the X-slot of the second vowel, and so on:



The next issue concerns the  $\pi$ -node features that represent the tone contrast. Choguita Rarámuri has three contrastive tones (HL, M, and L), and each of them can be specified lexically. Within the present model, where features dividing the inventory of  $\pi$ -nodes are assigned via the successive division algorithm, a three-member inventory must be specified for two different contrastive features, where one member is specified for one feature and the other two members are

<sup>10</sup>While Caballero (2011) assumes foot-based representations of stress, she does not address tone at all. Caballero and Carroll (to appear) focus on the description and phonetic analysis of Choguita Rarámuri prosody, but do not develop a detailed system of phonological representations.

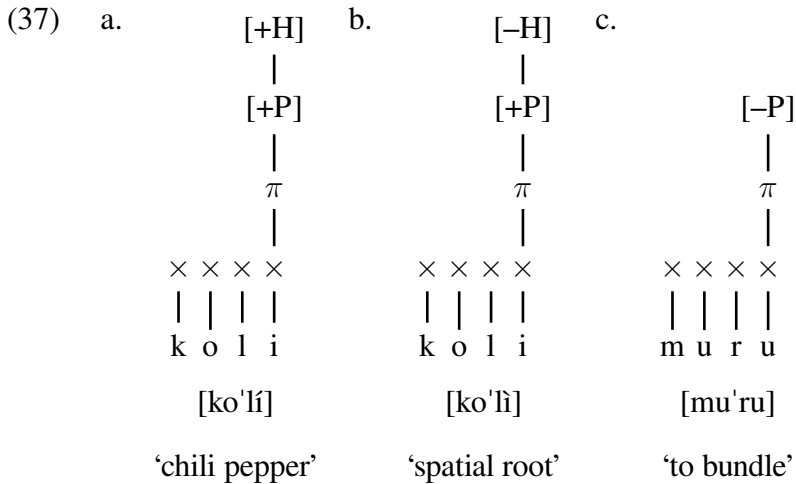
specified for both features. I propose that the first cut is made by a feature  $[\pm(\text{P})\text{eripheral}]$ , which separates M from the rest of the prosodic inventory by specifying it  $[-\text{P}]$ . This keeps M tone from deviating from neutral reference pitch, that is, it prevents it from being realized at the periphery of the pitch range. The other two tones, which are realized at the high and low ends of the pitch range, are differentiated by the feature  $[\pm\text{H}]$ ; HL receives  $[\text{+H}]$ , while L receives  $[-\text{H}]$ . These feature specifications result in the prosodic contrastive hierarchy in (36):



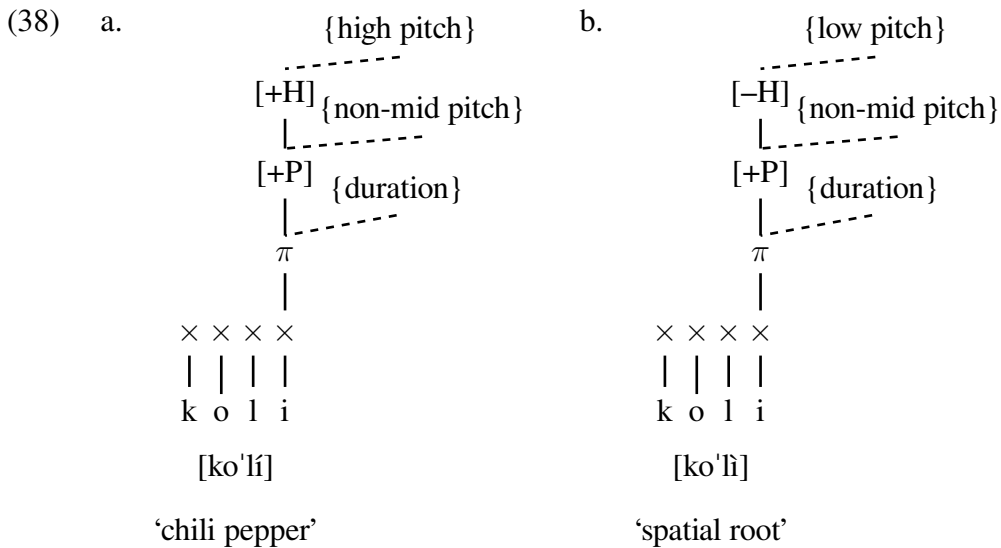
Note that this analysis abstracts away from the falling pitch of HL, treating it as a single category. I retain the label “HL” for this category in order to avoid needless deviation from Caballero and Carroll’s (to appear) notation. The phonetic reflex of the feature  $[\pm\text{H}]$  is simply high or low pitch within the non-mid space carved out by  $[\text{+P}]$ . HL, with  $[\text{+H}]$ , hits a peak above neutral reference pitch (starting in the high range of the pitch periphery), and L, with  $[-\text{H}]$ , remains below neutral reference pitch.

To represent word prosodic contrasts, the features specified in (36) form a feature chain hosted by the  $\pi$ -node, linking in order from the hierarchy’s root node to the relevant toneme’s terminal node. This applies as in (37):



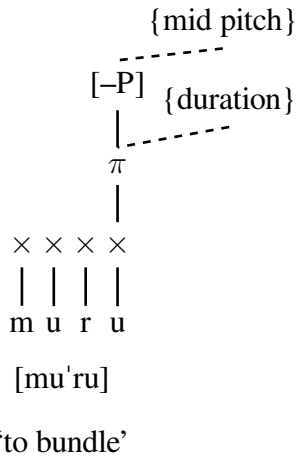


The relationships between the contrastive representations and the surface phonetics are as follows. In Choguita Rarámuri, the phonetic realization of the  $\pi$ -node is increased duration, that is, the primary correlate of stress in the language. This {duration} is applied to any position to which a  $\pi$ -node is linked. The specification of  $[+P]$  ensures that the onset of the tone has {non-mid pitch}. In (38a), the specification  $[+H]$  ensures that the tone's onset has {high pitch}, while in (38b),  $[-H]$  ensures that its onset has {low pitch}.



In (39), the stressed position shows M tone. Here, the only feature specification hosted by the  $\pi$ -node is  $[-P]$ . This simply carves out an area of {mid pitch} within the range of neutral pitch. No specification of pitch features (such as  $[\pm H]$ ) is present, so the onset of the M tone does not receive any instructions in phonetic implementation.

(39)

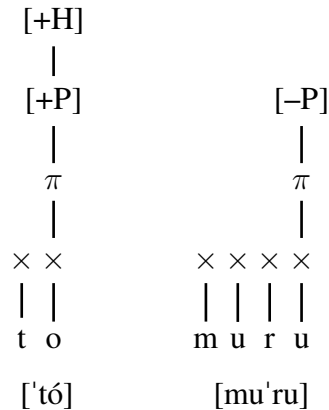


#### 4.3.4 Morphemic analysis

Based on the system of contrastive representations outlined above, I will now demonstrate how it is possible to derive the word prosodic alternations observed in Choguita Rarámuri using a compositional morphological analysis. There are no previous analyses that capture the formal relationship between stress and tone with which the analysis given here can be compared. Caballero (2011) considers only patterns of stress alternations, proposing that the behaviour of different classes of morphemes can be captured using cophologies. Caballero and Carroll (to appear), on the other hand, do not give an explicit account of how prosodic alternations are derived, offering a more theory-neutral description of the language.

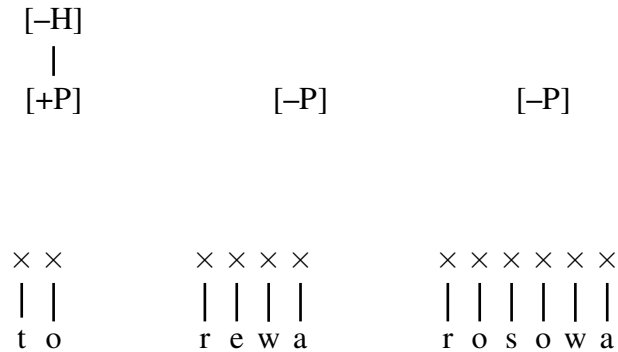
I will begin with the most clear-cut cases, namely stressed roots with fixed lexical tone. These roots undergo no prosodic alternations, regardless of phonological or morphological context. I assume that such roots are lexically specified for fixed stress and tone; the position of stress is marked with a lexical accent taking the form of a  $\pi$ -node linked to the accented vowel's X-slot, and tone is specified by prosodic contrastive features associated underlyingly with that  $\pi$ -node. Morphemes of this type are illustrated in (40). When these roots appear in their unsuffixed forms, they are phonetically interpreted according to the conventions laid out above, giving rise to their surface pronunciations.

(40) a. 'bury' b. 'bundle'



Unstressed roots do not bear lexical accent, and accordingly their underlying forms do not contain  $\pi$ -nodes. Although unstressed roots contain no lexical information about the location of stress, they do have lexically specified tone. I assume that these tones are stored as floating prosodic feature chains in their lexical entries:

(41) a. 'take' b. 'see' c. 'cough'



Roots like these receive specification for accent after application of the default  $\pi$ -node insertion process in (42). I state this process as a descriptive rule which isolates the default ("Else") condition of the stress assignment algorithm given in (27) at the end of Section 4.3.1 above. I remain agnostic as to the exact principles causing post-initial placement of stress. What is important is the nature of the representations derived from default stress placement, that is, a  $\pi$ -node linked to the word's second vocalic X-slot.

(42) *Default  $\pi$ -node insertion ( $\pi$ -ins)*

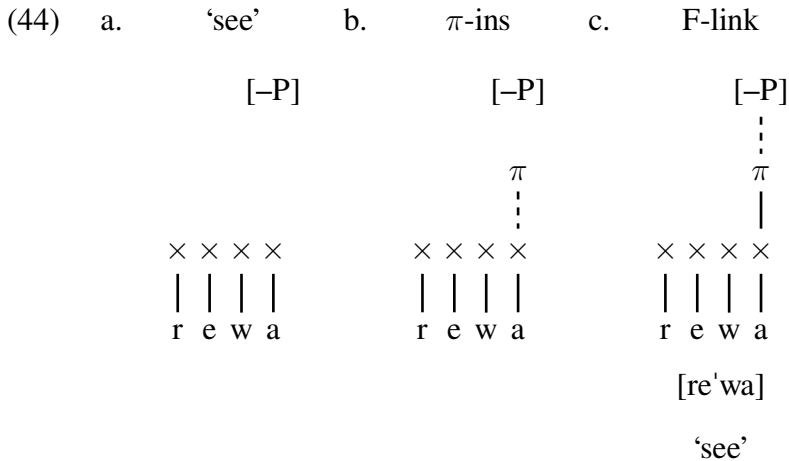
If the word contains no  $\pi$ -node, insert a  $\pi$ -node and link it to the second (or only) vocalic X-slot of the root

I further assume the process in (43), which links the floating prosodic features of the root to the newly inserted  $\pi$ -node. This rule applies from right to left, for reasons which will become clear below.

(43) *Linking of prosodic features (F-link)*

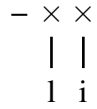
Link the rightmost floating prosodic feature to a docked featureless  $\pi$ -node

To derive the unsuffixed form of the root ‘see’, the  $\pi$ -nodeless underlying root receives a  $\pi$ -node on its second vocalic X-slot due to application of  $\pi$ -ins in (44b). The stressed position marked by this  $\pi$ -node then receives its tonal specification when the root’s floating [-P] feature is linked by application of F-link in (44c). This results in an output form with default post-initial stress and M tone:

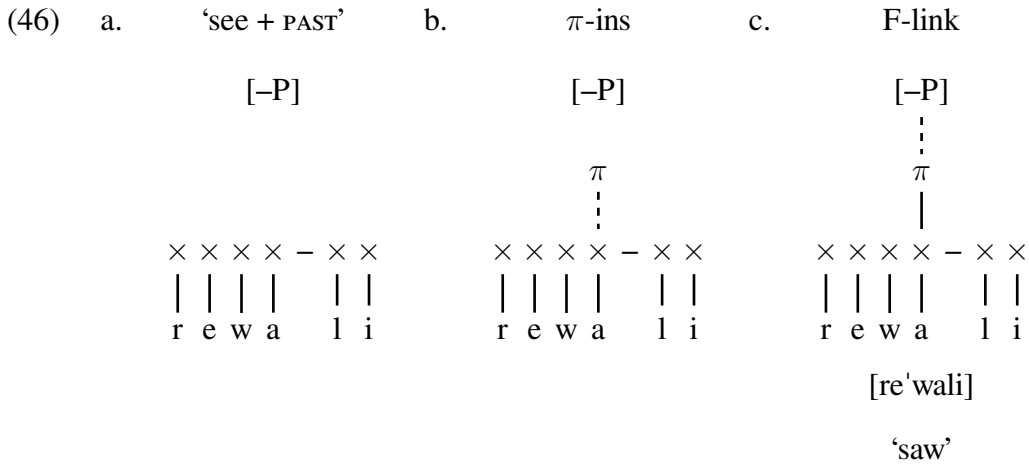


Lexical entries for stress-neutral suffixes contain no information about stress and generally contain no specification for tone; they contribute only segmental information (segmental contrastive features and their associated X-slots):

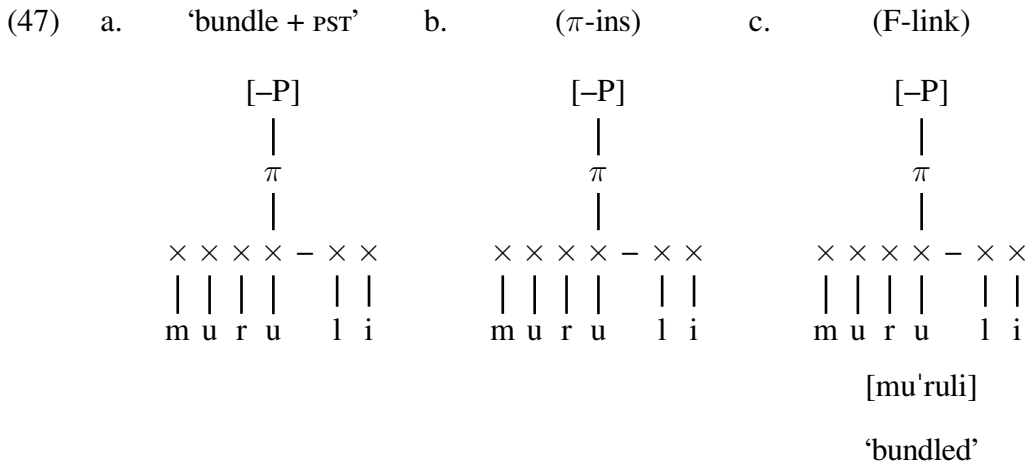
(45) 'PAST'



The presence of such a stress-neutral suffix on an unstressed root does not affect the derivation of prosodic contrasts. Application of  $\pi$ -ins still puts default stress on the second (or only) syllable of the root, and F-link still realizes the root's lexical tone contrast in stressed position:



When a stressed root is interpreted, it enters the derivation with specifications for stress and tone. Since there is no need to insert a  $\pi$ -node and no floating prosodic features to link, prosodic representations are not manipulated by any processes. A stress-neutral suffix, if present, likewise has no effect on stress or tone:



Let us now consider the suffixes that trigger prosodic alternations, the most common of which are the stress-shifting suffixes. I propose that the stress-shifting properties of these suffixes are not due to their triggering certain morphological rules, but rather derived directly from phonological information contained in their lexical entries. This can be accomplished by specifying floating  $\pi$ -nodes in the underlying forms of stress-shifting suffixes. When stress falls on stress-shifting suffixes, a tone lexically associated with the suffix is realized. This information is encoded as floating prosodic features also present in the underlying form of the suffix. Examples of underlying forms for stress-shifting suffixes are given in (48).

(48) a. 'FUT.SG'	b. 'COND'	c. 'IMP.PL'
[+H]	[+H]	[-H]
[+P]	[+P]	[+P]
$\pi$	$\pi$	$\pi$
- × ×	- × ×	- × ×
m a	s a	s i

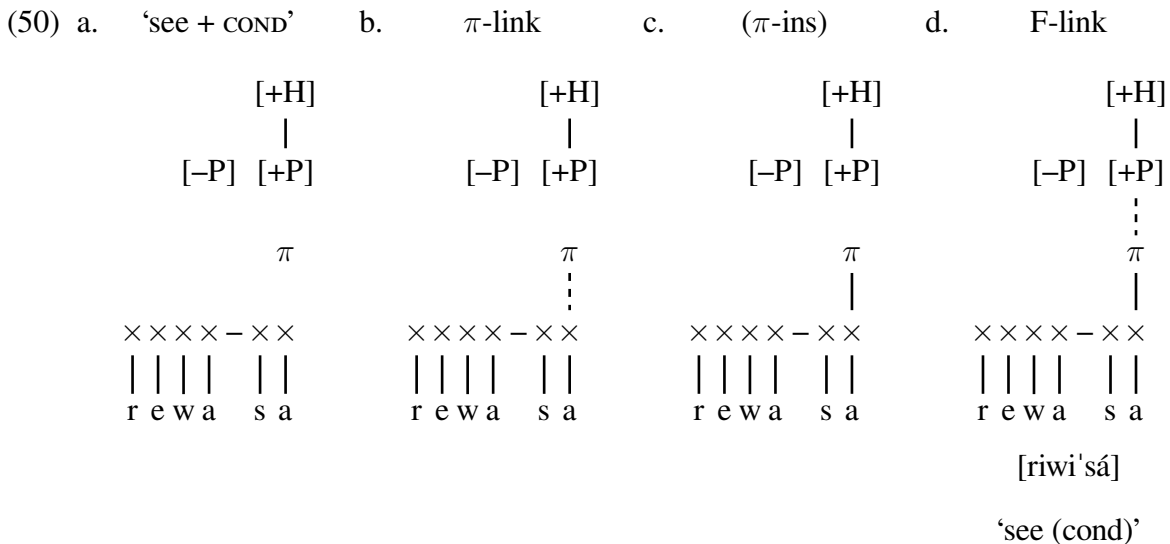
The alternation observed with stress-shifting suffixes can be captured if processes which link  $\pi$ -nodes to X-slots can distinguish between the presence and absence of floating  $\pi$ -nodes in the input. While the default  $\pi$ -node insertion rule in (42) above targets the second possible accent-bearing position in the root, I propose that floating  $\pi$ -nodes which are present in the input (rather than being inserted) are operated on by the  $\pi$ -node linking process in (49), which targets the *third* vocalic X-slot in the *word*. While  $\pi$ -link does not distinguish between the root and the suffix as locations for linking  $\pi$ -nodes, it must link floating  $\pi$ -nodes only within the suffix which is closest to the root.

As with  $\pi$ -ins, I remain agnostic as to the exact mechanisms involved in locating the target of  $\pi$ -node linking, and take the same position with respect to the manner in which morphological domains are defined. The concern of the present study is the contrastive phonological representations that these processes output.

(49) *Suffix  $\pi$ -node linking ( $\pi$ -link)*

If the leftmost suffix contains a floating  $\pi$ -node and there is no docked  $\pi$ -node in the word, link the suffixal  $\pi$ -node to the third vocalic X-slot of the word

The derivation of stress shift is demonstrated in (50). The  $\pi$ -node linking process applies first in (50b), blocking the application of  $\pi$ -node insertion in (50c). Unlike in the forms examined above, this word contains two sets of floating prosodic features—one from the root and one from the suffix. In (50d), F-link, searching from right to left, links the prosodic features of the suffix to the  $\pi$ -node. This results in a trisyllabic word with stress on the suffix and HL tone:<sup>11</sup>



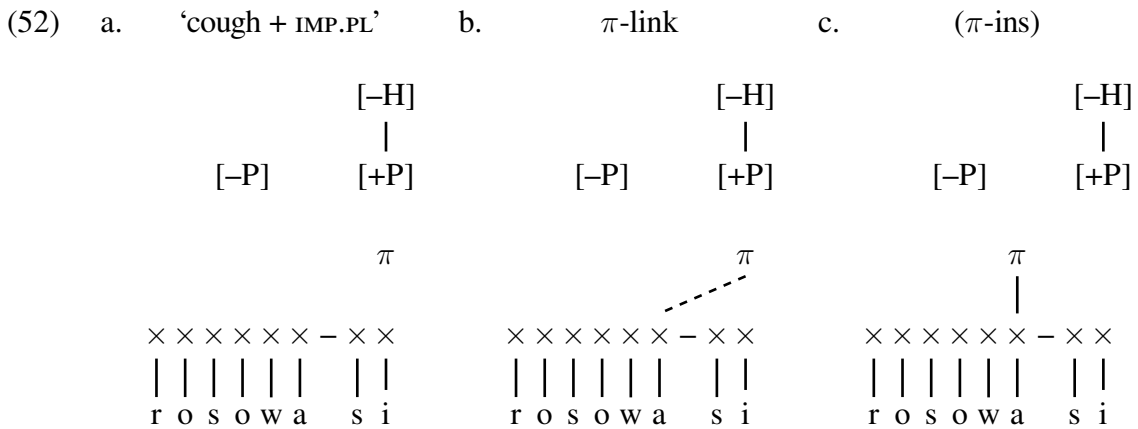
This analysis does not account for the tone pattern seen in trisyllabic unstressed roots. Stress-shifting suffixes cause stress to be realized on the root, since the root is long enough that  $\pi$ -link places the  $\pi$ -node within the root. Recall that the tone realized in such positions is always HL, regardless of the lexical tones of the root or suffix. Consider the monosyllabic root ‘take’ in (51a), which has lexical L tone. When this root occurs with a stress-shifting suffix, the suffix bears stress and realizes its own lexical L tone. As seen in (51b), the unstressed root ‘cough’ has lexical M tone, which it realizes when unsuffixed or combined with a stress-neutral suffix. When a stress-shifting suffix is added, however, stress falls on the root’s final syllable. The tone realized in this

<sup>11</sup>Note that segmental changes have also applied in deriving the surface form; these do not interact with the relevant aspects of the prosodic system, and I do not address them here.

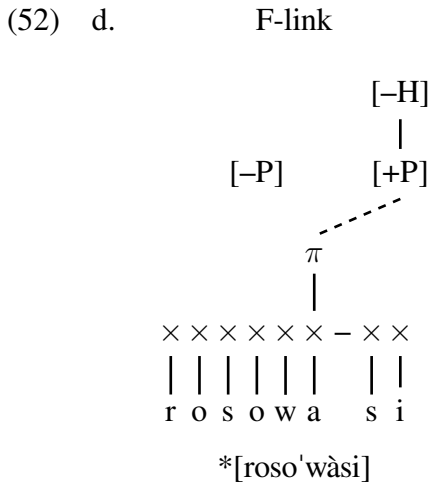
position is neither the root's lexical M tone nor the suffix's lexical L tone—instead, HL tone is realized (Caballero, p.c.).

- (51)      Bare root              Stress neutral (pst)      Stress-shifting L (imp.pl)
- a.    'tò 'take'              'tò-li                      to-'sì
- b.    ro'sowa 'cough'      ro'sowa-li              roso'wà-si

If the input form of 'cough (imp.pl)' is derived via the same processes outlined above, a wrong prediction is made. The suffix's lexical  $\pi$ -node will first be linked to the end of the root as in (52b). This allows the (attested) third-syllable stress and blocks default  $\pi$ -node insertion in (52c):



However, F-link will apply as in (52d), causing the suffix's features (which represent L tone) to link to the  $\pi$ -node. This wrongly realizes L tone on the stressed third syllable of the root, rather than HL tone:



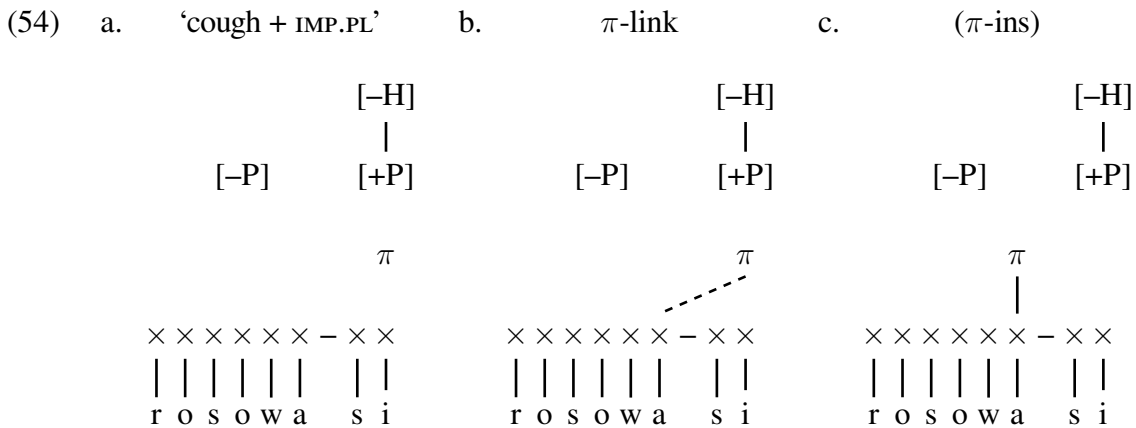


I propose that the curious tonal behaviour of trisyllabic stems is captured by the rule in (53), which ensures that any stressed third root syllable that is not specified for prosodic features receives HL tone by default. In other words, only stressed roots (those bearing lexical accent and tone specifications) can realize a tone other than HL on their third vocalic X-slots.

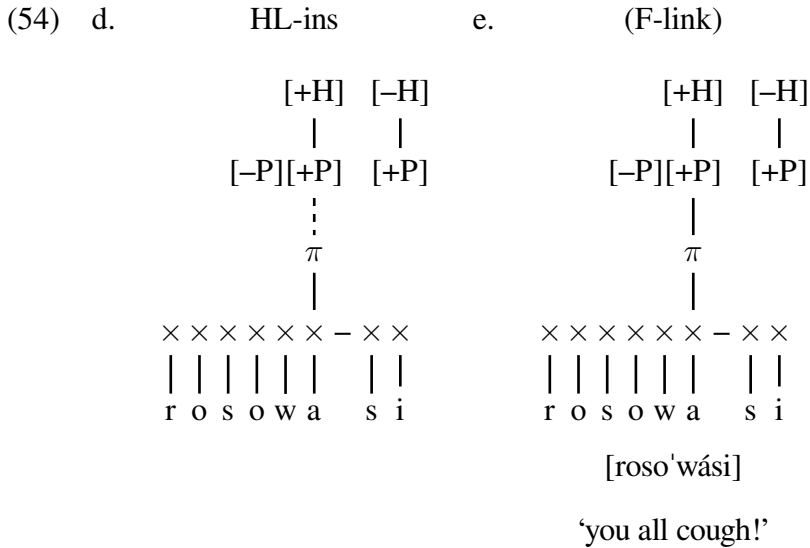
(53) *Root-final HL tone insertion (HL-ins)*

If a bare  $\pi$ -node is associated with the third vocalic X-slot of a root, insert a feature chain [+P, +H] and associate it with that  $\pi$ -node

The derivation begins the same as in (52). First, the suffix's floating  $\pi$ -node links to the root's third vocalic X-slot, blocking application of default  $\pi$ -node insertion:



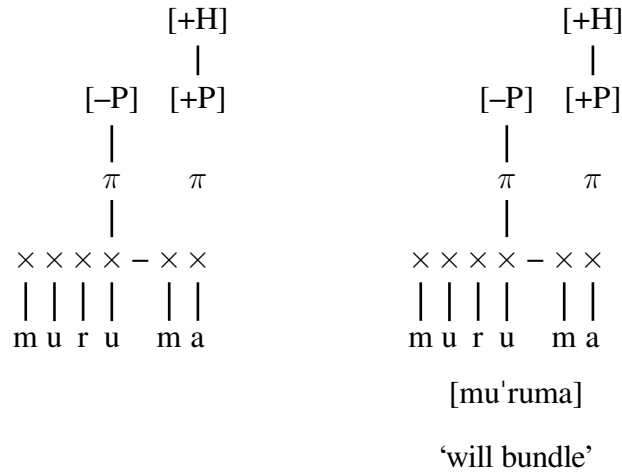
Before F-link can link the floating L tone to the  $\pi$ -node, however, default HL tone insertion creates and attaches the features for HL tone to the  $\pi$ -node. This is spelled out as the (correct) surface HL tone in root-final position:



### 4.3.5 In-place tone alternations

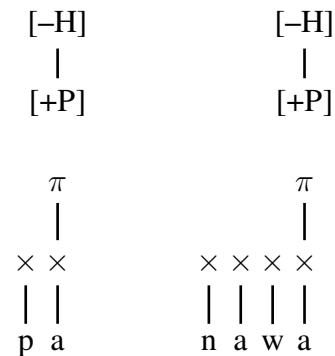
I will now turn to prosodic alternations involving a change in tone without a shift in the position of stress. Recall that stressed roots, which have an underlyingly linked  $\pi$ -node marking the position of lexical accent, usually keep the same prosodic contrasts across morphological contexts; the addition of a stress-shifting suffix causes neither a shift in stress nor a change in tone. This behaviour is predicted by the analysis outlined above, because none of the phonological processes can affect roots in which both  $\pi$ -nodes and prosodic features are already associated. This is demonstrated in (55). Since a  $\pi$ -node is already linked in the root, the suffix's floating  $\pi$ -node is not linked by  $\pi$ -link and a default  $\pi$ -node is not inserted by  $\pi$ -ins. Likewise, since the root's  $\pi$ -node already bears prosodic features, default features are not inserted by HL-ins, nor are the suffix's floating features linked to it by F-link. Thus the output realizes only the prosodic information which is associated in the lexical entry for the root.

(55) a. ‘bundle + FUT.SG’    b. ( $\pi$ -link,  $\pi$ -ins, HL-ins, F-link)

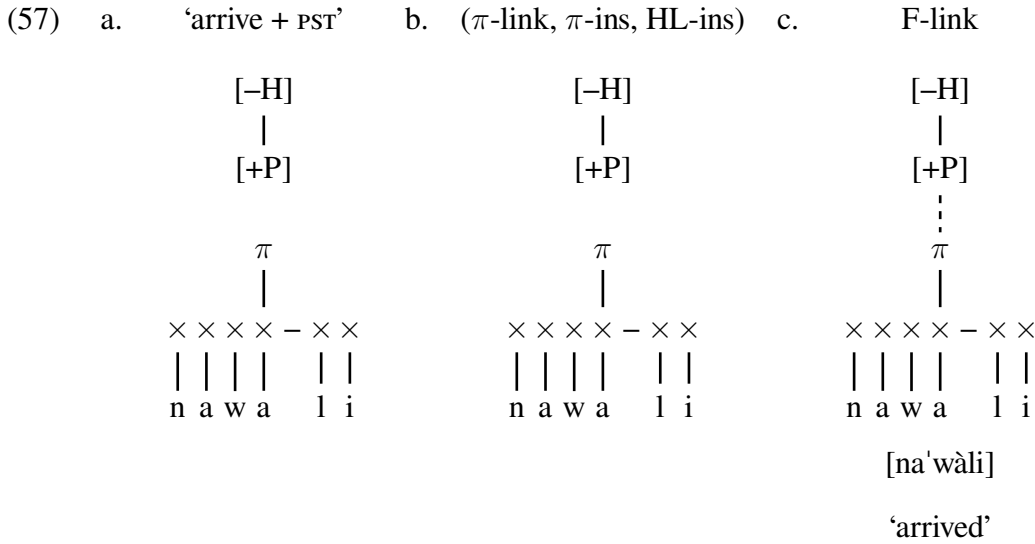


Some stressed roots behave differently, however. Caballero and Carroll (to appear) refer to these as “alternating L roots”; while the location of stress always remains the same—the position of stress is marked as lexical accent—the addition of stress-shifting suffixes triggers an alternation from L tone to HL tone at the lexical location of stress. The present model offers an elegant way of accounting for such roots using the same phonological processes outlined above. The only difference is that in alternating L roots, prosodic features are not linked to the  $\pi$ -node. Some examples are given in (56):

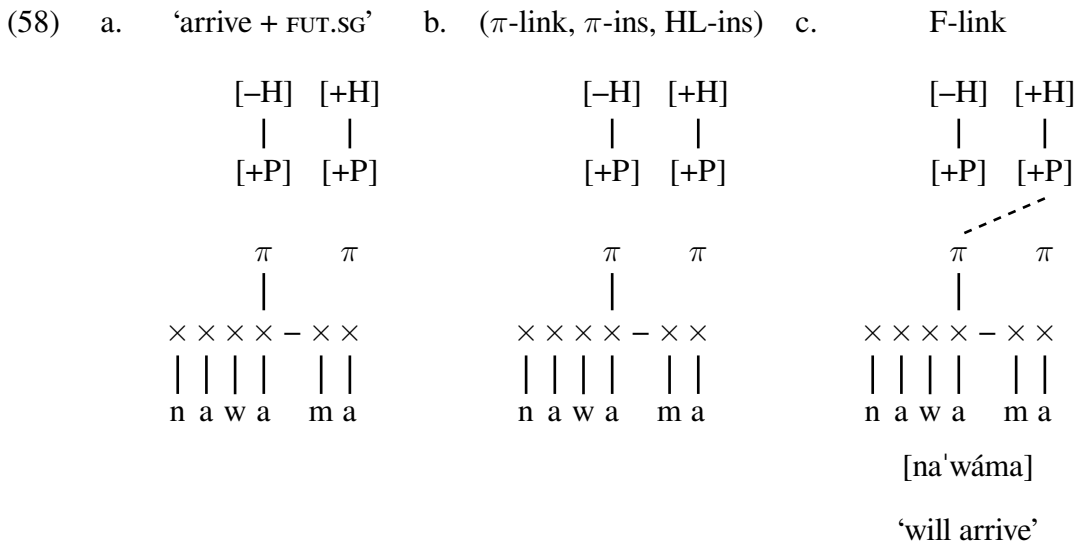
(56) a. ‘bring’    b. ‘arrive’



When these occur with stress-neutral suffixes, no tone alternation occurs, and the root’s lexical tone is realized at the location of lexical accent when F-link causes the floating prosodic features to link to the root’s  $\pi$ -node:

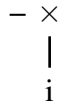
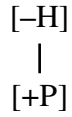


When alternating L roots are combined with stress-shifting suffixes, however, their distinct behaviour can be seen. F-link, applying from right to left, links the suffix’s prosodic feature chain to the root’s  $\pi$ -node, realizing the lexical tone of the suffix instead of that of the root:



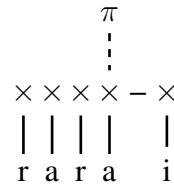
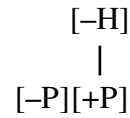
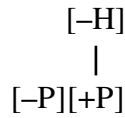
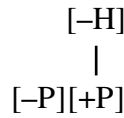
Finally, I will consider the in-place tone alternations affecting unstressed roots. Recall that the imperfective suffix leaves unstressed M tone roots with stress in the default position, but causes an alternation to L tone. This can be captured if the imperfective suffix contains segmental information and a floating feature chain representing L tone, but has no  $\pi$ -node:

(59) ‘IMPERF’

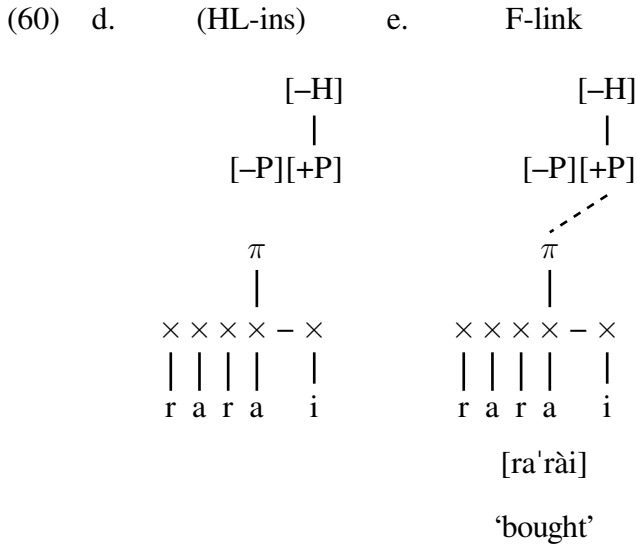


When this suffix is combined with an unstressed root, stress-shift is not triggered because there is no  $\pi$ -node present to be linked by  $\pi$ -link. Instead, one is inserted by  $\pi$ -ins, marking the position of default stress:

(60) a. ‘buy + IMPERF’    b. ( $\pi$ -link)    c.  $\pi$ -ins

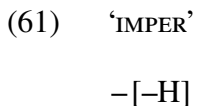


Then F-link applies, linking the rightmost prosodic features (those for L tone, floating on the suffix) to the root’s newly inserted  $\pi$ -node. This results in the realization of L tone rather than the root’s lexically specified M tone, along with the suffix’s segmental /-i/ marking:



The last tone alternation to be accounted for is the imperative L tone allomorph form, in which the imperative morpheme is realized as only an alternation from HL to L tone in the lexically accented position of stressed roots. Although we have seen a number of alternations that are triggered by floating prosodic material from suffixes linking to roots which lack associated features and/or  $\pi$ -nodes, the L tone imperative is complicated by the fact that it affects stressed roots. Stressed roots, with the exception of alternating L roots, have underlying forms in which all prosodic specifications are fully linked, and as a result they are not affected by the various linking and default insertion processes that normally result in prosodic alternations.

What is needed to account for the L tone imperative is a way to get the feature chain for L tone onto a  $\pi$ -node which is associated with an X-slot and lexically fully specified for features. This can be accomplished if the L tone imperative is a suffix consisting not of the L tone feature chain [+P, -H], but of just a floating [-H] feature:

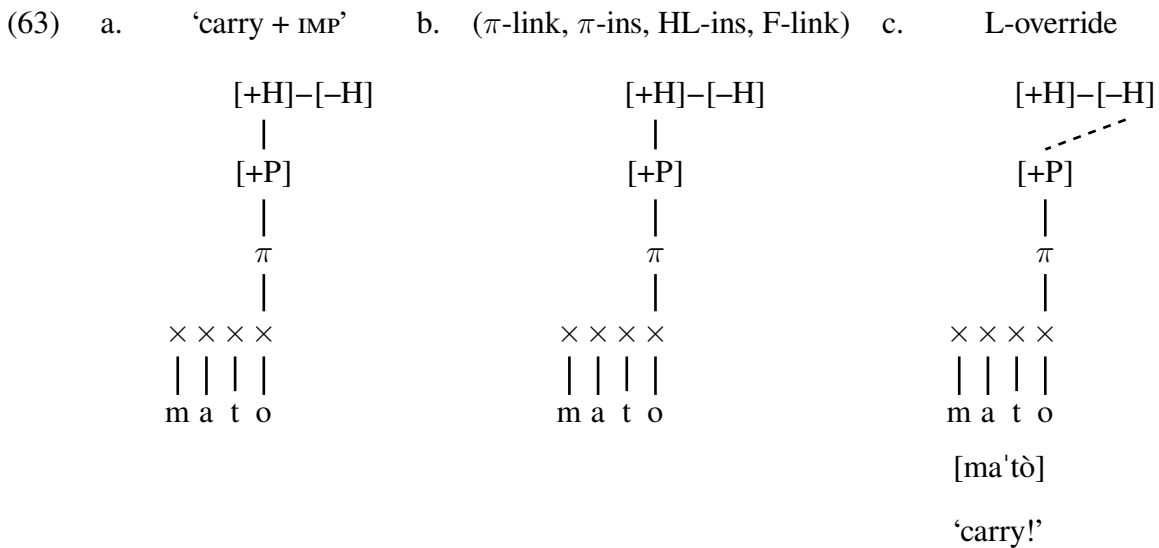


The presence of only a floating [-H] feature without an accompanying [+P] feature triggers the process in (62). This causes the [-H] feature to link to the [+P] feature to its left, replacing any existing specification for [ $\pm$ H] there.

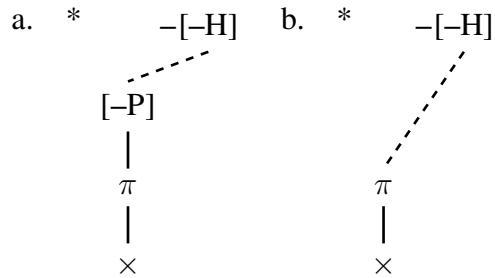
(62) *L tone feature replacement (L-override)*

Link a lone floating  $[-H]$  feature to a valid position to its left, delinking any specification for  $[\pm H]$  there

Because the  $\pi$ -node specifying root-final stress and the feature chain specifying HL tone are linked in the input, none of the prosodic linking or insertion processes apply. The environment for L-override is present, however, and it does apply, deriving L tone in the stressed position:



This analysis, in which the L tone imperative is derived from only a floating  $[-H]$  feature rather than the full feature chain representing L tone ( $[+P, -H]$ ), accounts for the fact that the alternation only affects roots with underlying HL tone and crucially not those with M tone. The “valid position to its left” to which floating  $[-H]$  links can only be a  $[+P]$  feature—the representations in (64) are not possible. (64a), in which  $[-H]$  attempts to link to an M tone’s  $[-P]$ , is ruled out because the feature chain  $[-P, -H]$  is not present in Choguítá Rarámuri’s contrastive hierarchy. In other words, the feature  $[\pm H]$  is not contrastive for  $[-P]$  autosegments in the language’s prosodic inventory, so  $[-P]$  is not a valid position for  $[\pm H]$  to link to. (64b) is ruled out for the same reason. If the suffix’s  $[-H]$  attempts to link to a  $\pi$ -node which has not received any feature specifications, it cannot alone cause that  $\pi$ -node to represent a well-formed member of the prosodic inventory—a  $\pi$ -node must bear a specification for  $[\pm P]$  as its first feature.

(64) *Ill-formed representations in Choguïta Rarámuri*

### 4.3.6 Exceptional stress alternations

There is, finally, a marginal class of roots that shows a different pattern of stress alternations from those seen above, exemplified in (65). These behave like unstressed roots in that the position of stress changes when combined with stress-shifting suffixes, but the positions in which stress appears are different from those of typical unstressed roots. These roots show default stress on the first (rather than expected second) syllable, and shifted stress on the second (rather than expected third) syllable.

(65) *Exceptional stress patterns* (Caballero, p.c.)

Root	Stress-neutral (pst)	Stress-shifting (fut.sg)
a. 'tʃùta 'sharpen'	'tʃùta-li	tʃu'tá-ma
b. 'soma 'wash one's head'	'soma-li	so'má-ma

In other words, stress in these roots shows a  $\sigma_1 \rightarrow \sigma_2$  alternation instead of the  $\sigma_2 \rightarrow \sigma_3$  alternation of typical unstressed roots. Furthermore, they undergo the alternation to HL tone when shifted stress falls on the root-final vocalic X-slot. This behaviour would not be unexpected if these roots covertly contained an additional syllable at their left edge; in such a case, stress would follow the typical pattern:



(66) *Exceptional stress as covert trisyllabicity*

	Root	Stress- neutral (pst)	Stress- shifting (fut.sg)
a.	V'tfùta 'sharpen'	V'tfùta-li	Vtju'tá-ma
b.	V'soma 'wash one's head'	V'soma-li	Vso'má-ma

I propose that such a covert third syllable is in fact present on these roots, as illustrated by the underlying forms in (67). The initial X-slot bears not a full set of segmental features, but a single [+V(ocalic)] feature. This means that syllable-counting processes identify it as a “vocalic X-slot”, but it is prevented from being overtly realized as a segment. I assume that this feature makes the first split in the Choguita Rarámuri segmental inventory, which means that it is possible for it to be the sole feature present in an underspecified feature chain, attaching directly to the segmental root node with no other features.<sup>12</sup>

(67)	a.	'sharpen'	b.	'wash'
		[-H]		
		[+P]		[-P]
		× × × × ×		× × × × ×
		t f u t a		s o m a
		[+V]		[+V]

In addition to capturing the synchronic facts, this analysis also makes sense diachronically, since these roots are derived historically from the truncation of initial syllables. However, the notion of covert extra syllables does not seem to be exploited elsewhere in the synchronic phonology, and there are only a handful of roots of this class in the language. Given the situation, we would expect them to show some instability in terms of the classes in which they pattern, and this

<sup>12</sup>Although not identical, this analysis bears resemblance to the concept of “catalexis” (Kiparsky 1991), which exploits prosodic material with no segmental content to achieve different metrical parsing. It is thus a kind of “logical counterpart of extrametricality” (Kager 1995a).





ical processes. Furthermore, it was done within an explicit and integrated model of accent–tone interaction.

This case study has shown that it is possible for a single language to show a large amount of the complexity predicted by the present model of word prosodic contrastive representations without exceeding the maximum level of complexity that it allows.

## 4.4 Commonalities of tonal accent languages

In this chapter I have applied the present model of contrastive prosodic representations to two tonal accent languages, Goizueta Basque and Choguita Rarámuri. These languages have much in common: both show a contrast in the position of surface stress which co-occurs with a contrast in tone, but in both languages the position in which tone is realized does not vary independently from the position of stress—a single position is syntagmatically contrastive for accent (realized as stress), and that position is paradigmatically contrastive for tone (realized as distinct pitch levels and/or contours). The present model predicts systems of this type because  $\pi$ -nodes can have phonetic realizations (in this case, stress) but at the same time can bear contrastive features with independent phonetic realizations (here, tone). When these coexist in a system where  $\pi$ -nodes themselves are culminative/obligatory, the result is a tonal accent system.

Goizueta Basque and Choguita Rarámuri are also similar in terms of their phonological activity. Both have productive morphophonological alternations in tone that apply independently of stress—a single position can retain its accent while changing its tonal category. Furthermore, both languages allow alternations in the position of stress. These processes are accounted for without morphological rules that change tonal categories or foot structure, but rather by positing suffixes that consist of floating phonological elements and roots with varyingly associated or floating  $\pi$ -nodes. By employing standard autosegmental notions such as linking and delinking, the data are captured in a phonological analysis whose only knowledge of morphology is morpheme boundaries. The languages differ only in the size of their autosegmental inventories: Goizueta Basque

has two different tonal categories while Choguita Rarámuri has three, and thus the number of autosegmental processes acting on tonal categories is greater in the latter.

# Chapter 5

## Mixed systems on two tiers

### 5.1 Introduction

In this chapter, I consider prosodic systems that use both the X-tier and the  $\pi$ -tier in which  $\pi$ -nodes do not bear any contrastive features. Depending on the configurations of X-slots and  $\pi$ -nodes allowed in a given system, a diverse range of phenomena can be explained without overstepping the bounds of the present model. I will consider two languages in detail.

The first is Papiamentu, which has been described as showing both contrastive tone and contrastive stress. Unlike the tonal accent languages considered in the preceding chapter, however, tone in Papiamentu is *not* directly dependent on the location of stress—the two are said to be independently manipulable. In the analysis I present, Papiamentu is a privative tone language, where tone is represented with  $\pi$ -nodes. Stress, on the other hand, is formally a length contrast, represented by doubling of X-slots. This is supported both by the phonetic realization of stress (primarily duration) and the shapes of roots that undergo stress alternations.

The second language examined is Estonian, which is said to have a ternary length contrast. Because the present model does not allow three-to-one linking of X-slots to segmental feature chains, a literally ternary analysis is ruled out, so the contrast must be split across both the X-tier and the  $\pi$ -tier. I argue that the phonetic realizations of  $\pi$ -nodes can be determined syntagmat-

ically, so that the second  $\pi$ -node in the word receives additional phonetic duration. This gives rise to the “overlong” category of length, and also accounts for the realization of falling pitch that accompanies overlength, since tonal properties, too, are regularly mapped across the string of  $\pi$ -nodes. I show that, through the use of morphemes containing bare X-slots, floating  $\pi$ -nodes, and floating vowel feature chains, a morpheme-by-morpheme account of a number of Estonian quantity alternations can be put forth without recourse to rules which reference morphological features—something that has not been accomplished in previous studies, even when resorting to ad hoc representational devices.

## 5.2 Papiamentu

Papiamentu is widely cited as a language with independent lexically contrastive tone and stress (Römer 1983; Kouwenberg and Murray 1994; Rivera-Castillo and Pickering 2004; Remijsen and van Heuven 2005; Rivera-Castillo 2009, and many others). The language is a creole whose vocabulary is based largely on Dutch, Portuguese, and Spanish. It is spoken on the islands of Aruba (where it is known as “Papiamentu”), Bonaire, and Curaçao. While there is some variation between dialects, they remain mutually intelligible (Kouwenberg and Murray 1994). The variety most widely described in the linguistic literature seems to be Curaçao Papiamentu, which I address in this section.

I will argue that Papiamentu’s mixed prosodic system arises from a tonal contrast (represented on the  $\pi$ -tier) and a formal quantity contrast (represented on the X-tier), the latter of which surfaces as “stress”. I will propose an analysis that accounts for both lexical specification and morphophonological alternations of the prosodic features, with an explicit compositional account of the individual morphemes from which the surface contrasts are derived.

### 5.2.1 Contrastive tone and stress

In Papiamentu, the locations of tone (surface pitch) and stress (surface duration/intensity) can be independently co-varied in order to realize lexical contrasts. Tone and stress can distinguish lexical items and can be affected by inflectional and derivational morphology. While the exact interpretation of the facts differs, sources agree that the fullest distribution of both the tone and stress contrasts occurs in a number of bisyllabic triplets such as those seen in (1):<sup>1</sup>

(1) *Papiamentu prosodic triplets* (Remijsen and van Heuven 2005; Rivera-Castillo 2009)

a. 'σ̇σ̇	b. 'σ̇σ̇	c. σ̇'σ̇
'lòrà 'parrot'	'lòrá 'to turn'	lò'rá 'turned'
'pàrà 'bird'	'pàrá 'stop'	pà'rá 'stopped'

Words like (1a) are typical for all classes of words except verbs; verbs generally follow the pattern in (1b), which contrasts with (1a) only in tone. Past participles of bisyllabic verbs in this class are derived by shifting stress to the final position, as in (1c).

While some tonal minimal pairs are unrelated morphemes, shifting of the tone melody in bisyllables can be used to indicate a change from a noun to a related verb. In some cases, this involves a change in the final vowel as well, as seen in 'lining (n)' ~ 'line (v)':

(2) *Noun–verb conversion through tone melody* (Kouwenberg and Murray 1994: 25)

a. Noun:	b. Verb:
'σ̇σ̇	'σ̇σ̇
'káskà 'peel (n)'	'kàská 'peel (v)'
'rímà 'rhyme (n)'	'rímá 'rhyme (v)'
'húmà 'smoke (n)'	'hùmá 'smoke (v)'
'fùrù 'lining (n)'	'fùrá 'line (v)'

<sup>1</sup>Because examples in this section are from a number of sources, I adopt a single set of conventions for the transcription of Papiamentu. Segments are given more or less in IPA transcription. Stressed positions are preceded by “'” or followed by “:” and high and low pitch are marked with acute and grave accents, respectively. Positions marked with an acute accent are those which Remijsen and van Heuven (2005) find to be realized with rising pitch, at least in most phrasal positions. The phonological interpretation of these categories by other authors and in the present study will be discussed in turn. The following abbreviations are used in glosses: agent. = agentive, imp. = imperative, infl. = basic inflectional suffix, n = noun, part. = participle, v = verb.



A similar alternation can be seen within verbs, as bisyllabic verbs undergo a shift in tone melody to differentiate the imperative from the “uninflected” form of the verb:

(3) *Tone shift in imperative* (Kouwenberg and Murray 1994: 13)

- |    |              |    |               |
|----|--------------|----|---------------|
| a. | 'òó          | b. | 'óò           |
|    | 'pàrá 'stop' |    | 'pàrà 'stop!' |

While these patterns hold for all verbs whose basic forms are of the shape 'òó, the pattern is not without exceptions.<sup>2</sup> A number of bisyllabic verbs have a HL tone pattern, as in (4a). Interestingly, verbs of this class do not form their participles by undergoing a stress shift, but rather take a prefix *he-* (in Aruba) or *di-* (in Curaçao) (Kouwenberg and Murray 1994: 20). I have derived such forms in (4b) according to Kouwenberg and Murray’s (1994) description.

(4) *Bisyllabic HL verbs* (Kouwenberg and Murray 1994: 13, 20, adapted)

- |    |                          |    |                      |
|----|--------------------------|----|----------------------|
| a. | 'óò                      | b. | dì'óò                |
|    | 'súntfì 'kiss'           |    | dì'súntfì 'kissed'   |
|    | 'skéirù 'brush'          |    | dì'skéirù 'brushed'  |
|    | 'fètèr 'lace'            |    | dì'fètèr 'laced'     |
|    | 'béitèl 'chisel'         |    | dì'béitèl 'chiseled' |
|    | 'wéldèr 'weld'           |    | dì'wéldèr 'welded'   |
|    | 'frétù 'stuff, gorge on' |    | dì'frétù 'stuffed'   |
|    | 'fángù 'catch'           |    | dì'fángù 'caught'    |

Etymologically speaking, verbs that take the pattern in (4) are of Dutch or English origin, while LH verbs that show the stress shift are Portuguese or Spanish in origin. However, etymology is not the deciding factor, as there exist some etymologically Dutch or English verbs which are LH and shift their stress (Kouwenberg and Murray 1994: 20). What stress-shifting verbs do seem to have in common is that their stems are all vowel-final, though not all vowel-final roots shift stress.

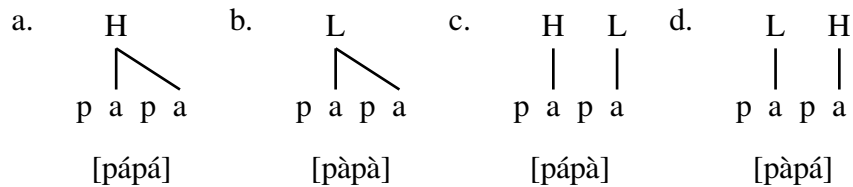
<sup>2</sup>The “basic” form of the verb here refers to that which is used as a citation form and in most morphosyntactic contexts—the majority of tense/aspect/mood morphemes are separate particles (Kouwenberg and Murray 1994: 42) and there is no separate verbal morphology for person or number marking. The only other verb forms are the past participle and the imperative, both of which I address here.

### 5.2.2 Theoretical accounts

The “stress” marked in the transcriptions above is assumed by most authors to constitute a single phonological category. Stressed positions are characterized by longer phonetic duration, along with increased intensity (Kouwenberg and Murray 1994; Rivera-Castillo and Pickering 2004; Remijsen and van Heuven 2005). Papiamentu stress is not said to have pitch as a correlate, and this is understood to be because tone is an independent phonological category.

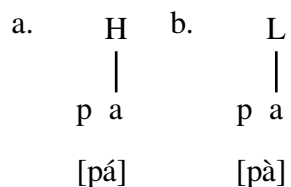
There are several ways to approach the representation of Papiamentu prosody. The first is to assume that all surface stress and tone is fully specified phonologically, which Remijsen and van Heuven (2005) claim is done by Römer (1991) and Rivera-Castillo (1998). Similar assumptions are made (at least implicitly) by Rivera-Castillo and Pickering (2004) and Rivera-Castillo (2009). If tone were taken to be a binary contrast between /H/ and /L/, then without further provisions, we might expect a greater range of possible tonal contrasts than are actually attested in Papiamentu:

(5) *Non-attested Papiamentu tone contrasts*



A binary tone system of this nature would also imply a contrast in monosyllabic words, which is not available in Papiamentu:

(6) *Non-attested Papiamentu monosyllabic tone contrasts*



Monosyllables instead show no prosodic contrasts at all—content words are generally described as showing surface stress and high tone. While stress is contrastive for words of greater than one syllable, the distribution of tone in Papiamentu is in fact quite limited: the HL–LH contrast is only available in bisyllabic words. Furthermore, the tonal contrast is constrained by the

location of stress. Whereas tone is contrastive in bisyllables with initial stress, those with final stress only show high tone in the position of stress; penultimate high tone is not allowed before final stress.

(7) *Restrictions on Papiamentu tone distribution* (Rivera-Castillo 2009: 440, adapted)

'pàrà 'bird'

'pàrá 'to stop'

pà'rá 'stopped'

\*pá'rà

Rivera-Castillo (2009: 440) interprets this as tone impinging on stress, stating that “tone attracts stress to penultimate position, overriding any contrastive role stress might have had in those minimal pairs.” On further consideration, however, it is not clear that this conclusion can be drawn. It could be, instead, that it is stress that attracts high tone to final position. The only thing that can be said for sure is that presence of final stress is correlated with the lack of tonal contrasts in bisyllabic words.

Devonish (1989) attempts to account for Papiamentu word prosody using purely tonal underlying forms. He first observes that in almost all cases, surface high pitch and surface stress occur in the same position, and assumes that the basic representation for stress (along with high pitch) is a formal H-tone. For bisyllabic LH words with penultimate stress (i.e., most bisyllabic verbs), two consecutive H-tones are specified underlyingly. The second is realized as high pitch, while the first is “reduced in the environment of an immediately following H-tone” (Devonish 1989: 60). Because this first H-tone appears in penultimate position, a typical position for stress cross-linguistically, it leaves behind duration, the non-pitch correlate for stress. Due to being “reduced”, however, it does not surface as a high tone.

While the interpretation of tone in Devonish’s (1989) account of Papiamentu is quite stipulative and the nature of “reduction” is not made clear, there is something to be said for the spirit of his approach: he attempts to capture the entire prosodic system with a restricted set of formal elements, doing so based on the assumption that stress and tone should only be separated

representationally when they constitute independent contrasts.

Remijsen and van Heuven (2005) present an alternative to fully-specified accounts of Papiamentu tone by describing the contrast in terms of “tone I” and “tone II”, taking inspiration from the literature on Scandinavian pitch accent. In citation form and in focused position, tone I involves rising pitch on the stressed syllable, while tone II involves rising pitch on the final syllable. In typical contexts, then, the high tones in most descriptions correspond, at least phonetically, to a rising (versus falling) pitch contour. They analyze this pattern in terms of a privative tone contrast encoded in a lexical HL falling tone, which is present in tone II words and absent in tone I words. The rising pitch under focus comes from an (LH) prominence tone, which links to the final syllable in tone II because of the presence of the lexical HL, as seen in (8a) and (8b). In tone I, the lexical tone is not present, so the prominence tone links to the stressed (penultimate) syllable, as in (8c).

(8) *Privative tone contrast in Papiamentu* (Remijsen and van Heuven 2005: 212, adapted)

a. $\text{II}$ 'lora ‘to turn’	b. $\text{II}$ lo'ra ‘turned’	c. $\text{I}$ 'lora ‘parrot’
$\begin{array}{cc} \text{'l} & \text{o} & \text{r} & \text{a} \\   &   & & \\ \text{HL} & (\underline{\text{LH}}) & & \end{array}$	$\begin{array}{cc} \text{l} & \text{o} & \text{'r} & \text{a} \\   &   & & \\ \text{HL} & (\underline{\text{LH}}) & & \end{array}$	$\begin{array}{cc} \text{'l} & \text{o} & \text{r} & \text{a} \\   & & & \\ & (\underline{\text{LH}}) & & \end{array}$
['lòrá]	[lò'rá]	['lórà]

Although Remijsen and van Heuven (2005) do not discuss it explicitly, this analysis seems to predict the lack of a tonal contrast in words with final stress, at least in focused positions. If the distinction between tone I and tone II is linking of the prominence tone to the stressed (versus final) syllable, then a word with final stress will realize its prominence tone on the final syllable no matter what—both stress and final position will attract it. However, given that their analysis conflates word-level and phrasal tone, and that their assumptions about the assignment and representation of stress are not made explicit, it is difficult to see what predictions (if any) they would make about the realization of words with final stress which do *not* have a lexical tone.

In monosyllabic words and words with more than two syllables, Remijsen and van Heuven (2005) generally transcribe tone I—rising tone on the stressed syllable. However, in cases where

suffixation adds a syllable to a tone II bisyllabic word (and shifts stress), they express some uncertainty:

“In affixation processes involving tone II word stems, it is unclear whether the tone pattern changes along with a shift in primary stress. For example, we have tentatively transcribed tone I in  $_{\text{I}},gana-'do$  ‘winner’, which is derived from  $_{\text{II}},gana$  ‘to win’. This transcription is impressionistic. It could be that the transcription  $_{\text{II}},gana-'do$  is actually more appropriate.” (Remijsen and van Heuven 2005: 210)

Again, all that can be said for sure is that no tonal contrasts are available in words of three syllables or greater. The formal representations of such words cannot be determined based solely on phonetics, impressionistic or otherwise.

Because Remijsen and van Heuven (2005) do not explicitly discuss how stress (lexical or predictable) is represented, it is impossible to infer an account of the morphophonology of stress alternations. Nonetheless, their claims that the lexical tone contrast is a privative opposition (with tone II as the marked member) and that surface rising tone can be due to several phonological factors (position of stress or final linking of prominence tone) offer some insight into the study of Papiamentu word prosody.

Dubina (2012, section 6.2) briefly addresses Papiamentu under a set of assumptions more similar to those of the present study. Dubina proposes that unpredictable morpheme-level stress is always represented using tonal autosegments, which predicts that only one of tone and stress may occur freely in a single language. Papiamentu is widely reported to be a counterexample to this claim because both tone and stress are known to be contrastive in the language. After considering a very limited data set of bisyllabic noun–verb–participle sets, he mistakenly concludes that tone, while contrastive, is always predictable based on morphological information. He claims that because of this, tone and stress can be used independently in the phonology. By allowing this possibility into his model, Dubina undermines the prediction that no language may use two kinds of prosodic autosegments at the same time, and he furthermore fails to illustrate what formal representations would be used when the two features come together in the same system.

From a survey of the existing literature on Papiamentu word prosody, we can draw two conclusions: (i) surface “stress” and “tone” are, at least in some formal configurations, each unpredictable and lexically contrastive, and (ii) no work on the language has given a thorough and explicit account of the representations and alternations involving both of the prosodic contrasts.

### 5.2.3 Papiamentu prosody in a unified model

I now turn to accommodating the Papiamentu prosodic system within the present model of word prosodic contrasts, informed by previous work on the language. I will begin with the representation of tone. It is clear that Papiamentu employs some kind of formal tone contrast in at least bisyllables with penultimate stress. Leaving aside for the moment the representation of stress, let us first consider the representation of tone in words where it is contrastive.

Where tone is contrastive, I propose that final high tone is reflected through privative lexical marking, by means of a featureless  $\pi$ -node. In citation form and in focus position, this  $\pi$ -node’s language-specific (and context-specific) phonetic realization is that of rising and then falling pitch. The  $\pi$ -node is floating in the underlying phonological representation, and by convention, undergoes linking to the right edge of the word, as in (9a). For words without marking on the  $\pi$ -tier, as in (9b), let us assume for the moment that a default tone pattern applies, with rising pitch on the penult.

- (9) a.  $\pi$  b.
- |         |         |
|---------|---------|
| × × × × | × × × × |
|         |         |
| p a p a | p a p a |
| [pàpá]  | [pápà]  |

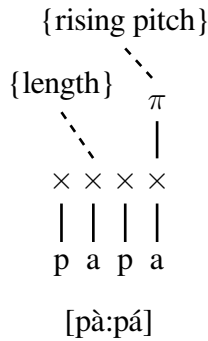
Similarly to the analysis given by Devonish (1989), the difference between LH and HL here is in the presence versus absence of an autosegment linked to the final syllable. The representations in (9) also draw on an insight from Remijsen and van Heuven’s (2005) analysis: the lexical distinction between their “tone II” and “tone I” is still encoded with a privative autosegment, where

the marked member of the opposition is that with a word-final pitch peak. Unlike the proposal by Remijsen and van Heuven (2005), however, this lexical marking is not composed of inherently substantive elements (H-tones and L-tones). Furthermore, the  $\pi$ -node links to the position which, in Remijsen and van Heuven's (2005) analysis, bears the prominence tone, not that which bears the lexical tone. Here being non-substantive is an advantage; although in citation form and focus position the  $\pi$ -node is realized as a rising–falling contour, nothing stops post-lexical rules from realizing it differently in other positions.<sup>3</sup>

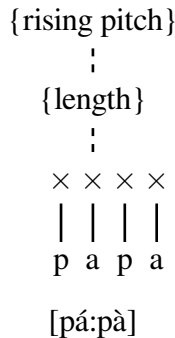
Let us turn now to the interpretation of Papiamentu stress. Several authors (Kouwenberg and Murray 1994; Remijsen and van Heuven 2005) have argued that Papiamentu stress is largely predictable, being quantity sensitive with default realization on the penultimate syllable. There is also a wide consensus that the primary phonetic correlate of “stress” in the language is increased duration (Devonish 1989; Kouwenberg and Murray 1994; Rivera-Castillo and Pickering 2004; Remijsen and van Heuven 2005; Rivera-Castillo 2009). In the presence of contrastively specified tone (a  $\pi$ -node), this default positional phonetic lengthening will occur independently, and it is this independence that has led CVCV words with final high tone to be described as having stress and tone in separate positions. However, when this lengthening-as-stress applies to default penultimate position, it does not actually constitute an independent contrast, but rather a default realization; it is only the final tone that is contrastively specified. This is illustrated in (10), where the phonetic correlate {length} is applied positionally by evaluating the string of vowel-bearing X-slots, and {rising pitch} is specified directly by the  $\pi$ -node. Note that in the transcription, the “stressed” position is henceforth notated as lengthened (with “:”) rather than with a primary stress mark in order to explicitly separate the relevant phonetic dimensions.

---

<sup>3</sup>Outside of focus positions, Remijsen and van Heuven (2005) say that prominence tone is not realized. They show that a flatter pitch occurs, rather than a rising pitch.

(10) *Default lengthening-as-stress and specified tone*

In the majority of cases, however, tone is not contrastively specified in Papiamentu, so the phonetic dimension of pitch is not given an independent locus. In such words, rising pitch and “stress” occur in the same position. I propose that pitch follows as a secondary correlate of positionally assigned duration—the different phonetic dimensions will tend to cluster together except when one is overtly specified elsewhere:

(11) *Pitch follows default lengthening-as-stress*

This understanding draws on the insight of previous analyses on the relationship between tone and stress in Papiamentu. For Remijsen and van Heuven (2005), words in which privative tone marking is absent realize their prominence tone on the “stressed” syllable. Here, a similar dependency relationship applies, but exists in the phonetic implementation rather than in the contrastive lexical phonology—stress corresponds to positional lengthening, and the prominence tone corresponds to the rising pitch which follows that lengthening. It is also in line with Devonish (1989) (and, presumably, Dubina 2012) in that tone and stress are the same entity (or in this case, non-entity) outside of situations where tone is represented separately.



Something should be said about the nature of the lengthening implied by “{length}” above. One possibility is that default stress involves phonological lengthening of the penultimate vowel at the contrastive lexical level by insertion of a second X-slot, which then attracts post-lexical rising pitch:

(12) *Pitch follows lexical lengthening-as-stress*

/Underlying/	→	Lexical lengthening	→	[Post-lexical pitch]
× × × ×         p a p a	→	× × × × ×     /     p a p a	→	{rising pitch} ↑ \ × × × × ×   /     p a p a [pá:pà]

The other possibility is that lengthening (like rising pitch) is entirely post-lexical, being conveyed merely schematically by the {phonetic realization} notation. Although double X-slot linking (virtual length) as in (12) will be exploited to specify exceptional stress below, in the case of default stress I know of no phonological evidence that can tease apart the two analyses. I thus remain agnostic as to the phonological status of lengthening in default stress words, and leave it an open question which possibility is more parsimonious.

While default “stress” in Papiamentu non-verbs typically falls on the penult, the language has been described as quantity sensitive (Kouwenberg and Murray 1994; Remijsen and van Heuven 2005). If the final syllable is closed (when the word ends in a consonant), then stress will instead fall on the final syllable. In monosyllabic content words, that single syllable is said to be stressed:<sup>4</sup>

<sup>4</sup>It is not clear what the word-final ⟨u⟩ in the words *bié:u* ‘old’ and *kòné:u* ‘rabbit’ represents here. However, the preceding vowel would receive stress according to Kouwenberg and Murray’s (1994) description in either of the likely cases: if ⟨u⟩ represents a separate syllable containing /u/, it is because the preceding vowel is in the penult and receives default stress there, while if ⟨u⟩ represents a coda /w/, the preceding vowel will be stressed because it appears in a closed final syllable.

(13) *Default stress patterns* (Kouwenberg and Murray 1994: 14, adapted)

monosyllabic items	final stress	penultimate stress
bié:u ‘old’	kòné:u ‘rabbit’	bá:hò ‘under’
stó:f ‘dust’	kàré:nt ‘currant’	làgàdí:fì ‘lizard’
bó:n ‘good’	fàlí:s ‘suitcase’	kué:rdè ‘spring (of water)’
ró:i ‘cleft, crevice’	kùràsó:n ‘heart’	pàlàbrú:à ‘owl’
bú:t ‘fine’	òmbèskó:p ‘insolent’	bùrá:kù ‘hole’

With the default pattern in (13), phonetic lengthening and (lacking a  $\pi$ -node to specify otherwise) rising pitch end up on a single vowel, but their position is not governed by an independent phonological contrast. The fact that these phonetic properties can be realized on vowels other than the penult is due to post-lexical phonetic interpretation of lexical segmental structure (X-slots and their features), *not* a separate phonological entity of “stress” or “accent”. In other words, the contrast is in quantity, not in the position of stress.

Kouwenberg and Murray (1994: 14) note that there are some apparent exceptions to this pattern, where consonant-final words receive penult stress. However, they point out that most of these end in /r/, /l/, or /n/, and that the orthographic ⟨e⟩ preceding the consonant is actually realized as a schwa, only being present due to phonetic implementation. This situation contrasts with words in which the same consonants occur after a full vowel, causing final stress:

(14) *(Non-)exceptional penult stress* (Kouwenberg and Murray 1994: 14, adapted)

final stress	penultimate stress
fàbó:r ‘favour’	bìná:gər ‘vinegar’
òspità:l ‘hospital’	ó:rxəl ‘organ’
màsàpá:n ‘marzipan’	sté:ndərt ‘standard (of car)’

While the apparent exceptions in (14) can be explained by post-lexical epenthesis, Kouwenberg and Murray (1994: 14–15) do list several types of “genuine” exceptions to the default stress pattern, giving the examples in (15). These include words that end in vowels with final stress,

words with antepenultimate stress, and words that end in a consonant with penultimate stress where the final vowel is not schwa:

(15) *Exceptional stress* (Kouwenberg and Murray 1994: 15, adapted)

irregular final stress	antepenultimate stress	penultimate stress
kàrnìsá: ‘pickled meat’	á:sìdò ‘acid’	difí:sìl ‘difficult’
màfjá: ‘very’	rá:pìdò ‘fast’	má:rtìr ‘martyr’
mùhé: ‘woman’	ìdé:ntìkò ‘identical’	

Since the words in (15) cannot be derived by the default stress rule, it is clear that some kind of lexical stress marking must be employed in the language. The main question here concerns what the formal nature of this marking should be—a question not typically addressed explicitly in the literature.

I argued above that default stress in Papiamentu does not constitute a separate phonological category, but rather that it is the result of the phonetic interpretation of segmental structure, which is independently needed in lexical representations. I propose that exceptionally stressed vowels are represented underlyingly with linking to two separate X-slots, as in (16): exceptional stress is phonological length. Because they are contrastively specified as long, they receive greater duration, giving them the primary phonetic correlate of stress. The other correlate, rising pitch, is associated with the long vowel, as in all words which do not have independent specification of tone.

(16) *Exceptional stress as vowel length*

a. {rising pitch}	b. {rising pitch}	c. {rising pitch}
× × × × ×         m u h e	× × × × × × ×           r a p i d o	× × × × × × × ×             d i f i s i l
[mùhé:]	[rá:pìdò]	[difí:sìl]

This analysis is based on two observations: (i) that length is not otherwise an independent phonological contrast in Papiamentu, leaving double X-slot linking available to perform other

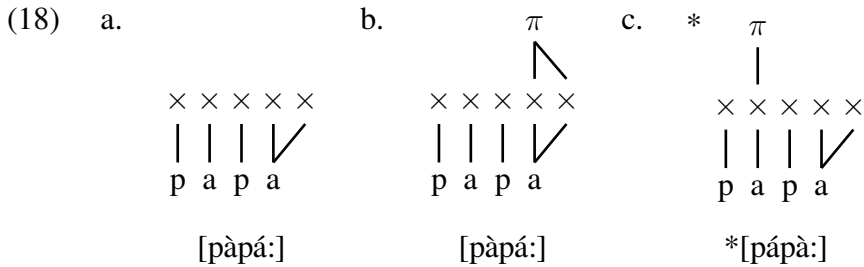
functions, and (ii) that the primary phonetic correlate of stress is duration, which, when contrastive (as stress is in Papiamentu) is typically represented with double X-slot linking. Such an approach is not without precedent. Bucci (2013) employs “virtual length” for accent in Corato Italian, which is represented by double linking to timing units (V-slots). Furthermore, Pater (1994) points out that such “honorary weight” (cast in terms of covert geminates) was employed as early as Chomsky and Halle (1968) in order to specify exceptional stressed light syllables in English.

In the context of the present study, this analysis makes a strong prediction about languages like Papiamentu: while they can have separate surface “stress” and “tone” contrasts, they cannot have an independent vowel length contrast in addition to this, since both tiers of root nodes are already spoken for.

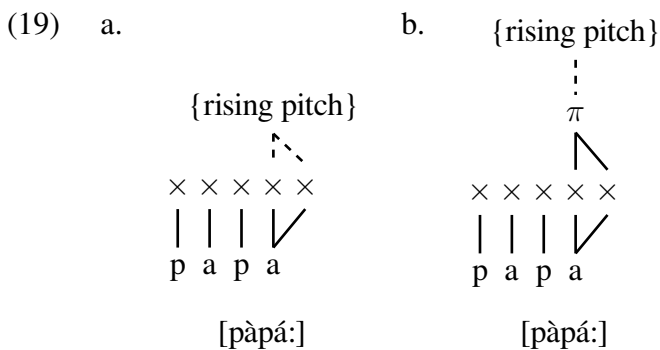
The analysis of Papiamentu word prosody put forth here also speaks to the restricted distribution of lexical tone. In bisyllables, which show the contrasts for both stress and tone, the tone contrast is only available in words with initial stress, as in (17a). In words with final stress, the contrast (at least on the surface) is neutralized, as in (17b).

- (17) a. Initial stress    b. Final stress
- |                         |                                   |
|-------------------------|-----------------------------------|
| $C\grave{V}:C\acute{V}$ | $C\grave{V}C\acute{V}: \quad LH$  |
| $C\acute{V}:C\grave{V}$ | $*C\acute{V}C\grave{V}: \quad HL$ |

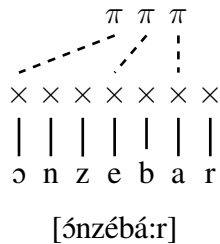
The fact that lexical tones can appear only at the right edge of a word follows from the language-specific property of association conventions. Just as autosegmental models of tone assume that association is either left-to-right or right-to-left, the present model allows the same restrictions for the association of  $\pi$ -nodes. In Papiamentu, association occurs from right to left, so that  $\pi$ -nodes, when present, are always aligned to the right edge of a word. Thus, the configuration in (18c) is ungrammatical in Papiamentu. The grammatical configurations in (18a) and (18b) do differ from each other structurally, but they produce identical surface outputs.



In both grammatical configurations, lexical “stress” is specified as phonological length (linking to two X-slots), but rising pitch comes from two different sources. In (19a), its phonetic realization follows the location of length, whereas in (19b), it is overtly specified by a  $\pi$ -node. While these two forms are indistinguishable on the surface, the possibility of course remains that there exists some kind of phonological evidence for a contrast between them. I leave this an open question for the time being.



Kouwenberg and Murray (1994: 15) report one other kind of exceptional prosodic pattern in addition to those thus far described. They claim that some polysyllabic words occur with high tone on every syllable, listing only the example *ónzébá:r* ‘poltergeist’. Although this word is exceptional in that it contains more than one vowel marked for tone, stress follows the default quantity sensitive pattern, that is, the vowel before a word-final consonant receives increased phonetic duration. Provided that this is actually an attested pattern in Papiamentu, it can be derived with the same representations and association conventions assumed above. The only difference is that rising (or high) pitch is assigned to each of the vowels hosting a  $\pi$ -node, because the word contains three floating  $\pi$ -nodes instead of only one:

(20) *Exceptional multiple H-tones*

The surface generalization to make about Papiamentu word prosody is that one position per word will have rising pitch (described by some authors as “high tone”, here notated  $\acute{V}$ ) and one position per word will have increased duration. In the absence of a  $\pi$ -node, rising pitch and duration will occur in the same position, but if a  $\pi$ -node is present, its host will receive rising pitch while increased duration is located separately. Increased duration can be specified phonologically either covertly by double linking of X-slots to vowel melodies, or overtly with a coda consonant. In the absence of any phonological quantity, increased duration is realized by default on the penultimate vowel. The assignment of surface phonetics to contrastive representations is schematized as the algorithm in (21):

(21) *Papiamentu surface word prosody algorithm*

1. If  $\pi$ -node present:
  - (a) Assign prominence tone to its host (default: rising pitch)
2. If vowel linked to two X-slots present:
  - (a) Realize greater duration of long vowel
  - (b) Assign prominence tone to long vowel if no  $\pi$ -node (default: rising pitch)
  - (c) End algorithm
3. If word ends in consonant:
  - (a) Assign greater duration to final vowel
  - (b) Assign prominence tone to final vowel if no  $\pi$ -node (default: rising pitch)
  - (c) End algorithm
4. Else:
  - (a) Lengthen penultimate vowel
  - (b) Assign prominence tone to penultimate vowel if no  $\pi$ -node (default: rising pitch)
  - (c) End algorithm

### 5.2.4 Morphemic analysis

While in many cases the phonological structure interpreted by the algorithm in (21) is simply present underlyingly on a single morpheme, in other cases it is derived from the concatenation of several morphemes. I will now address cases of the latter type, seeking to identify the phonological material provided by each morpheme involved in prosodic alternations in Papiamentu and demonstrating that the observed patterns can be elegantly accommodated within the present model.

Let us first consider alternations involving the tonal contrast. For most bisyllabic verbs, the different forms share the same segmental root, and are only distinguished prosodically. The imperative and basic inflected forms of these verbs both have stress in initial position, and are distinguished only by tonal melody; the imperative shows rising pitch in stressed position, while the basic form shows rising pitch in final position. This follows straightforwardly if the imperative is assumed to take a null suffix and the basic form takes a suffix consisting of a floating  $\pi$ -node. In (22a), the penultimate vowel receives default lengthening and rising pitch, while in (22b), the  $\pi$ -node links to the final vowel, causing rising pitch to be realized there:

- (22) a. ‘stop + IMP’      b. ‘stop + INFL’
- |   |  |
|---|--|
| $\times \times \times \times - \emptyset$<br>       <br>p a r a<br><br>[pá:rà]<br>‘stop!’ | $\times \times \times \times$<br>       <br>p a r a<br><br>[pà:rá]<br>‘stop’ |
|---|--|
- 

It is this same  $/-\pi/$  suffix that leads to tone being described as a marker distinguishing nouns from verbs. It is true that nouns and the basic forms of verbs differ only in tone, but this is not due to final rising tone distinguishing verbs from other morphological categories. Instead, it is due to the presence of an inflectional suffix consisting of a  $\pi$ -node on the verbal forms which happen to be used as citation forms, as in (23c). The unsuffixed noun in (23a) and the null-suffixed imperative

of the verb in (23b) are pronounced identically, with the penultimate vowel receiving both default lengthening-as-stress and rising pitch:

- (23) a. ‘rhyme’      b. ‘rhyme + IMP’      c. ‘rhyme + INFL’
- |  |   |  |
|--|---|--|
| $\begin{array}{cccc} \times & \times & \times & \times \\   &   &   &   \\ r & i & m & a \\ [r\acute{i}:m\grave{a}] \\ \text{‘rhyme (n)’} \end{array}$ | $\begin{array}{cccc} \times & \times & \times & \times - \emptyset \\   &   &   &   \\ r & i & m & a \\ [r\acute{i}:m\grave{a}] \\ \text{‘rhyme!’} \end{array}$ | $\begin{array}{cccc} \times & \times & \times & \times \\   &   &   &   \\ r & i & m & a \\ [r\grave{i}:m\acute{a}] \\ \text{‘rhyme (v)’} \end{array}$ <div style="text-align: right; margin-right: 20px;"> <math>\begin{array}{c} - \pi \\ \text{---} \end{array}</math> </div> |
|--|---|--|

Additional evidence that tone is not a marker of morphological category, but rather due to the phonological form of concatenated morphemes, comes from exceptions to the generalization that bisyllabic non-verbs have penultimate rising pitch and bisyllabic verbs have final rising pitch. Exceptional non-verbs (those with final rising pitch) do in fact differ from typical ones, but merely because they contain floating  $\pi$ -nodes in their underlying representations. Thus *pà:pá* ‘father’ has the same phonological shape as *pà:rá* ‘stop’, except its tonal specification (the  $\pi$ -node) is not contributed by a suffix:

- (24) ‘father’
- |  |
|--|
| $\begin{array}{cccc} & & \pi & \\ & & \vdots & \\ \times & \times & \times & \times \\   &   &   &   \\ p & a & p & a \\ [p\grave{a}:p\acute{a}] \\ \text{‘father’} \end{array}$ |
|--|

Exceptional verbs (those that have rising pitch on the penult), then, are not exceptional because they fail to be marked for their morphological category. Instead, they select for a null allomorph of the inflectional suffix. The lack of a  $\pi$ -node anywhere in the concatenated phonological form prevents rising pitch from being realized in final position, so default penultimate stress and rising pitch apply in the imperative and the basic inflected form, which both take null suffixes:



- (25) a. ‘kiss + IMP’      b. ‘kiss + INFL’
- |               |               |
|---------------|---------------|
| × × × × × – ∅ | × × × × × – ∅ |
|               |               |
| s u n tʃ i    | s u n tʃ i    |
| [sú:ntʃi]     | [sú:ntʃi]     |
| ‘kiss!’       | ‘kiss’        |

In addition to tone, prosodic alternations in Papiamentu can also affect the position of stress. The most notable (and apparently most productive) of these occurs when forming the past participle of bisyllabic verbs. This alternation involves moving stress to a final open syllable, which is an exception to the default pattern and requires, in the representations laid out above, double linking of X-slots to the final vowel melody. I propose that this occurs due to a participle suffix which consists of only an X-slot. This X-slot links to the final vowel melody, causing phonological length and the subsequent realization of final “stress”:

- (26) ‘stop + PART’
- |             |
|-------------|
| × × × × – × |
| - - -       |
| p a r a     |
| [pàrá:]     |
| ‘stopped’   |

Note that no  $\pi$ -node is present in (26).<sup>5</sup> Rising pitch is realized on the final syllable, but it is due to the presence of length there; tone therefore need not be contrastively specified. The verb never has more than a single suffix—it takes the basic inflectional suffix ( $/-\pi/$ ), the participle suffix ( $/-\times/$ ), or, in the imperative, a null suffix ( $/-\emptyset/$ ).

A participle suffix consisting of only an X-slot is only capable of causing a stress shift in vowel-final verbs, since it cannot form a virtual long vowel if it appears next to a consonant:

<sup>5</sup>If the final syllable of (26) did bear a  $\pi$ -node, its phonetic realization would not be any different. It is thus possible, in this analysis, for the past participle suffix to contain a  $\pi$ -node in addition to its X-slot. The presence of the X-slot, however, is crucial, since the presence of a  $\pi$ -node alone would result in penultimate stress (as in the basic inflected form of ‘stop’ in 22b).

- (27) \* × × × × × - ×  
 | | | | |  
 f e t e r  
 ‘laced’

This is not problematic, however, as all verbs that take the /-×/ allomorph of the participle morpheme end in vowels. The exceptional class of verbs—that which has rising pitch on its penultimate vowel in the basic inflected form and which is sometimes consonant-final—instead selects for a segmental prefix allomorph /di-/ of the participle morpheme. This does not interfere with the realization of stress or tone:

- (28) a. ‘PART + kiss’      b. ‘PART + lace’
- |                   |                  |
|-------------------|------------------|
| × × - × × × × ×   | × × - × × × × ×  |
|                   |                  |
| d i    s u n tʃ i | d i    f e t e r |
| [d̥isú:ntʃi]      | [d̥ifé:tèr]      |
| ‘kissed’          | ‘laced’          |

The past participle suffix is not the only morpheme in Papiamentu that causes a shift in the position of stress. The agentive suffix *-do* is “autostressed” (Kouwenberg and Murray 1994: 15), that is, it always bears stress. For example, the bisyllable *gà:ná* ‘win’ appears as *gànàdó*: ‘winner’ (Kouwenberg and Murray 1994: 23). This pattern can be accounted for if the suffix vowel /o/ is underlyingly linked to two X-slots. The suffix is attached directly to the root, as in (29b), where the phonologically long vowel causes the realization of final stress, along with final rising tone. This is in contrast with the basic inflected form of the verb in (29a), where instead the inflectional suffix /-π/ is attached to the root, causing final rising pitch, and the position of lengthening-as-stress is non-exceptional:

- (29) a. ‘win + INFL’      b. ‘win + AGENT’
- |   |  |
|---|--|
| $\begin{array}{cccc} \times & \times & \times & \times \\   &   &   &   \\ \text{g} & \text{a} & \text{n} & \text{a} \end{array}$ | $\begin{array}{cccccc} \times & \times & \times & \times & - & \times & \times & \times \\   &   &   &   &   &   & / \\ \text{g} & \text{a} & \text{n} & \text{a} & \text{d} & \text{o} \end{array}$ |
| [gà:ná]   | [gànàdó:]  |
| ‘win’   | ‘winner’   |

The final issue to be addressed in Papiamentu word prosody concerns words of three or more syllables. While in general tone contrasts are not available for such words, the location of stress can be affected by morphological considerations. While nouns receive default stress, with occasional exceptions being lexically marked by double X-slot linking, trisyllabic and longer verbs receive final stress in all forms. That is, the basic and imperative forms are homophonous (Kouwenberg and Murray 1994: 13), and are the same as the past participle (Kouwenberg and Murray 1994: 15). This can be accounted for if the inflectional suffixes for all three of these forms are phonologically identical and consist of only an X-slot. Much like in bisyllabic past participles, these suffixes link with the final vowel of the root, leading to both stress and rising pitch:

- (30) a. ‘squander + IMP’      b. ‘squander + INFL’      c. ‘squander + PART’
- |  |  |  |
|--|--|--|
| $\begin{array}{cccccccc} \times & \times & \times & \times & \times & \times & \times & \times & - & \times \\   &   &   &   &   &   &   &   & \backslash & \text{---} \\ \text{d} & \text{i} & \text{s} & \text{t} & \text{r} & \text{i} & \text{b} & \text{i} & & \end{array}$ | $\begin{array}{cccccccc} \times & \times & \times & \times & \times & \times & \times & \times & - & \times \\   &   &   &   &   &   &   &   & \backslash & \text{---} \\ \text{d} & \text{i} & \text{s} & \text{t} & \text{r} & \text{i} & \text{b} & \text{i} & & \end{array}$ | $\begin{array}{cccccccc} \times & \times & \times & \times & \times & \times & \times & \times & - & \times \\   &   &   &   &   &   &   &   & \backslash & \text{---} \\ \text{d} & \text{i} & \text{s} & \text{t} & \text{r} & \text{i} & \text{b} & \text{i} & & \end{array}$ |
| [dìstrìbí:]  | [dìstrìbí:]  | [dìstrìbí:]  |
| ‘squander!’  | ‘squander’   | ‘squandered’   |

Stress can still shift off the end of the root in such words, however. When the agentive suffix *-do* is used, the suffix receives stress and rising pitch, as in bisyllabic verbs (Kouwenberg and Murray 1994: 15). This pattern is derived in the same way for trisyllabic verbs as it is for bisyllabic ones; the phonologically long vowel on the suffix receives stress due to its lexical long vowel, and rising pitch follows. The root-final vowel in this case does *not* receive stress because it has no long vowel, morphologically derived or otherwise:

- (31) 'squander + AGENT'
- |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|
| × | × | × | × | × | × | × | × | - | × | × | × |
|   |   |   |   |   |   |   |   |   |   |   | / |
| d | i | s | t | r | i | b | i | d | o |   |   |
- [distribidó:]
- 'squanderer'

The existence of /-×/ verbal suffixes is further supported by stress minimal pairs with related nouns. One strategy (though not a fully productive one) for differentiating verbs and nouns of three or more syllables involves a change in the position of stress (Kouwenberg and Murray 1994: 26). While nouns will usually take default stress, verbs will take final stress. As with the examples above, the shift of stress to final position is due to suffixation:

- (32) a. 'vaccination'      b. 'vaccination + INFL'
- |   |   |   |   |   |   |  |  |  |  |  |  |
|---|---|---|---|---|---|--|--|--|--|--|--|
| × | × | × | × | × | × |  |  |  |  |  |  |
|   |   |   |   |   |   |  |  |  |  |  |  |
| v | a | k | u | n | a |  |  |  |  |  |  |
- [vákú:nà]
- 'vaccination'
- |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|
| × | × | × | × | × | × | - | × |   |   |   |   |
|   |   |   |   |   |   |   |   | - | - | - | - |
| v | a | k | u | n | a |   |   |   |   |   |   |
- [vákùná:]
- 'vaccinate'

### 5.2.5 Summary

In this section, I surveyed the literature on Papiamentu word prosody and demonstrated a formal analysis of the system that seeks to capture stress and tone within a single cross-linguistic model of prosodic contrasts. I proposed that neither surface stress nor tone are necessarily specified using contrastive phonological mechanisms outside of exceptions to the default implementation of these phonetic properties. Finally, I argued that morphological alternations which seem to involve only changes in prosodic structure can be derived by a straightforward compositional analysis if it is assumed that affixes may consist only of featureless X-slots and  $\pi$ -nodes.

## 5.3 Estonian

Estonian is famous for its three-way quantity contrast, but the short–long–overlong distinction is not simply a ternary property of the segment in the way that long–short distinctions seem to be in many languages. Estonian quantity is instead typically understood to be a combination of several levels such as segments, syllables, or feet. In this section I will argue that Estonian quantity is represented with a combination of X-slots and  $\pi$ -nodes. The  $\pi$ -nodes behave formally as in a tone language; they are floating underlyingly and link to the X-tier by association conventions. However, they have syntagmatically-determined phonetic realizations carrying durational and tonal characteristics which lead to the manifestation of “overlength”. In this analysis,  $\pi$ -nodes serve as the highest level of contrastive prosodic organization, and no formal reference is made to syllables or feet.

Based on this system of representations, I show that a compositional analysis of a number of nominal paradigms can be elegantly formulated using roots and suffixes that contain floating segments, X-slots, and  $\pi$ -nodes, without appeal to morphophonological rules that directly manipulate prosodic structure based on morphological categories. The only interaction with the morphology is through allomorph selection.

### 5.3.1 The quantity system

Estonian shows a three-way surface length contrast, and the language has many sets like those in (33).<sup>6</sup> The three degrees of length are known as “short”, “long”, and “overlong”, which are often referred to as Q1, Q2, and Q3, respectively (Q for “quantity”). The words in (33a) differ in progressively increasing duration of the vowel, while those in (33b) differ in progressively increasing

<sup>6</sup>For the Estonian transcriptions in this section, Q1 will always consist of (C)V sequences. Long segments are indicated by doubling the vowel or consonant symbol, and the overlength of Q3 is indicated by the addition of “:”. Note that these conventions do not follow the language’s orthography, which does not consistently indicate Q3. Vowel quality does follow the orthography, however: ⟨õ⟩ = /ɤ/ (a mid-back unrounded vowel) and ⟨ü⟩ = /y/. The following abbreviations are used in glosses: gen. = genitive, ill. = illative, nom. = nominative, part. = partitive, and shrt.ill. = short illative. All example words have minimally been checked against a morphologically annotated normative dictionary that *does* indicate Q3 (Raadik 2013).

duration of the medial consonant. The words in (33c) only show a contrast between Q2 and Q3; diphthongs cannot appear in Q1, but the Q2–Q3 contrast is available on the diphthong, typically transcribed on the diphthong’s second component.

(33) *Estonian quantity minimal sets*

	“short”	“long”	“overlong”
	Q1	Q2	Q3
a.	vina	viina	vii:na
	‘vapour (nom.)’	‘vodka (gen.)’	‘vodka (part.)’
b.	lina	linna	lin:na
	‘flax (nom.)’	‘city (gen.)’	‘city (part.)’
c.		laulu	lau:lu
		‘song (gen.)’	‘song (part.)’

It is widely recognized that quantity involves some element above or beyond the segment. This is evidenced by the distribution of contrasts in the system; if individual segments could appear as short, long, or overlong, then a great many more distinctions would be predicted than are actually available in Estonian. The fullest set of contrasts is available within a bisyllabic group. These are given schematically in (34), following Prince (1980):

(34) *Basic syllable types* (Prince 1980: 515, adapted)

	Q1	Q2	Q3	
	CV-CV	CVV-CV	CVV:-CV	a.
		CVVC-CV	CVV:C:-CV	b.
		CVC-CV	CVC:-CV	c.
			*CVVC:-CV	d.
			*CVV:C-CV	e.

A word containing only short segments and open syllables is referred to simply as Q1. Anything more complex—the presence of a coda, long segment, or diphthong—allows a contrast

between Q2 and Q3. The segmental length contrast for vowels (V versus VV) in initial position and for consonants (C versus CC) in medial position covary freely, but the same cannot be said of the Q2–Q3 distinction. Overlength is not a category that can be identified for individual segments. Furthermore, as indicated in (34d–e), the position of the extra length that is characteristic of Q3 is not distinctive. The segments marked with “:” in the Q3 shapes in (34a–c) should not be taken at face value either, however; Prillop (2013: 3–4) notes that there is some dispute in the literature as to which components of the rhyme actually receive the lengthening, and that there may be some variability depending on the segments involved. The important phonological fact is that the presence of this lengthening is itself a binary property: a bisyllabic group beginning with a complex syllable either has overlength or it does not.

Three-way contrasts are not available at all in monosyllables. Monosyllables are always said to appear with Q3, so whatever phonological property is characteristic of Q3, it is a minimal requirement of monosyllables. Although all monosyllables have Q3, they do show segmental length contrasts between vowels and final consonants, where one or both may be long:<sup>7</sup>

(35) *Segmental length in Q3 monosyllables*

- a. CVV      maa:    ‘land (nom.)’
- b. CVVC    vii:n    ‘vodka (nom.)’
- c. CVCC    lin:n    ‘city (nom.)’
- d. CVVCC   paa:t:t   ‘boat (nom.)’

Alternations in quantity often follow the expression of morphological categories. Consider the data in (36), which exemplifies an extremely common paradigm in Estonian morphophonology. Many nominals are Q3 monosyllables in the nominative case, but gain a final vowel in the genitive and partitive cases, which surface in Q2 and Q3, respectively.

---

<sup>7</sup>The only exceptions to this seem to be function words, which can be CVC or CV. I have observed that, in practice, those shaped CV are augmented to CVV or CVh when coerced into strong positions.

(36)	Nominative	Genitive	Partitive	
	Q3	Q2	Q3	
<hr/>				
a.	vii:n	viina	vii:na	‘vodka’
b.	lin:n	linna	lin:na	‘city’
c.	lau:l	laulu	lau:lu	‘song’

The final vowel in the genitive and partitive forms may be any of /a, e, i, u/, a subset of the language’s nine-vowel inventory; the vowel associated with any given word is not predictable (Blevins 2008: 248).<sup>8</sup> Because of the unpredictability of these vowels, there are many pairs like those in (37), where two different lexical items have homophonous nominative cases and are distinguished only in their other forms:

(37)	Nominative	Genitive	Partitive	
	Q3	Q2	Q3	
<hr/>				
a.	tuu:r	tuuri	tuu:ri	‘tour’
b.	tuu:r	tuura	tuu:ra	‘sturgeon’

In addition to duration, the surface phonetics of Estonian quantity is also in the process of acquiring tonal properties, with the consequence that the distinction between Q2 and Q3 is coming to be differentiated by pitch (Lehiste 2003; Lippus and Ross 2011). It has been shown in perception experiments that, when duration is kept identical, pitch contour alone serves as a reliable cue for distinguishing Q2 and Q3 (Lehiste 2003; Lippus et al. 2011). The generalization is that, at least where there are vowels to host the pitch, a fall in pitch occurs between the first and second syllable in Q1 and Q2, but within the first syllable in Q3. I schematize this pattern in (38):

(38)	Q1	Q2	Q3
<hr/>			
a.	vínà	víínà	vî:nà
b.		láúlù	láu:lù

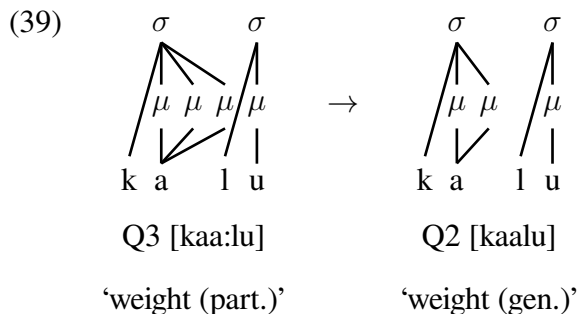
<sup>8</sup>Some lexical items show segmental alternations in final vowels between genitive and partitive, but because of the focus on quantity and for reasons of space, I will not address these here.



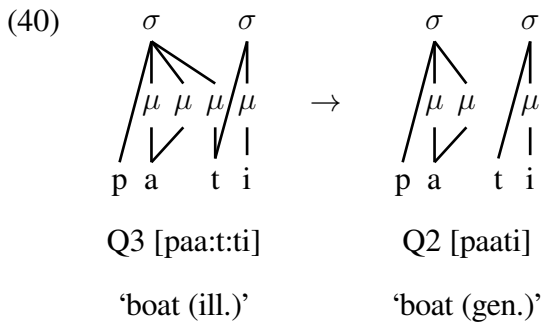
### 5.3.2 Representing three-way quantity

The literature contains many different phonological interpretations of the Estonian quantity system. Although assumptions vary from analysis to analysis, accounts of Estonian quantity generally come in one of two flavours: either they analyze the contrast as one of literal ternarity, or they split the contrast binarily across two representational tiers.

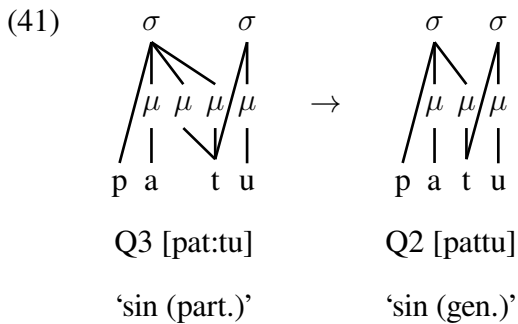
Hayes (1989, 1995) assumes that the three-way contrast is represented in a literally ternary manner. Q1 syllables contain one mora, Q2 syllables contain two moras, and Q3 syllables contain three moras. Hayes's system represents segmental quantity with the same device: short vowels have one mora and long vowels have two or more moras, while short consonants have no moras and long consonants have one or more moras. Q2 is derived from underlyingly trimoraic forms by the application of a morphological rule which deletes the third mora (Hayes 1989: 296), as in (39). No specifics of the rule are articulated, as it simply applies in "those morphological environments in which the grade alternation is manifested by overlength" (Hayes 1989: 295).



In addition to Q2/Q3 quantity alterations, many nominals show a process of consonant gradation, through which consonants are degeminated when they appear in their weak grade (Q2) versus their strong grade (Q3) forms. In Hayes's analysis, degemination follows automatically from the mora deletion rule because geminate consonants must be dominated by a mora, and losing this mora makes them revert to singletons:



Because Hayes also uses three moras to represent Q3 syllables consisting of a short vowel and a geminate, his deletion rule correctly predicts the behaviour of geminates in this context. The short vowel is dominated by a single mora, and the geminate is dominated by two moras. When the third mora is deleted in order to derive the Q2 weak grade form, the consonant retains a mora, and thus remains a geminate:



While this analysis elegantly captures several cases of co-occurrence of segmental quantity, syllabic quantity, and gemination, it fails to offer an account of CVVCCV words in Q2. Ehala (2003: 56) points out that there exist minimal pairs such as the following:

- (42) a. Q2 [aitta] ‘Aita (proper name)’  
 b. Q3 [ait:ta] ‘barn (part.)’

Representing (42a) would require a trimoraic syllable (as in 43a), but it is unclear why this should be realised in Q2 if trimoraicity is the defining characteristic of Q3. Ehala notes that Q3 (42b) would then require a quadrimoraic syllable, as in (43b), but such a possibility suggests no obvious upper bound for the number of moras allowed in a single syllable.

- (43) a. b.
- Q2 [aitta]                      Q3 [ait:ta]
- ‘Aita (proper name)’            ‘barn (part.)’

Even if we allow representations like (43b), Hayes’s mora deletion rule makes the wrong prediction for the weak grade of [ait:ta] ‘barn’. The actual attested weak grade form of this stem is Q2 [aita], with no geminate, but applying Hayes’s rule to (43b) leaves the consonant moraic, which would be realised as a geminate:

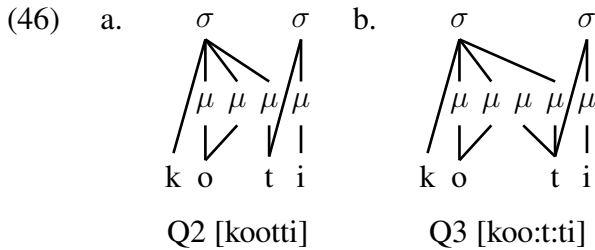
- (44)
- [ait:ta]                      \*[aitta]

Ehala (2003: 56–57) suggests that it would be undesirable but possible to use mora splitting for the diphthongs in words like Q2 [aitta] ‘Aita (proper name)’:

- (45)
- Q2 [aitta]

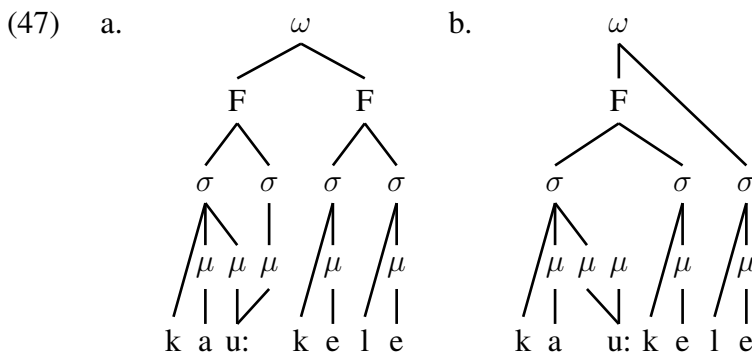
However, on closer inspection it is not so easy to save Hayes’s analysis. Lehiste (1977, 2003) notes a contrast between words such as Q2 [kootti] ‘gothic’ and Q3 [koo:t:ti] ‘flail (part.)’. Since Q2 [kootti] contains a long vowel which must remain in contrast with Q2 [kotti] ‘sack (gen.)’, and there is no obvious way to split a mora onto a single vowel melody, the long vowel must remain bimoraic, with Q3 [koo:t:ti] being quadrimoraic, as in (46b):<sup>9</sup>

<sup>9</sup>The only way around this is to assume that long vowels are represented as two consecutive short vowels (Hint 1997), but see Lehiste (1985) for an argument against this.



The weak grade of Q3 [koo:t:ti] ( $\rightarrow$  Q2 [kooti]), which Hayes would derive through mora deletion, is analogous to that of (42b) Q3 [ait:ta] ( $\rightarrow$  Q2 [aita]) in that it undergoes degemination. This cannot be derived by deleting any one mora from (46b).

Bye (1997) attempts to implement a kind of literal ternarity without sacrificing maximally bimoraic syllables. He does this by using a third free-standing mora which dominates the centre of overlength, as in (47b), but because this mora is not dominated by the syllable node, there are technically no trimoraic syllables. This optionally coexists with a representation in which overlength consists of a bisyllabic foot whose second member is formally a monomoraic syllable, as in (47a). Ehala (2003: 56) spells these out as follows:



Bye's proposal is based on the optionality of secondary stress patterns for this word (and others like it), which can be pronounced either with secondary stress immediately following the Q3 syllable (*ká:u:kèle*, 47a) or with no secondary stress (*ká:u:kele*, 47b). Ehala (2003) notes that this would cause problems for Q2 [aitta] 'Aita (proper name)' (mentioned above), which would require a trimoraic first syllable in order to represent the diphthong and the geminate consonant. Bye (1997: note 3) does consider the possibility of mora sharing more generally, but Ehala (2003: 56–57) argues that “the three devices together (degenerate syllables, free moras and splitting) would

seriously reduce the attractiveness of the theory allowing for far more possible combinations than there are attested syllable types”.

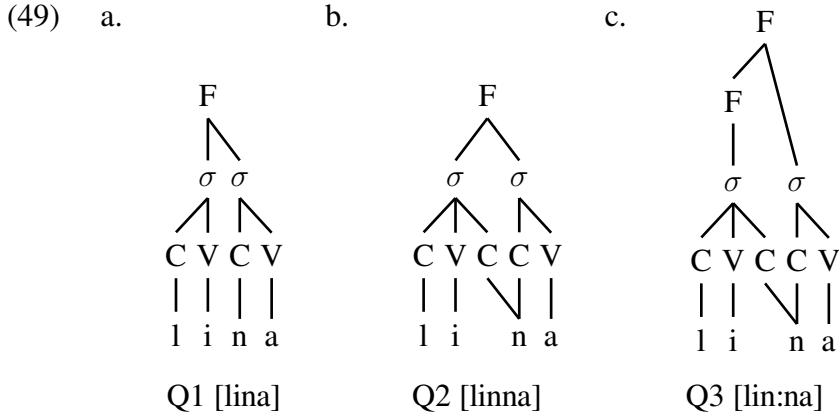
It is thus clear that using literal ternarity to represent the Estonian quantity system necessarily involves incorporating additional complications into any theory based solely on a single timing tier.

Viitso (2003: 11-12) describes the Estonian quantity system as a contrastive property of the syllable, reflected as a combination of length and weight as in (48). Light long syllables are in Q2 (long), while heavy long syllables are in Q3 (overlong). Short (CV) light syllables are all in Q1 (short); short heavy syllables do not exist.

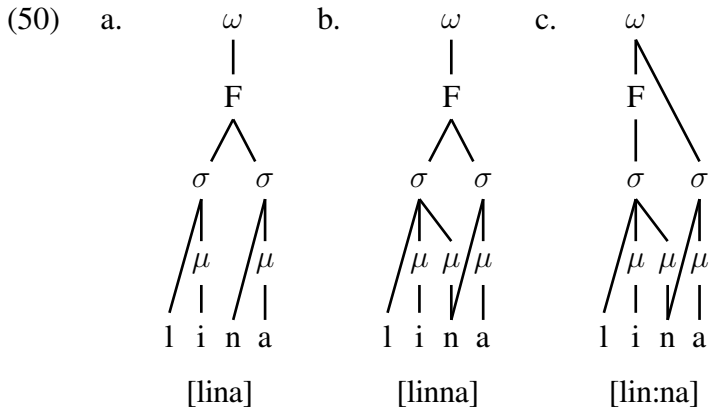
(48)		Short	Long
	Light	Q1	Q2
	Heavy		Q3

It is advantageous from a generative perspective to mirror the description in (48) representationally, by splitting “length” and “weight” onto different tiers, but there is disagreement as to which tiers these should be. Based on the distribution of stress patterns, work by Hint (1973), Prince (1980), Mürk (1982), Bye (1997), and Prillop (2013) has exploited foot structure to accomplish a three-way contrast. The three-way contrast exists only within a foot—monosyllables can appear only in Q3—and both Q1 and Q2 constitute bisyllabic feet, being distinguished from each other based on segmental/syllabic quantity (i.e., whether the syllable has a branching or non-branching rhyme). Q3 syllables themselves constitute monosyllabic feet.

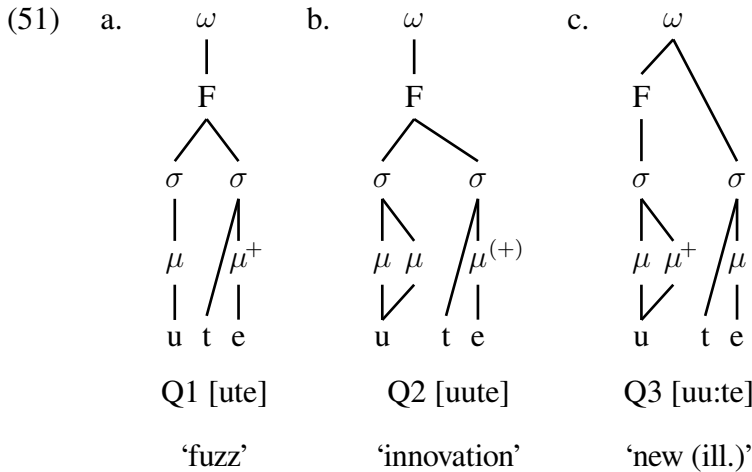
Prince (1980) implements Q3 bisyllables using recursive foot structure, by allowing the first syllable/foot to serve as the head of a larger foot, as in (49c). The additional duration in Q3 comes from a general phonetic rule that lengthens the weak branch of a foot, which in this case targets the second half of the syllable. Q1 (49a) and Q2 (49b) each consist of a single foot, and are differentiated based on information on the CV-tier:



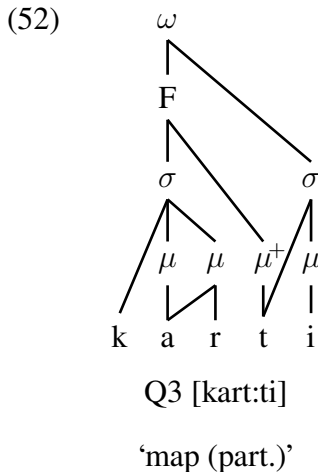
Odden (1997: 179) suggests that the use of recursive foot structure can be avoided while maintaining Prince’s insight if the second syllable of Q3 is instead extrametricaly adjoined directly to the prosodic word, as in (50c). The representations for Q1 and Q2 would presumably remain essentially identical to Prince’s, as in (50a) and (50b):



Prillop (2013) takes an approach similar to Odden’s by allowing the direct adjunction of extrametrical syllables, and follows Prince by representing Q3 syllables as exhausting their own feet. Foot-final lengthening, which she notates with a superscript “+”, affects the final mora of a foot. This lengthening occurs in Q1 (reflected by the presence of a half-long vowel in the second syllable) and in Q3 (reflected by overlength). It also applies optionally in Q2 (Prillop 2013: 16, phonetic transcriptions adapted):



While Prilop does not use literal ternarity to reflect the three-way quantity contrast, she does need to use a third mora within the Q3 foot in order to represent CVRC.CV words with a medial geminate. This geminate is moraic, but Prilop avoids using trimoraic syllables by adjoining this mora directly to the foot (Prilop 2013: 19):



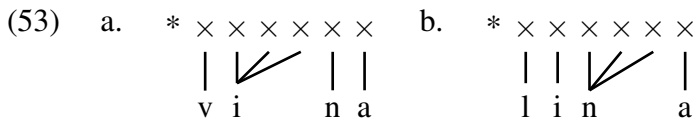
Prilop also proposes an additional theoretical tool: strong moras (those with “+”, which undergo lengthening) and weak moras (those without “+”, which are not lengthened) must be lexically distinct. A weak mora in the input causes the syllable to be realised in Q2, while a strong mora causes it to be realised in Q3. She likens such an opposition to languages in which stress is a lexical phenomenon (Prilop 2013: 20–21).<sup>10</sup> While she states in prose “I assume that Estonian only allows strong moras at the end of syllables, feet or words” (Prilop 2013: 19), it is not clear

<sup>10</sup>Cf. Mürk (1981, 1982), who proposes that foot structure itself can be lexically specified in Estonian.

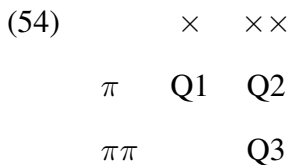
what part of her analysis actually captures this. If strong moras are lexically contrastive, then we would predict inputs that have a long vowel or diphthong in the first syllable whose first mora is strong ( $CV_{\mu^+}CV_{\mu}$ ), or short syllables that end with strong moras ( $CV_{\mu^+}CV_{\mu}$ ).

### 5.3.3 Three-way quantity in a minimal model

Having reviewed previous approaches to representing Estonian quantity, I will now establish a system of representations using the mechanisms provided by the present model. I follow the line of research that assumes that Estonian quantity is split across two tiers at the lexical level. A flat single-tier literal ternarity approach to Estonian quantity is not available in the present model because three-to-one associations between root nodes and feature chains are universally banned:



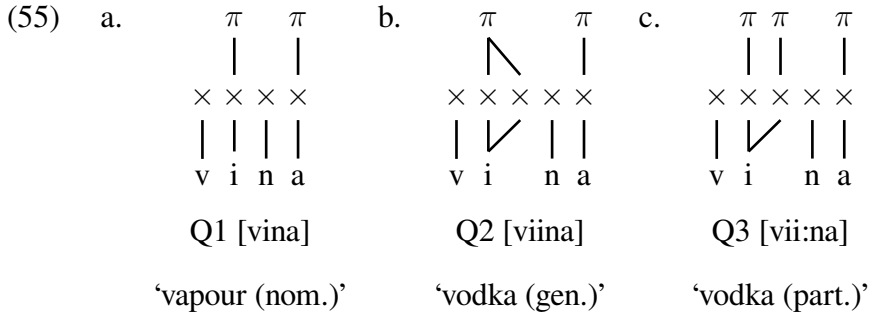
In the two-tier representations that I propose, X-slots represent the binary “length” distinction and  $\pi$ -nodes represent the “weight” distinction—the presence of two  $\pi$ -nodes linked to adjacent X-slots represents Q3:



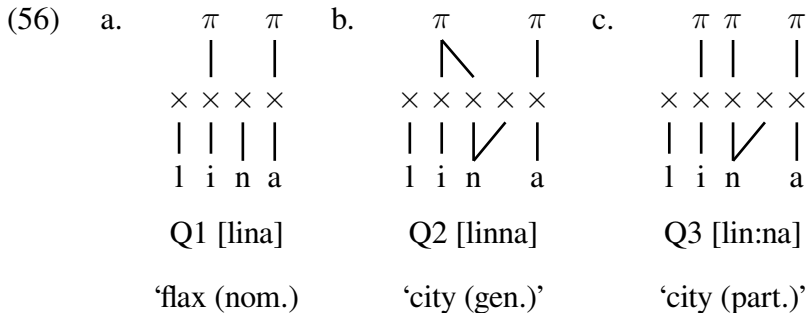
Q1 is then our term to describe a group of two  $\pi$ -nodes in which the  $\pi$ -nodes do not link to adjacent X-slots and the first  $\pi$ -node links to a single X-slot, as in (55a). Since the segmental feature chains are linked to a single X-slot each, all segments in Q1 are interpreted as short. Q2 is used to describe any group of two  $\pi$ -nodes in which the  $\pi$ -nodes are linked to non-adjacent X-slots and the first  $\pi$ -node is linked to more than one X-slot, as in (55b). In this case, both X-slots are linked to a single segmental feature chain, representing a long vowel. Finally, Q3 is now the term to refer to a group of two  $\pi$ -nodes in which there are no intervening  $\pi$ -nodeless X-slots, as in (55c). In all cases, the first  $\pi$ -node is minimally linked to a vocalic X-slot, but the second



$\pi$ -node, when adjacent to the first, may link to any X-slot. In order to realize a second syllable, the word in Q3 requires an additional  $\pi$ -node to host the word-final vowel.



The representation in (56a) differs from that in (55a) only in terms of the segmental features of the initial  $\pi$ -nodeless consonant; the forms are prosodically identical. The representation in (56b) shows that a single  $\pi$ -node in Q2 can be split between a vocalic X-slot and a consonantal X-slot, and that in (56c) shows that the second adjacent  $\pi$ -node can link to any X-slot, not only to vocalic ones.



These X-slots need not reflect long segments, however. They can link to coda consonants, or even to a combination of vocalic and consonantal segments. For example, a Q2–Q3 contrast is available in words (even relatively recent loans) in which the first syllable contains a short vowel and a complex coda:<sup>11</sup>

<sup>11</sup>Recall that, although associations between X-slots and feature chains are maximally binary (ruling out a literally ternary analysis of overlength), associations between the two tiers of root nodes (that is, between X-slots and  $\pi$ -nodes) may be unbounded. Thus, the representation in (57a) does not violate any well-formedness conditions.

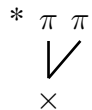
- (57) a.  $\pi$   $\pi$  b.  $\pi$   $\pi$   $\pi$
- |  |  |
|--|--|
| $\diagdown$<br>$\times \times \times \times \times \times$<br>$\mid \mid \mid \mid \mid \mid$<br>p u n k t i | $\mid \diagdown \mid$<br>$\times \times \times \times \times \times$<br>$\mid \mid \mid \mid \mid \mid$<br>p u n k t i |
| Q2 [punkti]  | Q3 [pun:k:ti]  |
| ‘point (gen.)’   | ‘point (part.)’  |

It is also available in words containing a diphthong followed by a long consonant:

- (58) a.  $\pi$   $\pi$  b.  $\pi$   $\pi$   $\pi$
- |  |  |
|--|--|
| $\diagdown$<br>$\times \times \times \times \times$<br>$\mid \mid \diagup \mid$<br>a i t a | $\mid \diagdown \mid$<br>$\times \times \times \times \times$<br>$\mid \mid \diagup \mid$<br>a i t a |
| Q2 [aitta]   | Q3 [ai:t:ta]   |
| ‘Aita (proper name)’   | ‘barn (part.)’   |

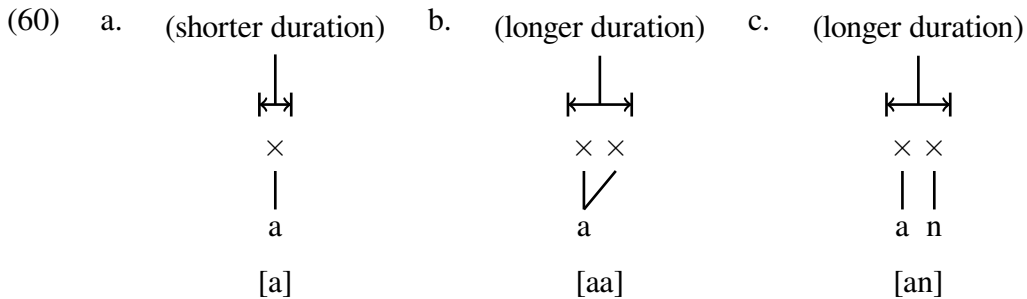
Note that there is a gap in the table in (54) above, in that there is no category with a single X-slot linked to two  $\pi$ -nodes. Given the formal nature of  $\pi$ -nodes in the present model—they are identical to tonal root nodes—there is a straightforward and cross-linguistically attested way to account for this gap. A single X-slot can bear only a single  $\pi$ -node due to the same constraint that (in languages where it is active) prevents the formation of short contour tones, as stated in (59). The only difference is that, unlike in languages with a tonal inventory that contains multiple members, in Estonian,  $\pi$ -nodes do not bear contrastive features. The formal reference that the constraint makes to  $\pi$ -nodes is the same, however.

(59) *X-slots may maximally bear one  $\pi$ -node* ( $\approx$  “no short contour tones”)



The phonetic correlates of the prosodic contrasts are determined as follows. The phonological length contrast (Q1 versus Q2) is distinguished by a binary contrast of X-slots (one versus two). Basic segmental duration is read cumulatively directly from X-slots; a segmental feature chain

linked to two X-slots has a longer duration than a segmental chain linked to a single X-slot. This holds whether the two X-slots are linked to a single feature chain or two consecutive feature chains:

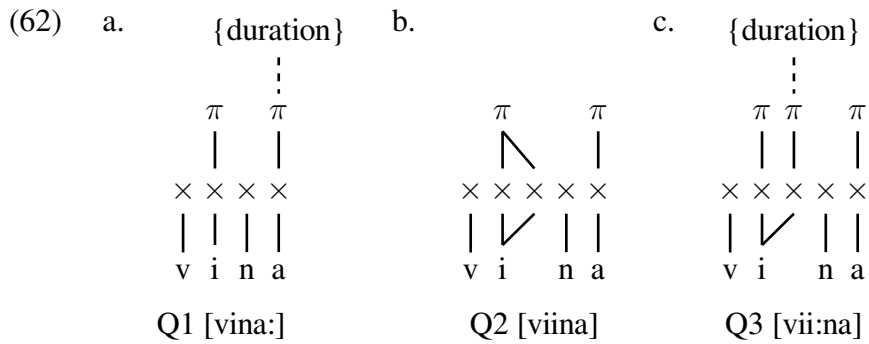


The phonetic realization of the  $\pi$ -node in Estonian is determined contextually, based on the configuration of  $\pi$ -nodes and X-slots. In Q1 and Q3, additional length is realized on the segments linked to the second  $\pi$ -node in the word. In Q2, no such lengthening applies—the length seen in Q2 is due to the spelling out of X-slots, not to lengthening from  $\pi$ -nodes. This generalization can be stated formally as in (61). In Q1, the first  $\pi$ -node is only linked to one X-slot (a short vowel), and in Q3, it is linked to a single X-slot, whether it belongs to a short vowel, long vowel, or diphthong. In Q2, however, the first  $\pi$ -node links to at least two X-slots (minimally, a short vowel and a coda consonant or the two X-slots of a long vowel or diphthong).

(61) *Durational realization of Estonian  $\pi$ -nodes*

If the first  $\pi$ -node in a word is linked to a single X-slot, apply additional length to the second  $\pi$ -node

The application of this phonetic lengthening to all three quantity degrees is schematized in (62). Here it can be seen that the additional phonetic length in Q3 applies on top of phonological length, giving rise to “overlength”, and distinguishing it from the corresponding “long” quantity seen in Q2.



Note that in (62a), phonetic lengthening is transcribed for the final (and phonologically short) vowel in Q1. This phenomenon is traditionally known as the “half-long vowel”, and has been incorporated into previous accounts of Estonian quantity in various ways. In Prince’s (1980) analysis, it is due to general foot-final lengthening—the position of the second  $\pi$ -node in the present representations corresponds to the right edge of the first foot in Prince’s account. For Prillop (2013), extra length comes from phonological “strong” moras, and thus is a lexically specified property. In my previous work on the topic (Spahr 2014b), it is due to a regular process of second mora lengthening. The  $\pi$ -nodes of the present account are equivalent to the “moras” proposed by Spahr (2013, 2014b,c).<sup>12</sup>

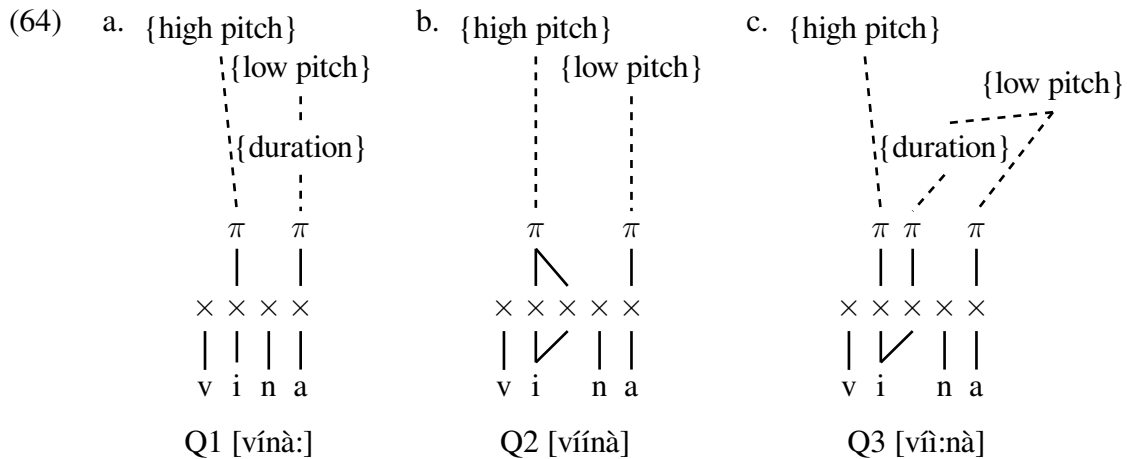
Phonetic tone is similar to phonetic duration in that it is mapped across the string of  $\pi$ -nodes. However, it differs from duration in that it is not sensitive to the number of X-slots to which the  $\pi$ -nodes are linked. The pattern is that a drop in pitch occurs from the first  $\pi$ -node to the second, as stated in (63).

(63) *Tonal realization of Estonian  $\pi$ -nodes*

Realize high pitch on the word’s first  $\pi$ -node and lower pitch on subsequent  $\pi$ -nodes

When these phonetic realizations are applied to the phonological representations, the result is that Q1 and Q2 show high pitch in the first syllable and low pitch in the second syllable, while Q3 shows falling pitch within the first syllable:

<sup>12</sup>The term “second mora lengthening” is borrowed from Suomi and Ylitalo (2004), who propose the phonetic lengthening process to explain durational and tonal properties present in a number of dialects of Finnish. Finnish differs from Estonian, however, in that the location of the second “mora” is not itself distinctive, so it does not give rise to a surface three-way quantity contrast. Nonetheless, phonetic lengthening has been connected with phonological processes in various dialects, including gemination (Nahkola 1987; Spahr 2011) and epenthesis (Karlin 2014).



Although lengthening of the second  $\pi$ -node and a fall in pitch between the first and second  $\pi$ -nodes are facts of Estonian phonetic realizations, in the analysis that follows, I will not notate tone, nor will I transcribe phonetic length outside of Q3. In terms of phonological representations, the following conventions are assumed:

- (65) **Q1:** The first  $\pi$ -node is linked to a single vocalic X-slot and is not adjacent to the second  $\pi$ -node (the first two  $\pi$ -nodes do not link to abutting X-slots)
- Q2:** The first two  $\pi$ -nodes are linked to non-adjacent X-slots, and the first  $\pi$ -node is linked to more than one X-slot
- Q3:** The first two  $\pi$ -nodes are linked to adjacent X-slots

### 5.3.4 Prosodic alternations and empty root nodes

I will now demonstrate how a number of Estonian quantity alternations can be accounted for using a compositional morphological analysis. Because of the complexity involved, I will introduce the data for each class of nominals in turn, followed by the technical details of the analysis. The principal forms to be accounted for in analyzing Estonian quantity alternations are the nominative, the genitive, and the partitive. With the exception of the short form illative (examined below), almost all other case and number marking is expressed segmentally, building on the genitive stem (Viitso 2003). Thus for each paradigm examined here, I will address only these three forms.

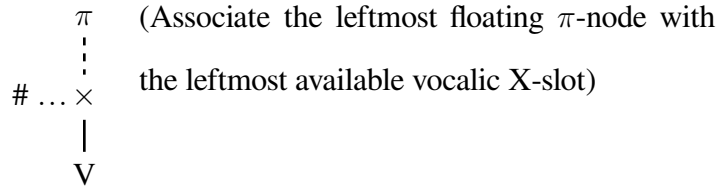
I begin with the basic alternations seen in (66), repeated from (36) above. For nouns of this class, a monosyllabic nominative alternates with bisyllabic genitive and partitive case forms. The genitive and partitive are segmentally identical, being differentiated only by quantity: the genitive is in Q2 while the partitive is in Q3.

(66)	Nominative	Genitive	Partitive	
	Q3	Q2	Q3	
a.	vii:n	viina	vii:na	‘vodka’
b.	lin:n	linna	lin:na	‘city’
c.	lau:l	laulu	lau:lu	‘song’

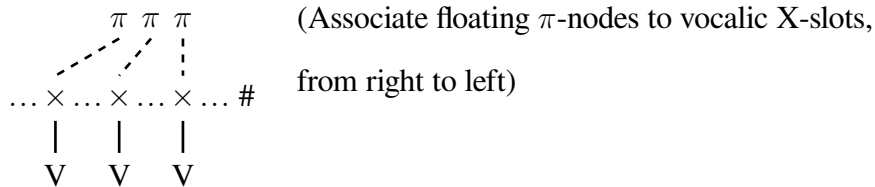
Recall from above that the final vowel in the bisyllabic forms, one of /a, e, i, u/, is not predictable based on the shape of the nominative. It must thus be the case that it is a lexical property of the root. I assume that these vowels are represented as floating segmental feature chains, that is, they are not linked to an X-slot in their underlying forms. All other segmental features are underlyingly linked to X-slots, however. I further assume that  $\pi$ -nodes are present underlyingly on roots, but that, like tone melodies in many tone languages, they are not underlyingly associated with X-slots. Thus, each morpheme may lexically contain one or more floating  $\pi$ -nodes. The underlying forms of the roots in (66) are shown in (67), where floating vowels are notated with underlining.

(67)	a.	‘vodka’	b.	‘city’	c.	‘song’
		$\pi$ $\pi$		$\pi$ $\pi$		$\pi$ $\pi$
		× × × ×		× × × ×		× × × ×
		v i n <u>a</u>		l i n <u>a</u>		l a u l <u>u</u>

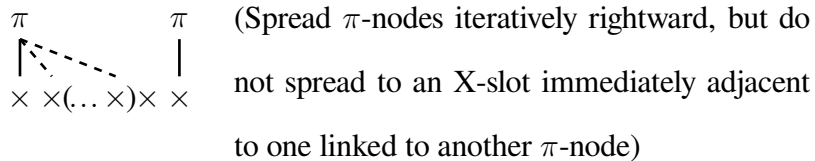
In order to reach the output representation, phonological processes are needed that link  $\pi$ -nodes to the X-tier. I propose that the first such process is the initial  $\pi$ -node association rule in (68), which links the leftmost floating  $\pi$ -node in the word to the leftmost vocalic X-slot that does not bear a  $\pi$ -node.

(68) *Initial  $\pi$ -node association rule ( $\pi$ 1-to-V1)*

Remaining floating  $\pi$ -nodes then associate by convention from right to left, linking to other vocalic X-slots. This process is reflected by the rule in (69):

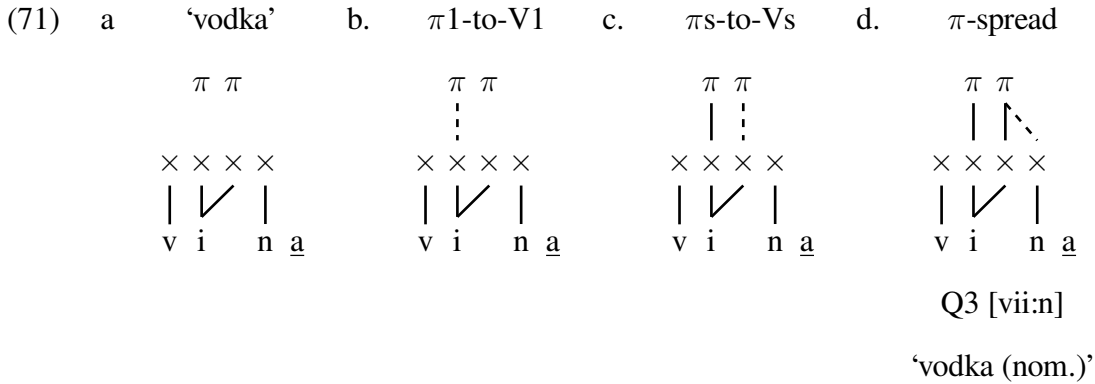
(69)  *$\pi$ -node association convention ( $\pi$ s-to-Vs)*

Finally, the rule in (70) causes  $\pi$ -nodes to spread rightward onto additional  $\pi$ -nodeless X-slots. Note that it will not spread a  $\pi$ -node onto an X-slot immediately adjacent to an X-slot which already bears a  $\pi$ -node.<sup>13</sup> The need for this will become apparent below when analyzing additional paradigms.

(70)  *$\pi$ -node spreading ( $\pi$ -spread)*

I assume that nominative case forms are unsuffixed. For example, the bare root for ‘vodka’ is interpreted directly, deriving the surface form as in (71). First,  $\pi$ 1-to-V1 causes the initial  $\pi$ -node to associate with the first X-slot linked to /i/. Then,  $\pi$ s-to-Vs links the root’s second  $\pi$ -node to the second X-slot linked to /i/. Finally,  $\pi$ -spread adds an additional association line between the final  $\pi$ -node and the word-final X-slot, which is linked to /n/. The floating feature chain representing /a/ receives no phonetic realization because it is not linked to an X-slot, and the result is a monosyllabic word in Q3.

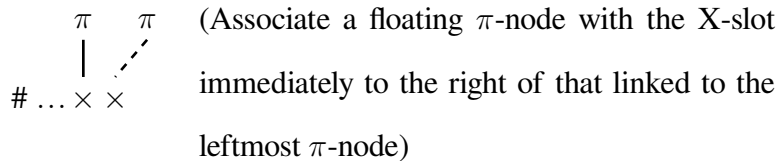
<sup>13</sup>Tone “retraction” effects causing configurations similar to those derived by (70) are known in the literature on OCP in autosegmental phonology (Myers 1997).



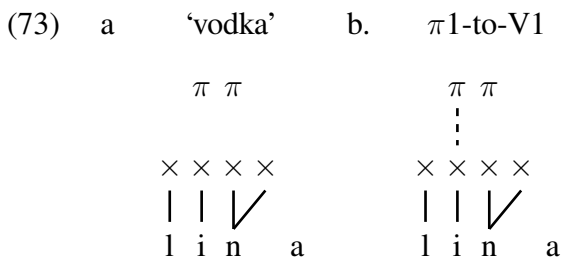
A derivation for Q3 [lau:l] 'song (nom.)' would proceed analogously. The only difference is that  $\pi$ s-to-Vs would link the second  $\pi$ -node to the X-slot bearing /u/ features—part of a diphthong rather than an X-slot representing the second half of a long vowel.

To derive words like Q3 [lin:n] 'city (nom.)', an additional rule is needed. Whereas roots of the shape CVVC have a second vocalic X-slot to which  $\pi$ s-to-Vs can attach the second  $\pi$ -node, in CVCC roots the second  $\pi$ -node must be associated with the postvocalic consonantal X-slot. This can be accomplished with the rule in (72), which causes a floating  $\pi$ -node to link to an X-slot to the right of an X-slot which already bears a  $\pi$ -node.

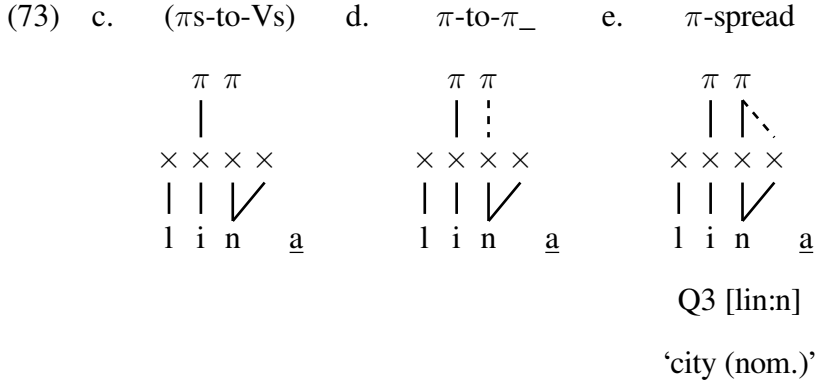
(72) *Post-vocalic  $\pi$ -node association ( $\pi$ -to- $\pi$ \_)*



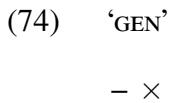
The derivation proceeds similarly to Q3 [vii:n] 'vodka (nom.)' and Q3 [lau:l] 'song (nom.)' above, except it is now  $\pi$ -to- $\pi$ \_ that links the second  $\pi$ -node to the X-tier. This  $\pi$ -node is still spread to the end of the word by  $\pi$ -spread, and results in a Q3 monosyllable:



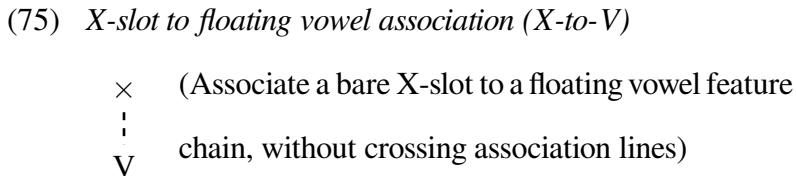




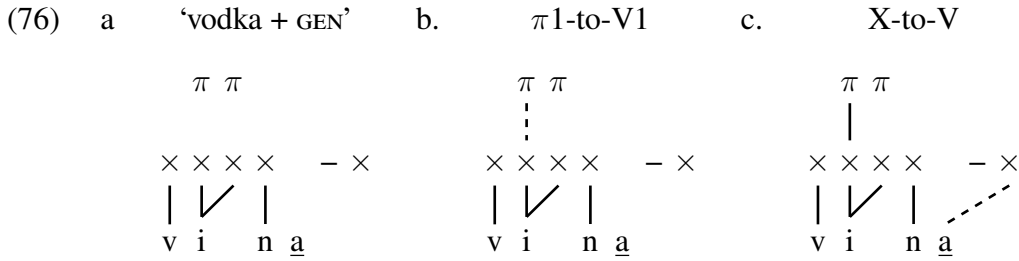
Let us now turn to the morphologically complex forms, that is, the genitive and partitive. While the nominative singular is simply the bare root, other case forms are marked with suffixes, which are often phonologically non-null. As is readily available in the present model, Estonian allows morphemes which contain only root nodes and have no feature specifications. I propose that the genitive case suffix applied to the roots discussed above has the underlying representation in (74)—it consists solely of an X-slot.



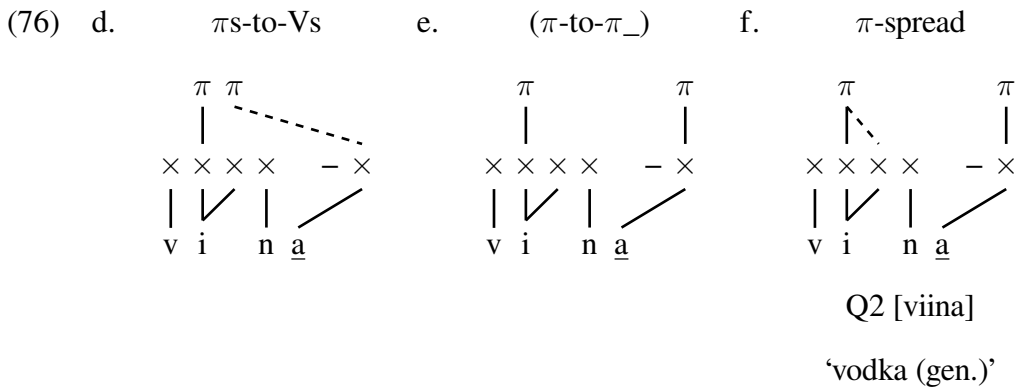
In deriving the genitive form, this bare X-slot must link to the segmental feature chain representing the floating vowel. This requires the linking process in (75), which will link a floating X-slot to a floating vowel.



When the genitive case suffix is concatenated with a root, initial  $\pi$ -node association first takes place as usual. This is followed by the application of X-to-V, which causes the suffix’s X-slot to link to the floating /a/:



When  $\pi$ s-to-Vs applies, the root’s second  $\pi$ -node links to the newly vocalic word-final X-slot. Thus a  $\pi$ -node must be “borrowed” from the root in order to realize the suffix. This leaves the root with only a single  $\pi$ -node, which associates to the second X-slot of the long /i/ because of  $\pi$ -spread. The  $\pi$ -node does not spread to the X-slot linked to /n/, because  $\pi$ -spread does not allow spreading to a position immediately adjacent to another  $\pi$ -node. The result is a bisyllable in Q2:

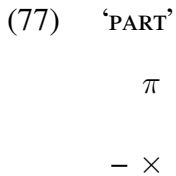


By deriving the Q2 genitive form in this way, several of the problems with previous analyses are avoided. First, unlike the foot-based accounts of Prince (1980) and Odden (1997), it is not necessary to assume that the weak versus strong grade (Q2 versus Q3) forms are derived from morphological footing rules—they result directly from the interpretation of the phonological material present in underlying forms. Second, unlike Hayes’s (1989) analysis, derivation of the weak grade involves no subtractive morphology—nothing need be deleted from the base form. Rather, the phonological elements (here,  $\pi$ -nodes) that are “missing” in the weak grade are simply not present in the input, because they are not part of the suffix.

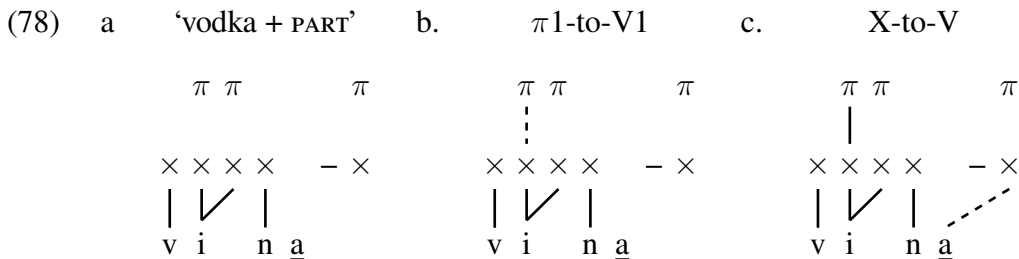
More importantly, this analysis reveals that the Estonian quantity system is less exotic than many previous studies have assumed. It does not require ad hoc devices such as multiple metrical parsings of the same segmental structure (Prince 1980; Odden 1997), weak/strong moras (Prillop

2013), or ternary use of timing units/moras (Hayes 1989). Instead, by using  $\pi$ -nodes and association conventions, Estonian is typologically the same as a tone language. Formally,  $\pi$ -nodes are equivalent to (and are operated on by the same kinds of processes as) tonal autosegments. Just like in a tone language, the tone melody will map across the available tone bearing units. With a two-tone melody, a monosyllabic form will carry a contour tone (as both tones map to the single syllable), but when a suffix provides an additional syllable, the tones will map across all of the available positions, preventing contour tone formation. This is precisely how Estonian quantity works, except the tonal (and length) properties of the different “melodies” are due to the syntagmatically determined phonetic realizations of  $\pi$ -nodes rather than to contrastive features on those  $\pi$ -nodes. The genitive provides only the timing unit necessary to provide a second syllable in which the root’s second  $\pi$ -node can link, preventing the formation of a “contour” tone which formally has no contour.

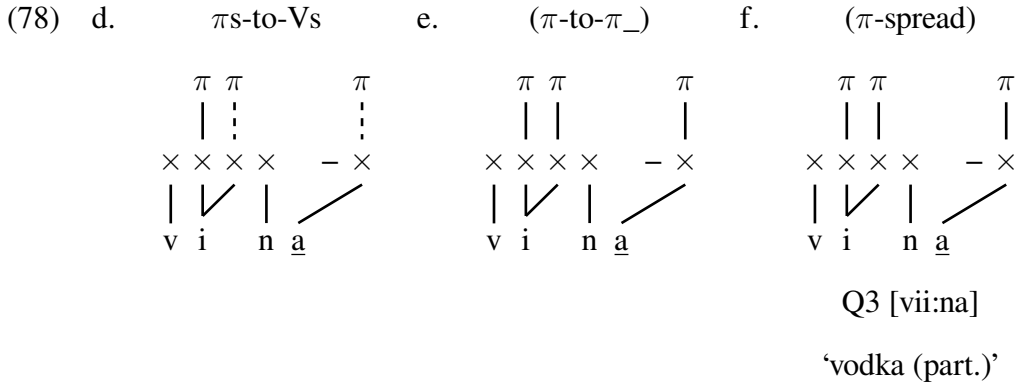
The partitive case is derived from a suffix with a different phonological shape. Like the genitive, the partitive suffix contains no segmental features, but in addition to an empty X-slot, it also contains a floating  $\pi$ -node, as in (77). In other words, the partitive contains a floating tone along with its segmental root node.



When this suffix is attached, initial  $\pi$ -node association and linking of the floating vowel features take place as normal:



However, the rightmost  $\pi$ -node, targeted by  $\pi$ s-to-Vs, is now that of the partitive case suffix. This prevents the root's second  $\pi$ -node from being borrowed, allowing it to remain attached within the root, as it does in the unsuffixed nominative case. The result is a bisyllable in Q3:



To return once more to the tone language analogy, the partitive case involves the addition of a suffix containing a floating tone, and consequently, a word with a tone melody containing one more member than that of the genitive. Because of the way tone mapping applies, the extra tone ends up in the first syllable, again giving rise to a contour.

### 5.3.5 Prosodic degemination

There exists another class of nominals which shows the same Q2–Q3 alternation, but furthermore shows shortening of root-final geminates in Q2. Geminates are realized finally in Q3 monosyllables and medially in Q3 bisyllables, but an alternation from Q3 to Q2 (in the genitive) coincides with an alternation from geminates to singletons.

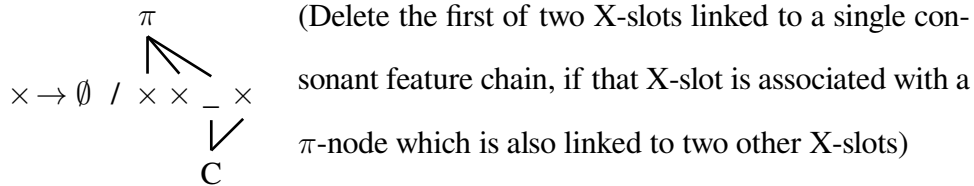
(79)

	Nominative	Genitive	Partitive	
	Q3	Q2	Q3	
a.	paa:t:t	paati	paa:t:ti	‘boat’
b.	tark:k	tarka	tark:ka	‘wise’
c.	poi:s:s	poisi	poi:s:i	‘boy’

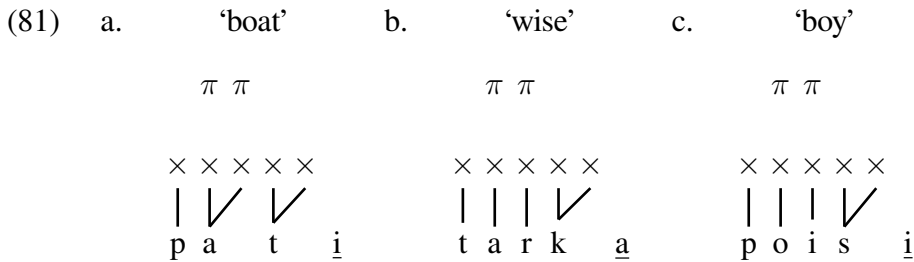
Prince (1980: 539) accounts for this with a rule of “Prosodic Degemination” which shortens a long consonant in the environment VX\_V. I adapt this in the present model as the rule in (80).

This deletes the first X-slot of a long consonant if it is linked to a  $\pi$ -node which is also linked to two other X-slots.<sup>14</sup>

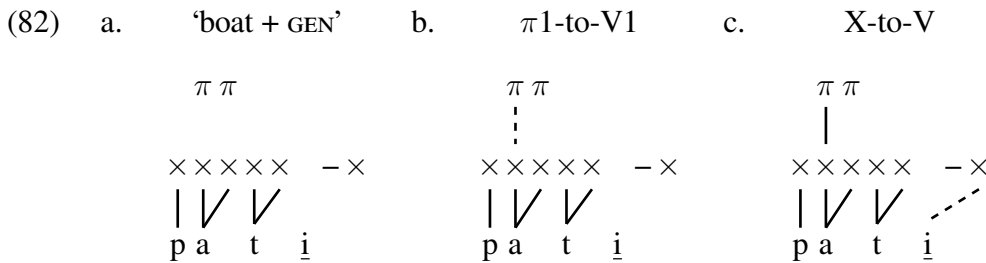
(80) *Prosodic degemination (Degem)*



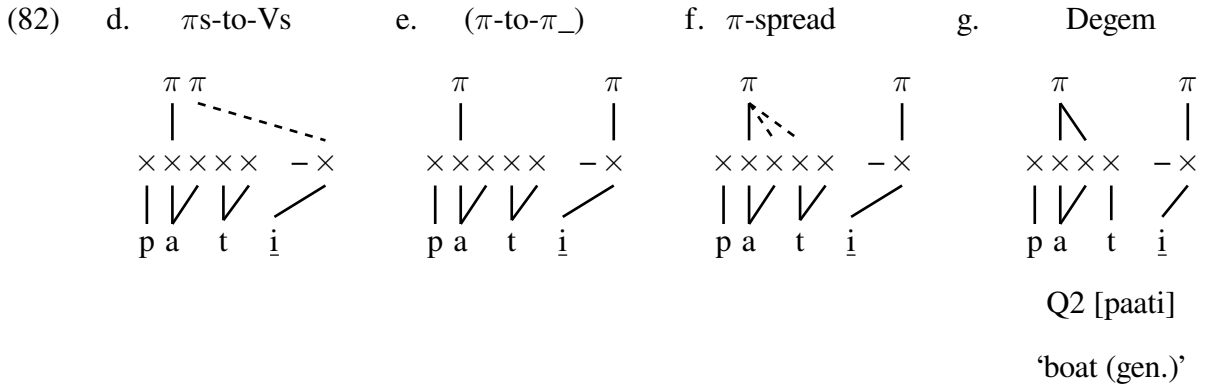
The root-final consonants are phonologically long in underlying representations:



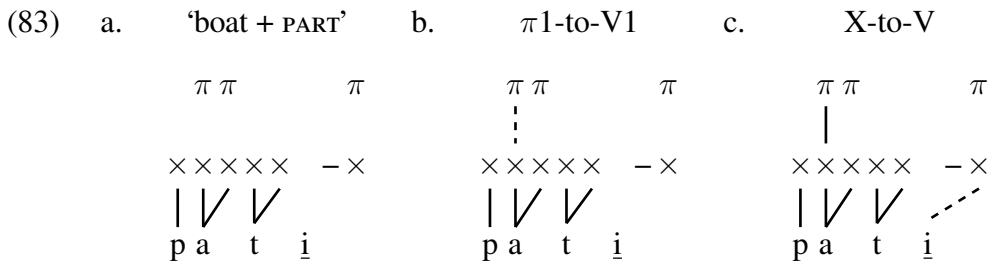
When the genitive suffix is added to these roots, the second  $\pi$ -node of the root attaches to the suffix's X-slot upon the application of  $\pi$ s-to-Vs. This leaves the root with a single  $\pi$ -node, which  $\pi$ -spread links to two additional X-slots, creating the environment for Degem. The result is a Q2 bisyllable with a medial singleton:



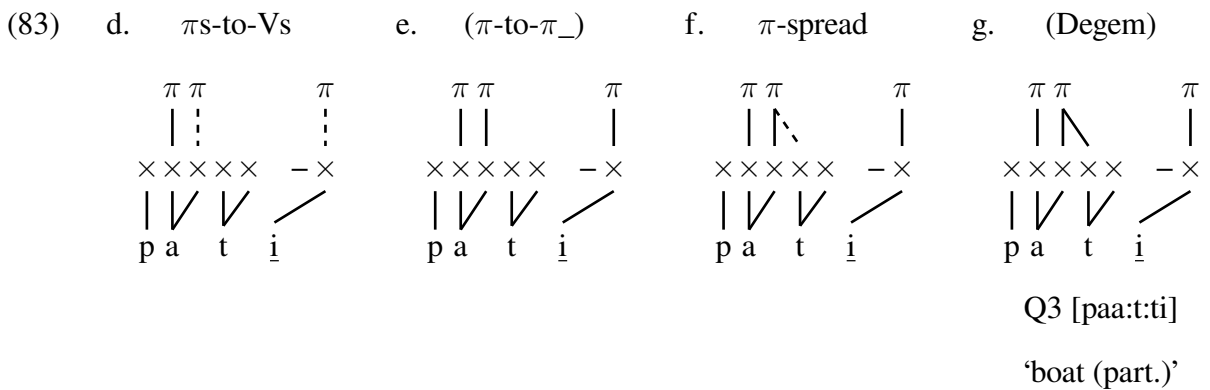
<sup>14</sup>Recall from the discussion above that Hayes (1989) accounts for degemination with the same mora deletion process that derives Q2 from Q3. Since Hayes's moras function essentially as timing units, his mora deletion rule and the X-slot deletion rule proposed here are not so dissimilar. The crucial difference concerns the representations. Hayes does not demonstrate how it is possible to represent words like Q2 [aitta] and Q2 [kootti] in his system, but such representations are readily available in the present model. The only question concerns why these Q2 forms do not undergo degemination.



In the partitive case, the suffix brings along an extra  $\pi$ -node, so that the  $\pi$ -node linked to the first X-slot of the geminate is only shared with one additional X-slot:



This fails to feed Degem, leaving a Q3 bisyllable with a medial geminate:



Thus the rule of prosodic degemination accounts for the alternation between short and long consonants that accompanies a shift from Q3 to Q2.

### 5.3.6 Prosodic epenthesis

A similar pattern of prosodic alternation is shown by the nominals in (84). For this class, a Q2 genitive and Q3 partitive alternate with a Q1 bisyllabic nominative, rather than a Q3 monosyllabic nominative.

(84)	Nominative	Genitive	Partitive	
	Q1	Q2	Q3	
a.	sōper	sōpra	sōp:ra	‘friend’
b.	puter	putru	put:ru	‘porridge’
c.	vakel	vakla	vak:la	‘grub’

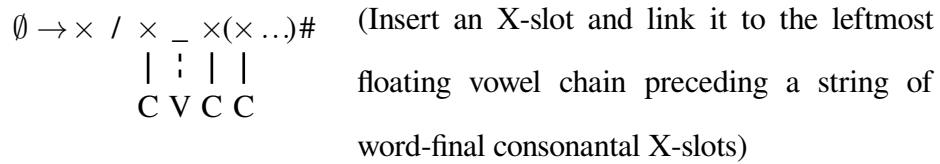
Note that in words of this class, one vowel appears word-finally in the genitive and partitive forms, and a different vowel appears between the final two consonants in the nominative form; the pattern is  $CVCV_{\alpha}C \sim CVCCV_{\beta} \sim CVC:CV_{\beta}$ .<sup>15</sup> I propose that these are represented underlyingly on roots as two separate floating vowels—one between the final two consonants and one word-finally:

(85)	a.	‘friend’	b.	‘porridge’	c.	‘grub’
		$\pi \pi$		$\pi \pi$		$\pi \pi$
		$\times \times \times \times$		$\times \times \times \times$		$\times \times \times \times$
		s <u>ō</u> p <u>e</u> r <u>a</u>		p u t <u>e</u> r <u>u</u>		v a k <u>e</u> l <u>a</u>

The first floating vowel can only be realized as a segment when it is hosted by an X-slot, and since it is non-final and the nominative is unsuffixed, this X-slot cannot be introduced by way of linear concatenation of a suffix. I propose that this vowel is realized via application of the rule in (86), which inserts an X-slot and links it to a medial floating vowel feature chain before one or more word-final consonantal X-slots.

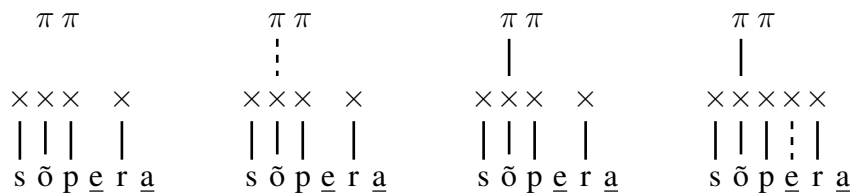
<sup>15</sup>The vowel that appears as  $V_{\alpha}$  behaves similarly to the so-called “ghost segments” seen with vowel–zero alternations in Slavic languages (see, for example, Szpyra 1992).

(86) *Realize medial floating vowel through X-slot insertion (V-epenth)*



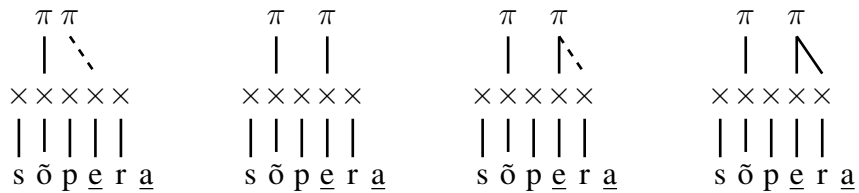
In the unsuffixed nominative case, the environment for V-epenth is present because the first floating vowel sits before a word-final consonantal X-slot:

(87) a. ‘friend’      b.  $\pi$ 1-to-V1      c. (X-to-V)      d. V-epenth



The root’s second  $\pi$ -node attaches to the newly inserted X-slot, which bears a vocalic feature chain. This results in a bisyllabic word with initial Q1:

(87) e.  $\pi$ s-to-Vs      f. ( $\pi$ -to- $\pi$ \_)      g.  $\pi$ -spread      h. (Degem)

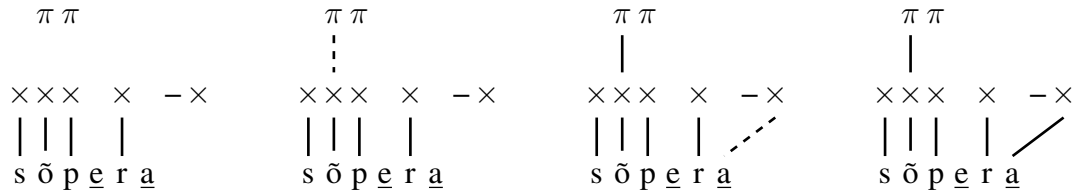


Q1 [sðper]  
‘friend (nom.)’

The derivation for the genitive form proceeds differently. After initial  $\pi$ -node association, the X-slot of the suffix links to the second floating vowel. X-to-V causes the X-slot of the suffix to link to the root-final floating vowel. This means that the first floating vowel no longer sits before a word-final consonant or consonant cluster, and the application of V-epenth is blocked:

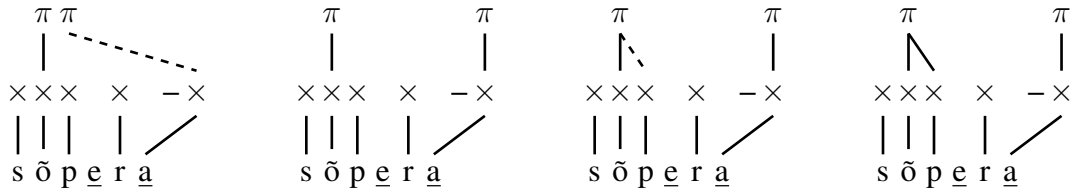


- (88) a. ‘friend + GEN’      b.  $\pi$ 1-to-V1      c. X-to-V      d. (V-epenth)



Instead of linking to an epenthetic X-slot, the second  $\pi$ -node links to the X-slot of the genitive suffix. The root’s first  $\pi$ -node then spreads to the X-slot linked to /p/, and the word is realized in Q2:

- (88) e.  $\pi$ s-to-Vs      f. ( $\pi$ -to- $\pi$ \_)      g.  $\pi$ -spread      h. (Degem)

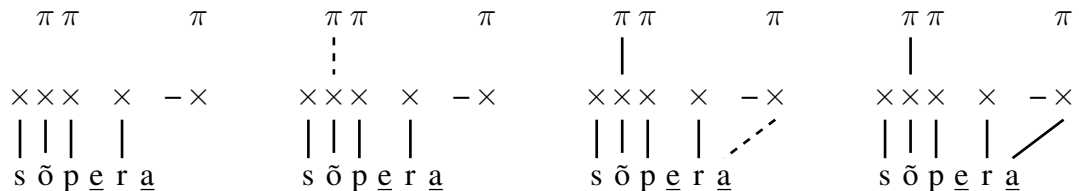


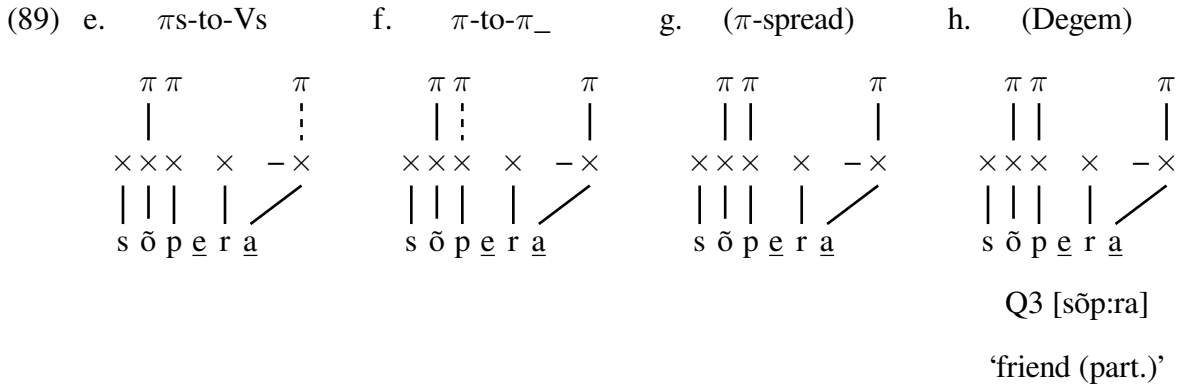
Q2 [sōpra]

‘friend (gen.)’

The partitive case is similar to the genitive, except it is the  $\pi$ -node of the suffix that attaches to the word-final vocalic X-slot. The root’s second  $\pi$ -node then links to the X-slot bearing /p/ by  $\pi$ -to- $\pi$ \_, and the word is realized in Q3 rather than Q2:

- (89) a. ‘friend + PART’      b.  $\pi$ 1-to-V1      c. X-to-V      d. (V-epenth)





As is clear from the demonstration above, only a phonological process of prosodic epenthesis must be introduced in order to account for an additional pattern of morphophonological marking. This class of nominals selects for the same suffix allomorphs needed for canonical Q3–Q2–Q3 alternations.

### 5.3.7 Case allomorphy

In the analysis presented thus far, different patterns of prosodic alternations were accounted for using a single set of phonological processes and a single phonological shape for each suffix (an X-slot for the genitive and an X-slot plus a  $\pi$ -node for the partitive). The behaviour of different classes of roots was captured by incorporating different kinds of abstraction into their underlying forms. I will now show that an even greater number of prosodic alternations can be accounted for if it is assumed that roots select for different case suffix allomorphs, which can have different phonological shapes. I will remain agnostic as to the precise morphosyntactic mechanism governing allomorph selection, but will assume that it is specified as a lexical property of the root morpheme and is not predictable (or at least not entirely predictable) based on the phonological shape of the root.

Consider the data in (90). Roots of this class retain their prosodic categories regardless of affixation, but it is clear that they are still derived from the same principles as the roots examined above. They all have an unpredictable final vowel which is not realized in the nominative case, so it must be underlyingly present on the root as a floating feature chain, which is only realized when

the root is combined with a suffix that contains an empty X-slot. Notice also that the partitive case has additional segmental marking in the form of a long /t/.

(90)		Nominative	Genitive	Partitive	
a.	Q1	jumal	jumala	jumalatt	‘god’
b.	Q2	raamatt	raamattu	raamattutt	‘book’
c.	Q3	alt:tar	alt:tari	alt:taritt	‘altar’

The fact that the roots in (90) do not alternate in quantity indicates that they retain the same number of  $\pi$ -nodes in all of their forms. In other words, both the genitive and the partitive suffix must have their own  $\pi$ -nodes, since no  $\pi$ -node is ever “borrowed” from the root. The underlying forms of roots are specified for segmental length distinctions and contain the number of  $\pi$ -nodes needed to realize the correct degree of quantity seen in the unsuffixed nominative (two for Q1 and Q2, three for Q3). They also contain the floating vowels realized in the non-nominative forms:

(91)	a.	‘god’	b.	‘book’	c.	‘altar’
		$\pi \pi$		$\pi \pi$		$\pi \pi \pi$
		×××××		×××××××		××××××
		j u m a l <u>a</u>		r a m a t <u>u</u>		a l t a r <u>i</u>

The roots in (91) select for a genitive allomorph that has both a bare X-slot and a  $\pi$ -node, as shown in (92). I notate this morpheme as ‘gen2’ in order to distinguish it from the basic genitive allomorph given in (74) above. Note that this suffix’s phonological form—an X-slot and a  $\pi$ -node—is phonologically identical (homophonous, in a sense) to that of the partitive allomorph used above.

(92)	‘GEN2’
	$\pi$
	– ×

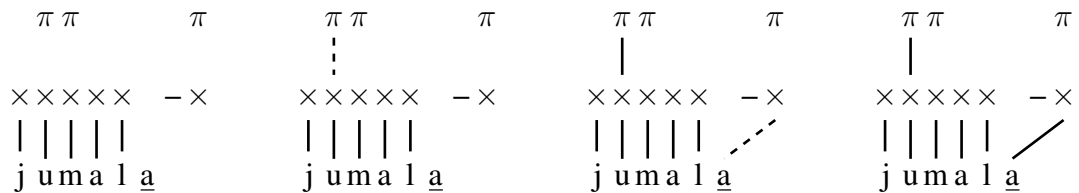
The partitive allomorph used with the roots in (91) must also have an X-slot (to realize the floating vowel) and a  $\pi$ -node (to preserve the quantity of the root), but it additionally has a fully specified segmental component: a long /t/, as in (93). As with gen2, I refer to this morpheme as ‘part2’ to distinguish it from the X+ $\pi$  partitive allomorph used above.

(93) ‘PART2’

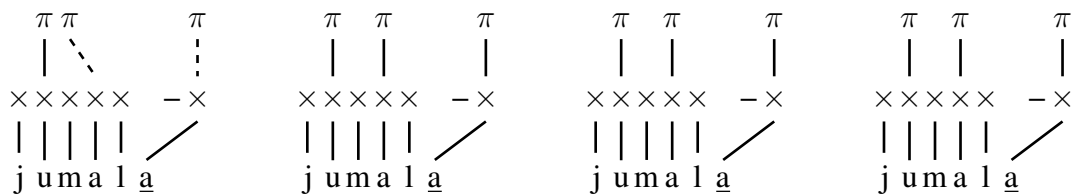


After the first  $\pi$ -node associates to the initial vocalic X-slot and the suffix’s X-slot links to the floating vowel,  $\pi$ s-to-Vs causes the remaining  $\pi$ -nodes to link to one X-slot each. The result is a trisyllabic word in Q1, with three short vowels:

(94) a. ‘god + GEN2’      b.  $\pi$ 1-to-V1      c. X-to-V      d. (V-epenth)



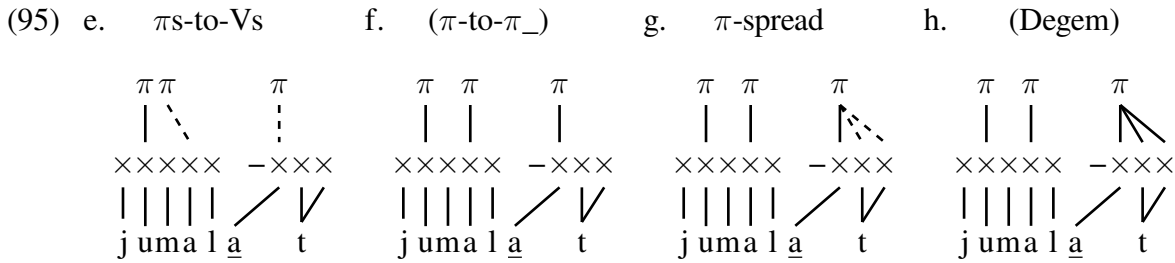
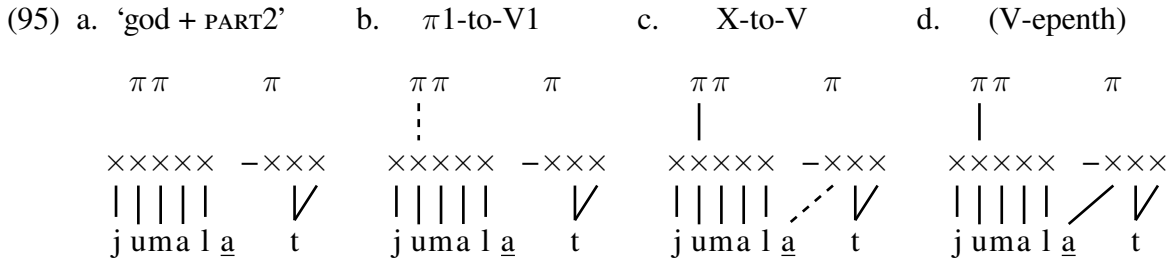
(94) e.  $\pi$ s-to-Vs      f. ( $\pi$ -to- $\pi$ \_)      g. ( $\pi$ -spread)      h. (Degem)



Q1 [jumala]

‘god (gen.)’

The partitive is derived identically, with the exception of  $\pi$ -spread causing the suffix’s  $\pi$ -node to associate with the two X-slots, which represent the suffix’s long /t/:

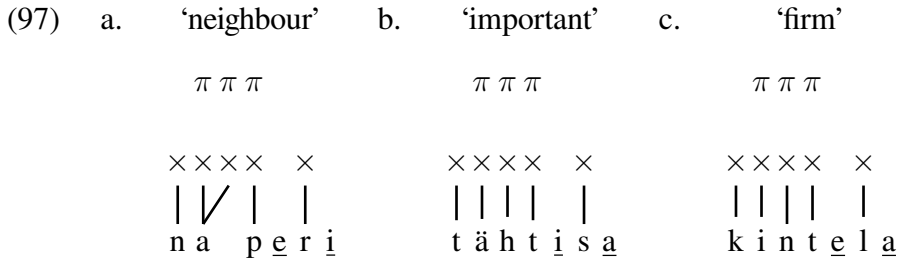


Q1 [jumalatt]  
 ‘god (part.)’

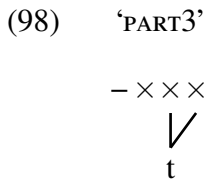
A similar but distinct class of nominals is illustrated in (96). Like the class that was above used as motivation for X-slot insertion by V-epenth, these words each show two separate unpredictable vowels: one in the nominative, and another in the genitive and partitive. However, unlike the words examined earlier, but much like the ones just analyzed, those in (96) do not show quantity alternations throughout the paradigm.

(96)	Nominative	Genitive	Partitive	
a. Q3	naa:per	naa:pri	naa:pritt	‘neighbour’
b. Q3	täh:tis	täh:tsa	täh:tsatt	‘important’
c. Q3	kin:tel	kin:tla	kin:tlatt	‘firm’

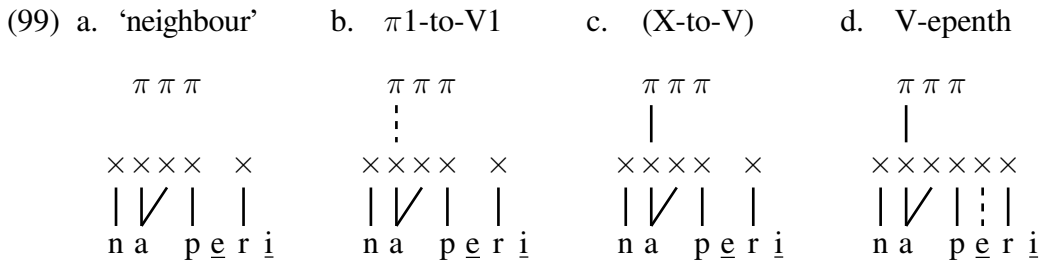
In order for such words to realize Q3 in their unsuffixed forms, the roots must contain three  $\pi$ -nodes each—two to link to adjacent X-slots, starting with the initial vowel, and one to realize the last vocalic X-slot in the word. Underlying forms for these roots are shown in (97). In the nominative, this X-slot will be inserted by V-epenth and linked to the first of the two floating vowels. In order to maintain a word with only three  $\pi$ -nodes, the genitive and partitive suffixes cannot contain  $\pi$ -nodes of their own, but must have X-slots to realize the final floating vowel and, in the case of the partitive, a segmental long /t/.



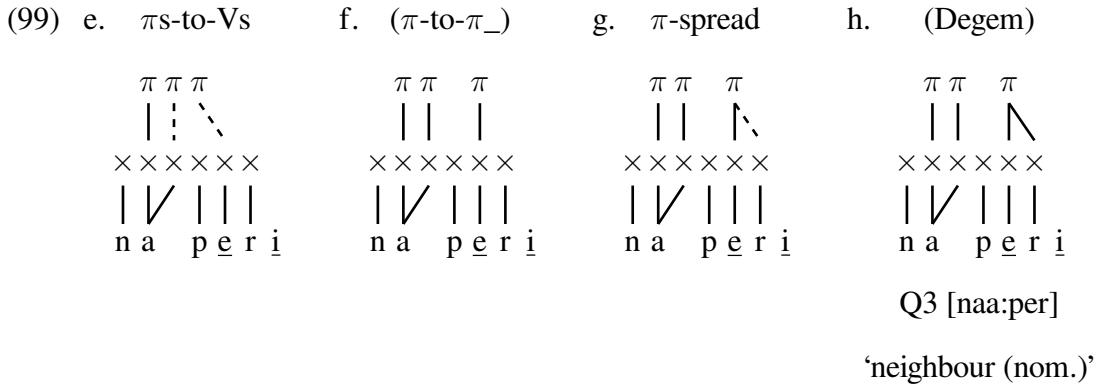
A genitive allomorph of the right phonological shape already exists; the basic genitive suffix proposed for Estonian contains only a bare X-slot. In order to account for the partitive forms, I posit an allomorph ‘part3’, given in (98).<sup>16</sup>



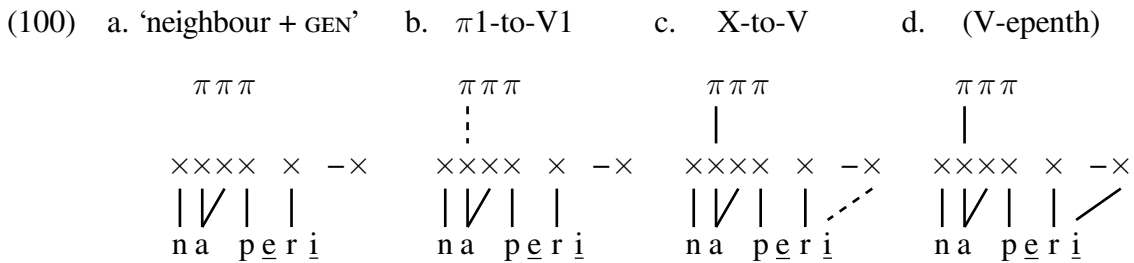
The unsuffixed nominatives of such words contain three  $\pi$ -nodes in the root alone. An X-slot is inserted by V-epenth to realize the first floating vowel, and the third floating  $\pi$ -node links to it. The second  $\pi$ -node is then realized adjacent to the first, and the result is a bisyllable in Q3 with a word-final consonant:



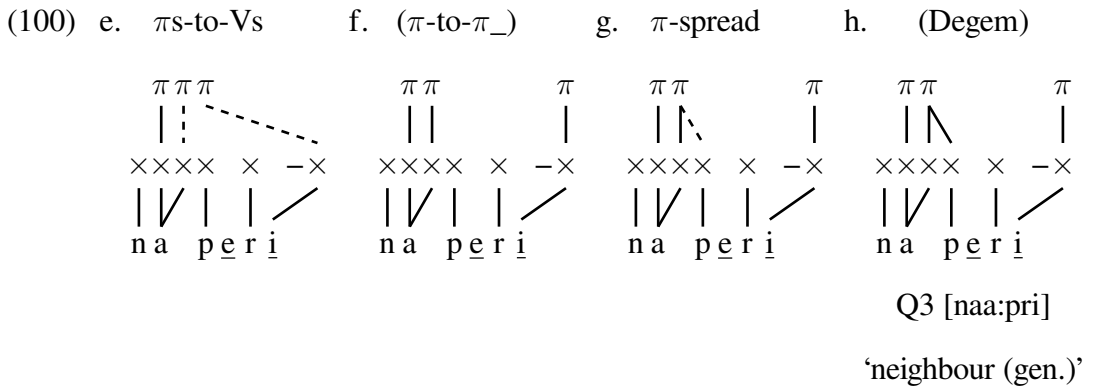
<sup>16</sup>Notice that the present analysis assumes that a given root can select for any genitive allomorph and any partitive allomorph. Thus the roots in the preceding subsections select for the basic X-slot genitive and the X+ $\pi$  partitive, the prosody-stable roots in (90) select for X+ $\pi$  gen2 and X+ $\pi$ +tt part2, and the roots in (96) select for the basic X-slot genitive and X+tt part3.



The genitive suffix provides an X-slot which can link to the final floating vowel via X-to-V, blocking the application of V-epenth:



The root’s third  $\pi$ -node links to this X-slot rather than an inserted interconsonantal one, but the result is still a bisyllable in Q3, albeit one without a word-final consonant:



The partitive is derived in exactly the same way, since the partitive allomorph used here (that given in 98) has an X-slot but lacks its own  $\pi$ -node, just like the genitive. The only difference is that the word-final long /t/ is linked to the third  $\pi$ -node by  $\pi$ -spread.

We have seen above that additional patterns of morphophonological alternations can be accounted for simply by introducing new suffix allomorphs. The morphologically-sensitive component of the grammar is limited to allomorph selection—the phonology interprets underlying representations using a single set of phonological processes, making no reference to morphological categories.

### 5.3.8 Optional allomorphy

That the phonological shapes of different suffix allomorphs play a role in triggering different quantity alternations is shown by the data in (101). These words have roots with two different floating vowels, the first realized in the Q2 nominative by V-epenth and the second realized in Q3 by affixing *gen2* ( $X+\pi$ ) and *part2* ( $X+\pi+tt$ ) in the genitive and partitive, respectively. However, these roots optionally take another partitive allomorph, which consists only of segmental short /t/, leaving the word in Q2 and realizing the first floating vowel.

(101)	Nominative	Genitive	Partitive	
	Q2	Q3	Q3/Q2	
a.	ihnus	ih:n:sa	ih:n:satt/ihnust	‘stingy’
b.	õilis	õi:l:sa	õi:l:satt/õilist	‘magnanimous’
c.	tõrkes	tõr:k:sa	tõr:k:satt/tõrkest	‘reluctant’

The underlying forms for these roots, then, are as follows:

(102)	a.	‘stingy’	b.	‘magnanimous’	c.	‘reluctant’
		$\pi \pi$		$\pi \pi$		$\pi \pi$
		$\times \times \times \times$		$\times \times \times \times$		$\times \times \times \times \times$
		i h n <u>u</u> s <u>a</u>		õ i l <u>i</u> s <u>a</u>		t õ r k <u>e</u> s <u>a</u>

The segmental /t/ partitive morpheme, ‘part4’, has the underlying form in (103). It consists of only the segmental feature chain for /t/ linked to a single X-slot.

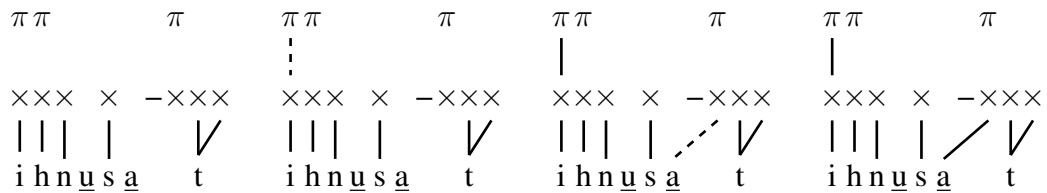


(103) ‘PART4’

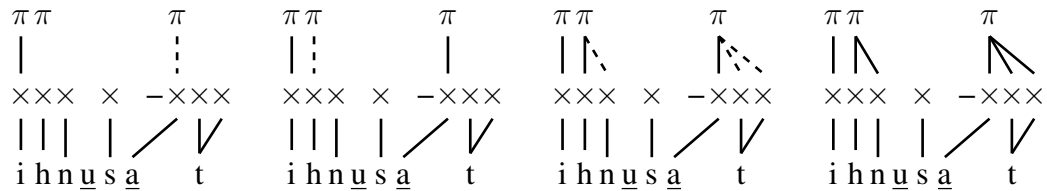


In the genitive (and the X+ $\pi$ +tt partitive), the second floating vowel links to the bare X-slot of the suffix, creating a position to which the suffix’s  $\pi$ -node can link. The second  $\pi$ -node of the suffix must then link within the stem adjacent to the first. This yields a bisyllable in Q3 realizing its second floating vowel (the nominative is not monosyllabic):

(104) a. ‘stingy + PART2’    b.  $\pi$ 1-to-V1    c. X-to-V    d. (V-epenth)



(104) e.  $\pi$ s-to-Vs    f.  $\pi$ -to- $\pi$ \_    g.  $\pi$ -spread    h. (Degem)

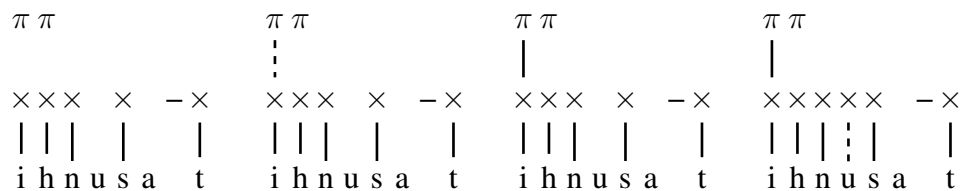


Q3 [ih:n:satt]

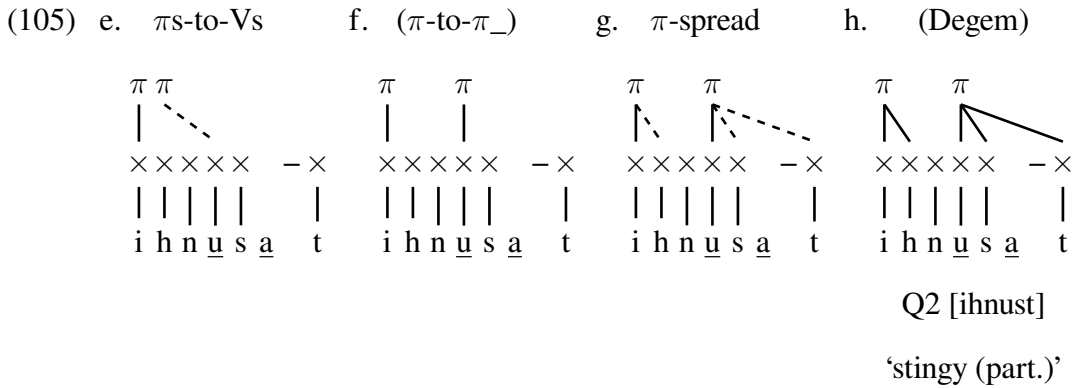
‘stingy (part.)’

In the optional form that takes part4, no bare X-slot is available to realize the second floating vowel. However, the environment for V-epenth is present, so an X-slot is inserted and linked to the first floating vowel:

(105) a. ‘stingy + PART4’    b.  $\pi$ 1-to-V1    c. (X-to-V)    d. V-epenth



The root's second  $\pi$ -node links to this X-slot, and the  $\pi$ -nodes spread rightward by application of  $\pi$ -spread. This yields a bisyllable in Q2 whose partitive is marked with only segmental /t/:



Thus two optional partitive suffix allomorphs cause the root to surface in different quantity degrees, because one (part2) brings along a  $\pi$ -node while the other (part4) consists only of segmental material.

### 5.3.9 Q2–Q3–Q2 alternations

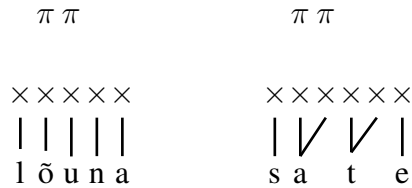
A final class of nominals relevant to a discussion of quantity alternations shows the pattern in (106). Here, it is the nominative and the partitive that are in Q2, while the genitive is in Q3—the exact opposite of the basic pattern. This class is different in another way from basic Q3–Q2–Q3 nominals: all three of the principle forms have the stem-final vowel, so the nominative is not monosyllabic.

(106)	Nominative	Genitive	Partitive	
	Q2	Q3	Q2	
a.	lõuna	lõu:na	lõunatt	'south'
b.	saate	saa:t:te	saatett	'accompaniment'

Based on the assumption that the nominative is derived from the unsuffixed root, we arrive at the lexical entries in (107). These roots contain two  $\pi$ -nodes, since their nominatives are realized

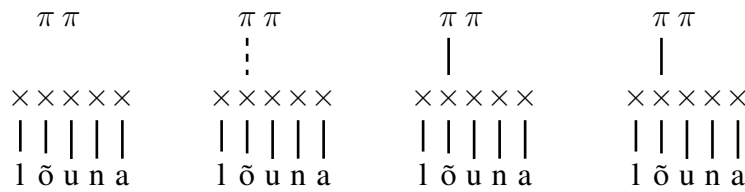
in Q2, and their final vowels are underlyingly linked to root-final X-slots, since they never fail to surface.<sup>17</sup>

(107) a. ‘south’      b. ‘accompaniment’

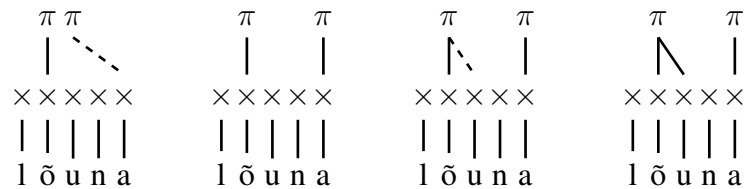


The derivation for the nominative simply shows initial  $\pi$ -node association, linking of the second  $\pi$ -node to the final vowel’s X-slot, and spreading of the initial  $\pi$ -node:

(108) a. ‘south’      b.  $\pi$ 1-to-V1      c. (X-to-V)      d. (V-epenth)



(108) e.  $\pi$ s-to-Vs      f. ( $\pi$ -to- $\pi$ \_)      g.  $\pi$ -spread      h. (Degem)



Q2 [lõuna]

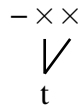
‘south (nom.)’

<sup>17</sup>Some words of this (or a closely related) class undergo other segmental alternations, for example showing an extra /s/ between the root and suffix, as in (i). Like other segmental alternations, I set these aside here.

(i)	Nominative	Genitive	Partitive	
	Q2	Q3	Q2	
	a. võðras	võð:ra	võðrast	‘foreign’
	b. rikkas	rik:ka	rikkast	‘rich’
	c. taevas	tae:va	taevast	‘sky’

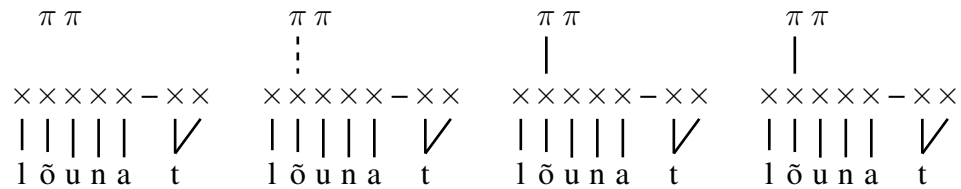
Partitive forms of this class do not differ prosodically from the unsuffixed nominative, being marked only segmentally with a long /t/. Because there is no need to realize a floating vowel, the partitive suffix must be the allomorph ‘part5’, which contains only a segmental feature chain for /t/ linked to two X-slots:

(109) ‘PART5’

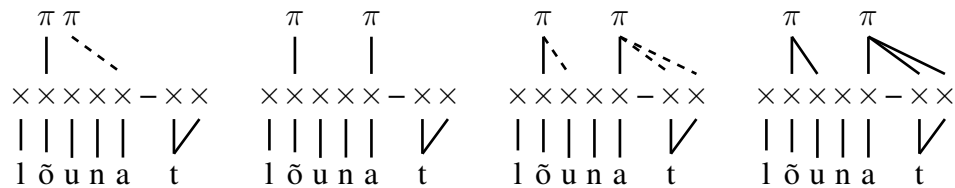


This yields a partitive derivation that differs from that of the nominative only in the application of  $\pi$ -spread, which now includes the two X-slots linked to the /t/:

(110) a. ‘south + PART5’    b.  $\pi$ 1-to-V1    c. (X-to-V)    d. (V-epenth)



(110) e.  $\pi$ s-to-Vs    f. ( $\pi$ -to- $\pi$ \_)    g.  $\pi$ -spread    h. (Degem)



Q2 [lönatt]

‘south (part.)’

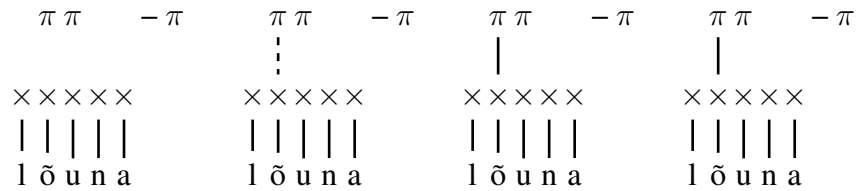
In contrast to the partitive, the genitive does trigger a quantity alternation. It does not, however, differ segmentally from the nominative case, so the phonological form of the genitive allomorph used here must consist of only the difference between the Q2 nominative and the Q3 genitive: a single  $\pi$ -node, like ‘gen3’ in (111).

(111) ‘GEN3’

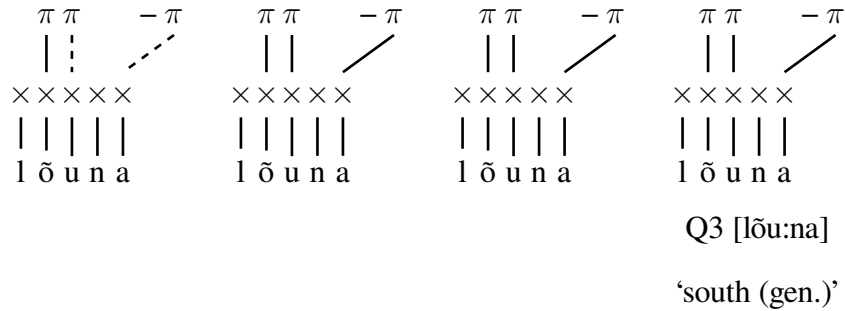


When the suffix in (111) is added to the same root, it is the suffix's  $\pi$ -node that links to the root-final vocalic X-slot. This leaves the root's second  $\pi$ -node to link adjacent to the first  $\pi$ -node, leading to a bisyllable in Q3 rather than Q2:

- (112) a. 'south + GEN3'    b.  $\pi$ 1-to-V1    c. (X-to-V)    d. (V-epenth)

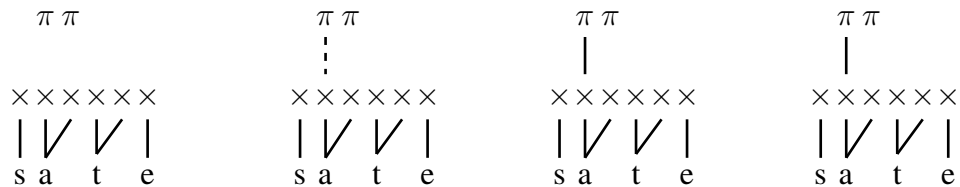


- (112) e.  $\pi$ s-to-Vs    f. ( $\pi$ -to- $\pi$ \_)    g. ( $\pi$ -spread)    h. (Degem)

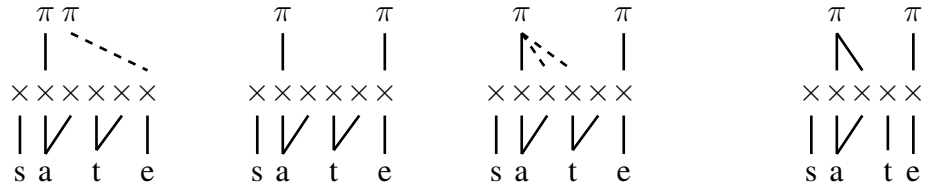


This analysis also accounts for the consonant length alternation observed for roots of this class with medial geminates. In the nominative and partitive cases, only the two  $\pi$ -nodes of the root link to the X-tier, and the environment for Degem exists, leading to degemination. Thus the unsuffixed nominative is not segmentally identical to the root's underlying form:

- (113) a. 'accompaniment'    b.  $\pi$ 1-to-V1    c. (X-to-V)    d. (V-epenth)



(113) e.  $\pi$ s-to-Vs      f.  $(\pi$ -to- $\pi$ \_)      g.  $\pi$ -spread      h. Degem

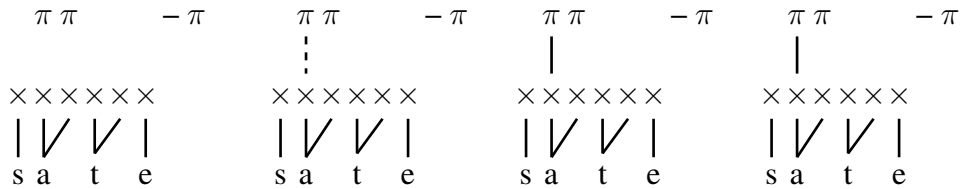


Q2 [saate]

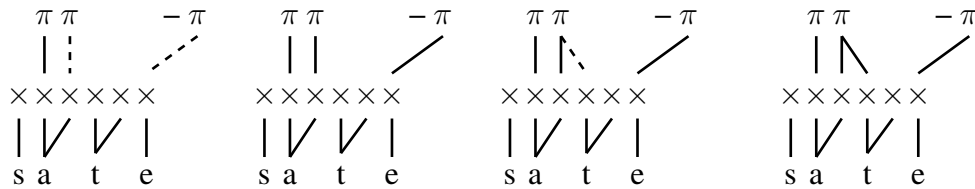
‘accompaniment (nom.)’

In the genitive case, however, the presence of an extra  $\pi$ -node brought by the suffix blocks the application of Degem, deriving a surface form that is segmentally identical to (but on the  $\pi$ -tier, prosodically different from) the underlying form of the root:

(114) a. ‘accomp. + GEN3’      b.  $\pi$ 1-to-V1      c. (X-to-V)      d. (V-epenth)



(114) e.  $\pi$ s-to-Vs      f.  $(\pi$ -to- $\pi$ \_)      g.  $\pi$ -spread      h. (Degem)



Q3 [saa:t:te]

‘accompaniment (gen.)’

We have seen here that by introducing a genitive case allomorph that brings along a  $\pi$ -node of its own, the genitive case can be made to trigger Q3 in a certain class of nominals. Furthermore, because of the purely phonological formulation of prosodic degemination, this class of nominals also undergoes long–short consonant alternations in its non-Q3 forms (the nominative and the partitive), rather than in the genitive, where it is typically seen to apply.

### 5.3.10 The short illative

In addition to the nominative, genitive, and partitive forms, the illative case can also trigger prosodic alternations. Like the majority of cases in Estonian, the illative is segmentally marked and is built on the genitive stem. It typically has the shape /-sse/. However, a subset of nouns can also take a short form of the illative (sometimes referred to as the “additive”). Rather than being added linearly to the weak grade stem, the short illative appears as the strong grade stem, and thus is often homophonous with the partitive:

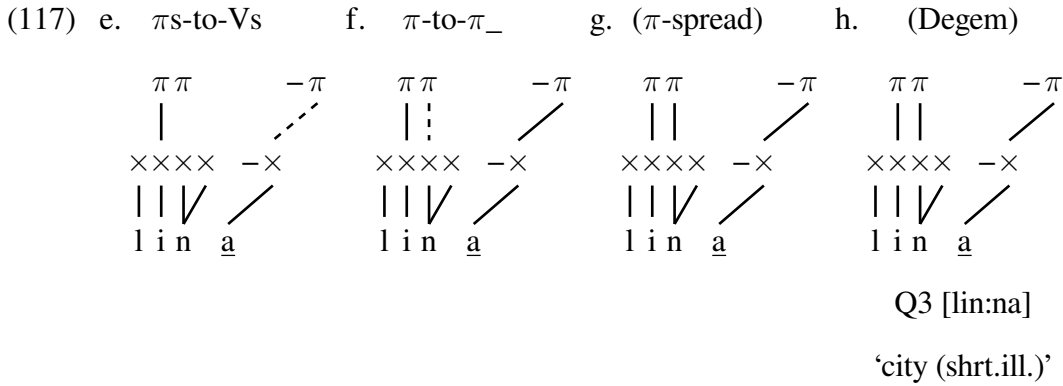
(115)	Long Ill.	Short Ill.	(Partitive)	
	Q2	Q3	Q3	
a.	linnasse	lin:na	lin:na	‘city’
b.	lippusse	lip:pu	lip:pu	‘flag’
c.	koolisse	koo:li	koo:li	‘school’

We know that the genitive suffix for this class of nouns consists of only an X-slot, and that the partitive suffix contains an X-slot and a  $\pi$ -node. If the total suffixed material in the short illative is equivalent to the partitive, then the short illative must contain only a  $\pi$ -node, as in (116):

(116)	‘SHRT.ILL’
	– $\pi$

The derivation of the short illative is the same as that of the partitive. The only difference is that the bare X-slot and the  $\pi$ -node that are not part of the root come from different suffixes. The genitive’s X-slot links to the floating vowel, and it is the short illative’s  $\pi$ -node that links to this X-slot:

(117)	a. ‘city + GEN + SHRT.ILL’	b. $\pi$ 1-to-V1	c. X-to-V	d. (V-epenth)
	$\pi \pi$ – $\pi$	$\pi \pi$ – $\pi$	$\pi \pi$ – $\pi$	$\pi \pi$ – $\pi$
	xxxxx –x	xxxxx –x	xxxxx –x	xxxxx –x
	/	/	/	/
	l i n <u>a</u>	l i n <u>a</u>	l i n <u>a</u>	l i n <u>a</u>

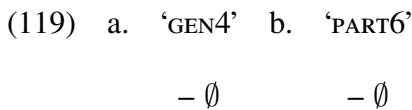


There is also a class of nominals which are prosodically and segmentally identical in the nominative, genitive, and partitive, all of them being of the shape CVCV:

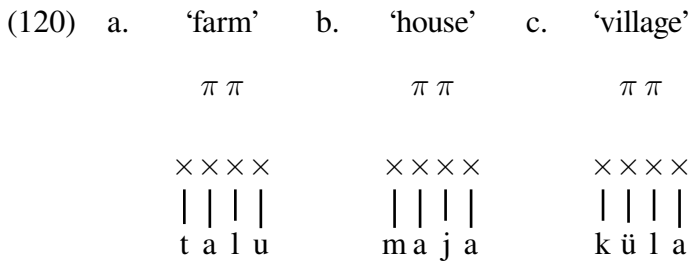
(118)

	Nominative	Genitive	Partitive	
	Q1	Q1	Q1	
a.	tal <u>u</u>	tal <u>u</u>	tal <u>u</u>	'farm'
b.	maj <u>a</u>	maj <u>a</u>	maj <u>a</u>	'house'
c.	kü <u>l</u> a	kü <u>l</u> a	kü <u>l</u> a	'village'

These roots, of course, must select for null allomorphs of the genitive and partitive:



The underlying forms of the roots are correspondingly simple; they contain four X-slots and two  $\pi$ -nodes, with no floating vowels:



Interestingly, this class of roots *does* show a quantity alternation in the short illative. This alternation involves both gemination of the medial consonant and a change to Q3:

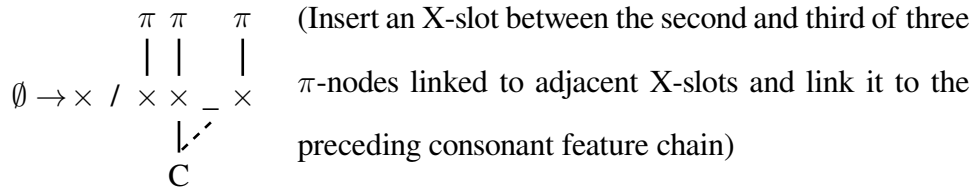


(121)      Long Ill.    Short Ill.

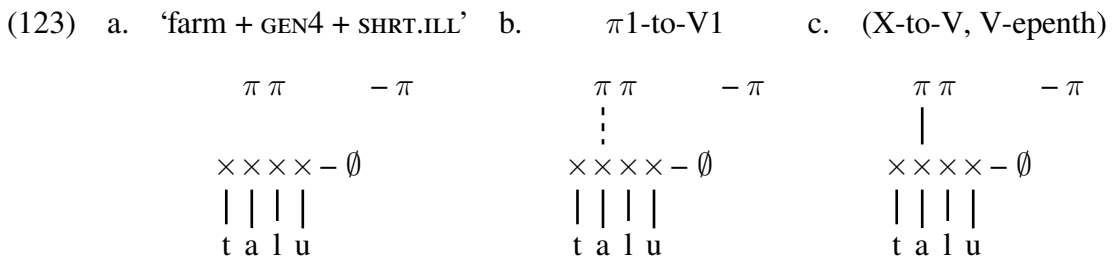
	Q1	Q3	
a.	talusse	tal:lu	‘farm’
b.	majasse	maj:ja	‘house’
c.	külasse	kül:la	‘village’

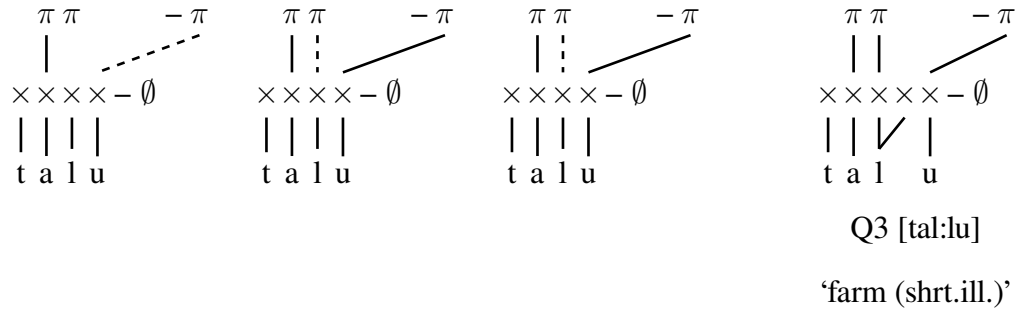
I propose that the short illative morpheme for CVCV roots is the same as that given above, and that it, too, attaches to the genitive stem. Medial gemination is caused by the rule in (122), which inserts an X-slot and links it to the preceding consonantal feature chain in order to break up a string of three  $\pi$ -nodes linked to adjacent X-slots. This provides a position for the  $\pi$ -node to link to and allows realization in Q3.

(122) *Medial consonant gemination (C-gem)*



When the short illative is derived, the second suffix’s  $\pi$ -node links to the final vowel’s X-slot. The second  $\pi$ -node of the root is then left with only one position to associate with, and  $\pi$ -to- $\pi$  causes it to do so. However, this creates the environment for C-gem, which inserts an X-slot and links it to the preceding consonant, resulting in a Q3 bisyllable with a medial geminate:



(123) d.  $\pi$ S-to-Vs      e.  $\pi$ -to- $\pi$ \_      f. C-gem      g. ( $\pi$ -spread, Degem)

Here we have seen that a short illative morpheme consisting of only a  $\pi$ -node can be used for bisyllabic nouns, creating a form which is, for most words, homophonous with the partitive. By introducing a consonant gemination process, the same short illative morpheme can be made to trigger an idiosyncratic Q1–Q3 alternation for a class of nominals whose other basic forms are zero-marked.

### 5.3.11 Summary

In this section, I have analyzed a number of nominal paradigms and their associated quantity alternations in Estonian. It is clear that, despite the complicated data, an elegant and morphologically compositional analysis of Estonian quantity is possible within the set of assumptions needed for a unified model of word prosodic contrasts. While I have by no means offered a comprehensive account of all aspects of Estonian quantity—even a modest attempt would easily fill an entire volume—it is clear that the present model has tools abstract enough to account for the prosodic system of the language without needing to posit additional formal primitives or stipulate a large number of morphophonological rules. With the present understanding of word prosody, Estonian is nothing special in terms of the formal mechanisms involved.

A possible objection to the analysis put forth here is that it employs a large number of genitive and partitive suffix allomorphs, as summarized in Table 5.1. However, it is unavoidably the case that there are many different patterns of prosodic alternations, and any account needs to accommodate each of them in some way. Rather than positing a large number of morphophonological

		‘GEN’	‘GEN2’	‘GEN3’	‘GEN4’	
			$\pi$	$-\pi$		
		$-\times$	$-\times$		$-\emptyset$	
‘PART’	‘PART2’	‘PART3’	‘PART4’	‘PART5’	‘PART6’	
$\pi$	$\pi$					
$-\times$	$-\times\times\times$	$-\times\times\times$	$-\times$	$-\times\times$	$-\emptyset$	
	$\checkmark$	$\checkmark$	$ $	$\checkmark$		
	t	t	t	t		

Table 5.1: Summary of Estonian genitive and partitive suffix allomorphs

rules, I have advocated for an approach in which variation in phonological behaviour is due to differences in the lexical input. This makes it possible to posit a single phonological grammar based on standard autosegmental operations, keeping the arbitrariness in lexical entries where it belongs. It is my hope that, in addition to providing further evidence for the unified model of word prosodic contrasts defended in the present study, this work will make new headway in developing a fuller understanding of Estonian phonology on the whole.

## 5.4 The diversity of privative $\pi$

In this chapter, I have examined the prosodic systems of two very different languages: Papiamentu and Estonian. Papiamentu is said to have independent tone and stress contrasts, a combination which is nearly unattested and has always required special treatment. Estonian, on the other hand, has a relatively rare surface three-way length contrast, and this, too, has required special consideration whenever it has been addressed in the literature on quantity.

At the level of abstraction taken in the current study, however, these languages have much in common. Both employ featureless  $\pi$ -nodes and a quantity contrast on the X-tier, and both allow suffixes that consist only of these floating prosodic elements. While a great deal of variation

exists with respect to language-specific phonetic realizations, the formal mechanisms at play are identical. Thus, although Papiamentu and Estonian seem to possess quite exotic prosodic systems, neither is more marked typologically than a privative tone language with an independent quantity contrast.

# Chapter 6

## Conclusions

It has long been understood that word prosodic features—tone, stress, and length—behave differently from segmental features, and accordingly prosody has been represented using different kinds of models. However, prosodic features have traditionally been posited based on (often impressionistic) surface phonetics and accounted for with disparate models. The result is that no clear predictions can be made about how different prosodic features should interact or how many of them can coexist in a single system.

In this thesis, I have developed a unified and restrictive model of word prosodic contrasts, capturing a diverse range of systems with a single set of abstract representational elements based on formal properties that are well established. My proposal draws on insights from work that assumes that phonological representations are limited by phonological contrasts and are not phonetically deterministic (e.g., Morén 2003; Blaho 2008; Dresher 2009, 2014), as well as Hyman's (2006) proposal that word prosodic typology should be based on phonological behaviour rather than phonetic properties.

The model consists of two tiers of root nodes: X-slots, which host segmental contrastive features and are double linked to represent segmental length, and  $\pi$ -nodes, which host prosodic contrastive features and have language-specific phonetic realizations. In lexical stress accent languages and privative tone languages, featureless  $\pi$ -nodes act as privative elements in a model that

otherwise operates on binary features. For languages with only one of contrastive tone, stress, or length, this model provides notational variants of standard analyses under a single set of assumptions, capturing the “easy” cases in a unified way. More importantly, as I have demonstrated in four case studies, it allows compositional analyses of apparently non-concatenative prosodic phenomena, such as stress shifts and alternations in tone and quantity, in languages with mixed prosodic systems. Finally, because of the restrictions on the complexity of the representations, all of this is done without predicting the greater degree of complexity that would be expected to be possible in a single system if it were assumed that all of the standard prosodic representational mechanisms could be simultaneously active in a language.

This is accomplished because one tier of root nodes is needed to represent each independent prosodic contrast. The X-tier canonically represents quantity contrasts, and the  $\pi$ -tier represents tonal autosegments (in non-accentual tone languages) and accented positions (in lexical stress accent languages). It is possible for a language to have both lexical tone and lexical stress accent, as with Papiamentu, but such a language must represent “stress” as covert quantity on the X-tier. The consequence of this is that the language has no means to represent a third independent prosodic feature of length.

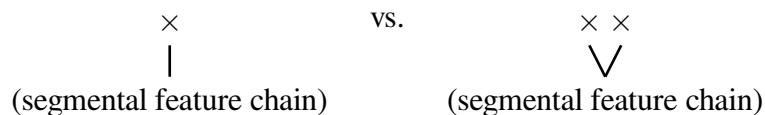
Also built into the model is a restriction on greater than binary linking from root node to feature chains, which means that segmental quantity (represented with X-slots) is maximally binary. This requires the ternary quantity contrast of Estonian to be split across both tiers of root nodes—a system which is justified by the behaviour of morphophonological quantity alternations in the language.

Finally, the model straightforwardly predicts a class of tonal accent languages, such as Goizueta Basque and Choguita Rarámuri, in which stress is lexically determined, and which allows a tonal contrast where the location of the tonal contrast is entirely dependent on the location of stress. In such languages, the  $\pi$ -node simultaneously serves as the autosegment marking the accented position and as the root node bearing the feature chains that divide the tonal inventory.

## 6.1 Towards a minimalist typology of prosodic systems

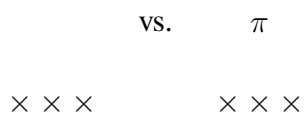
In Chapter 3 of this thesis, I developed a unified model of prosodic contrasts and showed its notational equivalents of standard cases of word prosody. Subsequently in Chapters 4 and 5, I demonstrated that this model allows for elegant and theoretically integrated analyses of an eclectic set of prosodic systems without requiring the introduction of ad hoc formal mechanisms. I will now consider more explicitly the range of diversity of prosodic systems predicted by the present model. In doing so, I will refer, as a term of convenience, to a number of “parameters”. A given language may either lack or have each parameter. By having a parameter, a particular formal configuration of X-slots,  $\pi$ -nodes, and contrastive feature chains is either ruled out or allowed. I will go through each of the major typological parameters in turn, followed by some examples of languages and their parameter settings.

(1) *Allow multiple X-slot to feature chain linking (useXX)?*



Languages with useXX are those that allow phonologically long segments; having useXX makes it possible to distinguish between one and two X-slots linking to a single segmental feature chain. Recall that linking more than two X-slots to a single feature chain is universally banned in the present model, so no upper bound of X-slot to feature chain linking needs to be specified for this parameter; it is simply binary.

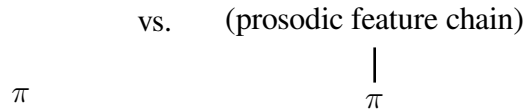
(2) *Employ a tier of  $\pi$ -nodes (use $\pi$ )?*



Languages with use $\pi$  are those that employ a tier of  $\pi$ -nodes in their contrastive representations. A language that has a  $\pi$ -tier is one which has an inventory of prosodic autosegments. These include but are not limited to, the usual typological categories of tone languages and lexical stress languages. The lack of the use $\pi$  parameter in a language entails that it has no settings

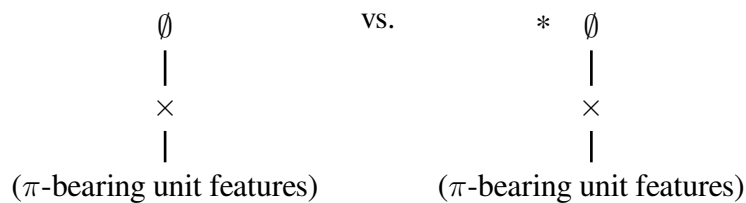
for any of the other parameters discussed below, which all define properties of  $\pi$ -nodes and are thus meaningless in the absence of  $\pi$ -nodes.

(3) *Contrastive features present on  $\pi$ -nodes ( $\pi$ feat)?*



Languages with  $\pi$ feat are those that have prosodic inventories containing more than one member. There must be an inventory of at least two prosodic autosegments in order for that inventory to be divided by specifications of binary contrastive features.

(4) *Require linking of  $\pi$ -nodes to all valid  $\pi$ -bearing units ( $\pi$ all)?*



Languages with  $\pi$ all require that, in the output of the lexical phonology, all positions that may bear  $\pi$ -nodes *do* bear  $\pi$ -nodes. This parameter accounts for, among other patterns, the requirement in many tone languages for all tone bearing units to be specified for tones in the output.

(5) *Culminative/obligatory  $\pi$ -nodes within the phonological word ( $\pi$ culm)?*

\*  $\pi \neq 1$  per word

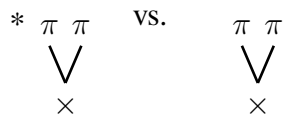
In languages with  $\pi$ culm,  $\pi$ -nodes are culminative/obligatory in the sense of Hyman (2006): each word must contain one and only one  $\pi$ -node linked to the X-tier. This parameter is active in, for example, languages with lexical stress accent. A single language may only have one of  $\pi$ all and  $\pi$ culm, since it is an inherent contradiction to require  $\pi$ -nodes on all valid positions and simultaneously allow only one  $\pi$ -node per word. The only place this contradiction would be impossible to detect is in a language where all words contain only a single  $\pi$ -bearing unit.



	GB	CR	Pap	Est	Fin	Sp	Pol
useXX		X	✓	✓	✓	X	X
use $\pi$	✓	✓	✓	✓	X	✓	X
$\pi$ feat	✓	✓	X	X		X	
$\pi$ all	X	X	X	✓		X	
$\pi$ culm	✓	✓	X	X		✓	
short $\pi\pi$			X	X			
spread $\pi$	X	X	X	✓		X	

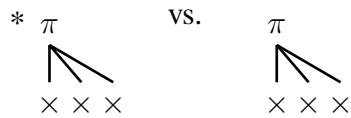
Table 6.1: Parameter settings for several languages. GB = Goizueta Basque, CR = Choguita Rarámuri, Pap = Papiamentu, Est = Estonian, Fin = Finnish, Sp = Spanish, Pol = Polish.

(6) *Multiple  $\pi$ -nodes may link to a single X-slot (short $\pi\pi$ )?*



Languages with short $\pi\pi$  allow two  $\pi$ -nodes to link to a single X-slot. For example, a tone language that has short $\pi\pi$  allows short composed contour tones. It is possible that this should be universally disallowed in languages that do not have  $\pi$ feat—a statement equivalent to saying that OCP for  $\pi$ -nodes is non-negotiable within a single  $\pi$ -bearing unit.

(7) *Multiple X-slots may share a single  $\pi$ -node (spread $\pi$ )?*



Languages with spread $\pi$  allow  $\pi$ -nodes to simultaneously link to more than one X-slot. The way that such representations are derived can, of course, vary from language to language.

Examples of the parameter settings in a number of languages are given in Table 6.1, beginning with the studies in Chapters 4 and 5. Here, parameters that the language has are marked with “✓” and those that the language does not have are marked with “X”. Parameters that cannot be determined due to the setting of another parameter are left blank.

In tonal accent languages like Goizueta Basque and Choguita Rarámuri, the presence of culminative/obligatory accent is due to the combination of  $\text{use}\pi$  and  $\pi\text{culm}$ . The tonal properties that are dependent on accent are allowed because those  $\pi$ -nodes can host features due to  $\pi\text{feat}$ .<sup>1</sup> These can be contrasted with a language like Spanish, which differs only in its setting for  $\pi\text{feat}$ —Spanish has a single-member prosodic inventory, consisting only of a bare  $\pi$ -node used to represent accent. Those  $\pi$ -nodes do not bear features because no paradigmatic contrast (tonal or otherwise) is dependent on the position of accent in Spanish.

In Papiamentu, phonological quantity is used to represent “stress”, so  $\text{useXX}$  is set to “✓”. The language also employs  $\pi$ -nodes to represent its tonal contrast, meaning that  $\text{use}\pi$  is set to “✓” as well, but since  $\pi$ -nodes do not bear features,  $\pi\text{feat}$  is set to “X”. Due to the morphemes involved, Papiamentu words generally contain one or zero  $\pi$ -nodes, although certain exceptional words may contain several consecutive  $\pi$ -nodes. Because of the lack of restrictions on how many  $\pi$ -nodes may be present in a word and where they may (or may not) link, both  $\pi\text{all}$  and  $\pi\text{culm}$  are set to “X”.

The Estonian quantity system is split across two tiers—segmental length on the X-tier and the long–overlong distinction on the  $\pi$ -tier. For this reason, both  $\text{useXX}$  and  $\text{use}\pi$  are set to “✓”. However,  $\pi$ -nodes cannot bear features, so  $\pi\text{feat}$  is set to “X”. The unique nature of the system comes from the lack of  $\pi$ -node features together with a requirement (due to  $\pi\text{all}$ ) for all valid X-slots to bear  $\pi$ -nodes. This motivates processes that cause  $\pi$ -nodes to link to the X-tier in order to realize segments, triggering long–overlong alternations when  $\pi$ -nodes are borrowed between roots and suffixes. Setting of  $\text{short}\pi\pi$  to “X” ensures that overlong (Q3) quantity is not possible when only one  $\pi$ -bearing unit is available: short quantity cannot alternate to long quantity without the insertion of an additional X-slot, as was seen in deriving the short illative forms of CVCV roots in Section 5.3.10.

---

<sup>1</sup>Note that Choguita Rarámuri is described by Caballero and Carroll (to appear) as having derived long vowels, which I have not addressed in my discussion of the language in Chapter 4. If these are in fact phonologically long vowels, then Choguita Rarámuri would of course have  $\text{useXX}$  set to “✓”.

For comparison with these more complicated systems, I also show the parameter settings for Finnish and Polish. Accent is contrastive in neither language, with stress being merely a surface phenomenon. Thus,  $\pi$ -nodes are not active in the grammar and  $\text{use}\pi$  is set to “X” in both, nullifying all subsequent parameters. The languages differ from each other only in their settings for  $\text{useXX}$ ; Finnish has a robust segmental length contrast, while Polish does not. Thus Finnish has  $\text{useXX}$  set to “✓” and Polish has it set to “X”.

The typology outlined here is by no means an exhaustive investigation of the patterns predicted by the present model, but merely a sketch of the kind of variation that the model predicts with a focus on the more complicated cases examined in this study.

## 6.2 Future work

This thesis has attempted to extend our understanding of the amount of prosodic variation that can be accounted for under a single set of assumptions, bringing together wisdom from the literature on several disparate phenomena. It is, of course, not the last word on contrastive representations of word prosody, and there are several topics that need to be addressed before a more complete theory of word-level contrast can be reached.

The first of these concerns phonetic implementation. One of the strengths of a contrast-based approach is that a range of surface phonetic forms can be derived from identical phonological representations, requiring a more restricted (and non-universal) set of formal elements to be posited in the grammar. However, this alone does not account for the reality of the speech stream, which contains phonetic variation that crucially cannot be explained as functional consequences of the vocal tract. In the analyses presented here, I have explicitly discussed the distinction between formal representations and phonetic realizations, but I have done so only schematically, speaking of properties such as “{high pitch}” and “{increased duration}”. In order to reach a fuller understanding of language competence, an explicit (and, presumably, quantitative) account of the relationship between the formal categorical and the non-categorical is needed—a “phonetic

grammar” of sorts.

Another task is to further refine the formal typology sketched in the preceding section. The major consideration in doing this is the division of labour between restrictions built into the representational model and the processes that derive representations in particular languages. The relationship between rules and representations is by no means a new topic (see, for example, Anderson 1985), but it is certainly one that must be revisited each time new insight is offered into the nature of representations.

Finally, it will also be necessary to conduct more contrast-focused case studies of word prosodic systems in order to rigorously test the model and arrive at a more refined typology. This includes both more recently described systems like those studied in Chapter 4 and revisiting the prosodic systems of dialect continua with notable variation in their prosodic systems. Some possible candidates on which there already exists a substantial literature include varieties of Japanese, the mainland Scandinavian languages (Lahiri et al. 2005), and South Slavic languages such as Slovenian and Serbo-Croatian (Inkelas and Zec 1988; Zec 1999, 2009).

For languages with length and/or tone, the present model provides notational equivalents for standard analyses of these phenomena; X-slots serve as binary timing units and  $\pi$ -nodes serve as tonal root nodes. It is in languages with other combinations of prosodic features that this approach has theoretical implications and makes stronger predictions: in a language with both stress and autosegmental tone, stress must be represented as a quantity contrast (as in Papiamentu), leaving no room for an additional independent length contrast. Furthermore, though it is not the central point of this study, an important outcome is the understanding of prototypical contrastive lexical stress accent languages as operating on autosegmental principles, which has been proposed independently by Hagberg (2006) and Dubina (2012).

# References

- Anderson, Stephen R. 1985. *Phonology in the twentieth century: Theories of rules and theories of representations*. Chicago: University of Chicago Press.
- Archangeli, Diana, and Douglas Pulleyblank. 1986. The content and structure of phonological representations. Ms., University of Arizona and University of British Columbia.
- Bennett, Ryan. 2013. The uniqueness of metrical structure: Rhythmic phonotactics in Huariapano. *Phonology* 30:355–398.
- Blaho, Sylvia. 2008. The syntax of phonology: A radically substance-free approach. Doctoral Dissertation, University of Tromsø, Tromsø.
- Blevins, James P. 2008. Declension classes in Estonian. *Linguistica Uralica* 44:241–267.
- Broselow, Ellen. 1995. Skeletal positions and moras. In *The handbook of phonological theory*, ed. John A. Goldsmith, 175–205. Cambridge, Massachusetts: Blackwell.
- Broselow, Ellen, Sui-I Chen, and Marie Huffman. 1997. Syllable weight: Convergence of phonology and phonetics. *Phonology* 14:47–82.
- Bucci, Jonathan. 2013. Voyelles longues virtuelles et réduction vocalique en coratin. *Canadian Journal of Linguistics* 58:397–414.
- Bye, Patrik. 1997. A generative perspective on ‘overlength’ in Estonian and Saami. In *Estonian prosody: Papers from a symposium*, ed. Ilse Lehiste and Jaan Ross, 36–70. Tallinn: IEL.

- Bye, Patrik, and Peter Svenonius. 2012. Non-concatenative morphology as epiphenomenon. In *The morphology and phonology of exponence*, ed. Jochen Trommer, 426–495. Oxford: Oxford University Press.
- Byrnes, James. 2011. Towards a phonetically sound phonological framework of pitch accent languages: A case study of Tokyo Japanese. Master's thesis, University of Toronto.
- Caballero, Gabriela. 2008. Choguita Rarámuri (Tarahumara) phonology and morphology. Doctoral Dissertation, University of Southern California, Berkeley.
- Caballero, Gabriela. 2011. Morphologically conditioned stress assignment in Choguita Rarámuri. *Linguistics* 49:749–790.
- Caballero, Gabriela, and Lucien Carroll. to appear. Tone and stress in Choguita Rarámuri (Tarahumara) word prosody. *International Journal of American Linguistics*.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
- Clements, G. N. 1985. The geometry of phonological features. *Phonology Yearbook* 2:225–252.
- Clements, G. N., and Elizabeth V. Hume. 1995. The internal organization of speech sounds. In *The handbook of phonological theory*, ed. John A. Goldsmith, 245–306. Cambridge, Massachusetts: Blackwell.
- Devonish, Hubert. 1989. *Talking in tones: A study of tone in Afro-European Creole languages*. Christ Church, Barbados: Karia Press & Caribbean Academic Publications.
- Doner, Julianne. 2013. The unification of Spanish verbal and non-verbal stress. General Paper, University of Toronto.
- Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

- Dresher, B. Elan. 2014. The arch not the stones: Universal feature theory without universal features. *Nordlyd* 41:165–181.
- Dresher, B. Elan, and Aditi Lahiri. 1991. The Germanic Foot: Metrical coherence in Germanic. *Linguistic Inquiry* 22:251–286.
- Dresher, B. Elan, Glyne Piggott, and Keren Rice. 1994. Contrast in phonology: Overview. *Toronto Working Papers in Linguistics* 13:iii–xvii.
- Dubina, Andrei. 2012. Towards a tonal analysis of free stress. Doctoral Dissertation, Radboud University, Nijmegen, Netherlands.
- Ehala, Martin. 2003. Estonian quantity: Implications for Moraic Theory. In *Generative approaches to Finnic and Saami linguistics*, ed. Diane Nelson and Satu Manninen, 51–80. Stanford: CSLI.
- Fox, Anthony. 2000. *Prosodic features and prosodic structures*. Oxford: Oxford University Press.
- French, Koleen Matsuda. 1988. *Insights into Tagalog reduplication, infixation, and stress from nonlinear phonology*. Dallas, TX: Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics 84.
- Fry, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 35:765–769.
- Godfrey, Ross. 2012. How abstract are tonal features? Generals Paper, University of Toronto.
- Goedemans, Rob, and Harry van der Hulst. 2014. The separation of accent and rhythm: Evidence from StressTyp. In *Word stress: Theoretical and typological issues*, ed. Harry van der Hulst, 119–145. New York, NY: Cambridge University Press.
- Goldsmith, John A. 1976. Autosegmental phonology. Doctoral Dissertation, MIT, Cambridge, MA.

- Goldsmith, John A. 1990. *Autosegmental and metrical phonology*. Oxford: Basil Blackwell.
- Gordon, Matthew. 2002. A factorial typology of quantity-insensitive stress. *Natural Language and Linguistic Theory* 10:491–552.
- Gordon, Matthew. 2006. *Syllable weight: Phonetics, phonology, typology*. New York: Routledge.
- Gordon, Matthew. 2014. Disentangling stress and pitch accent. In *Word stress: Theoretical and typological issues*, ed. Harry van der Hulst, 83–118. New York, NY: Cambridge University Press.
- Green, M. M., and G. E. Igwe. 1963. *A descriptive grammar of Igbo*. London: Oxford University Press.
- Gussenhoven, Carlos. 2006. The prosody of Nubi between stress and tone. *Phonology* 23:193–223.
- Hagberg, Lawrence Raymond. 1993. An autosegmental theory of stress. Doctoral Dissertation, University of Arizona, Tucson.
- Hagberg, Lawrence Raymond. 2006. *An autosegmental theory of stress*. SIL e-Book.
- Hall, Daniel Currie. 2007. The role and representation of contrast in phonological theory. Doctoral Dissertation, University of Toronto, Toronto.
- Hall, Daniel Currie. 2011. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28:1–54.
- Halle, Morris, and William Idsardi. 1995. General properties of stress and metrical structure. In *The handbook of phonological theory*, ed. John A. Goldsmith, 403–443. Cambridge, MA: Oxford.
- Halle, Morris, and Jean-Roger Vergnaud. 1978. Metrical structures in phonology. Ms., MIT.
- Halle, Morris, and Jean-Roger Vergnaud. 1987. *An essay on stress*. Cambridge, MA: MIT Press.



- Haraguchi, Shosuke. 1977. *The tone pattern of Japanese: An autosegmental theory of tonology*. Tokyo: Kaitakusha.
- Harris, James W. 1995. Projection and edge marking in the computation of stress in Spanish. In *The handbook of phonological theory*, ed. John A. Goldsmith, 867–887. Cambridge, MA: Oxford.
- Hayes, Bruce. 1980. A metrical theory of stress rules. Doctoral Dissertation, MIT, Cambridge, MA.
- Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20:253–306.
- Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: Chicago University Press.
- Hint, Mati. 1973. *Eesti keele sõnafonoloogia I*. Tallinn: Eesti NSV Teaduste Akadeemia Keele ja Kirjanduse Instituut.
- Hint, Mati. 1997. The Estonian quantity degrees in prosody and morphophonology. In *Estonian prosody: Papers from a symposium*, ed. Ilse Lehist and Jaan Ross, 125–135. Tallinn: IEL.
- Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55:37–58.
- Hualde, José Ignacio. 2012. Two Basque accentual systems and the notion of pitch-accent language. *Lingua* 122:1335–1351.
- Hualde, José Ignacio, Oihana Lujanbio, and Francisco Torreira. 2008. Lexical tone and stress in Goizueta Basque. *Journal of the International Phonetic Association* 38:1–24.
- Hualde, José Ignacio, Oihana Lujanbio, and Juan Joxe Zubiri. 2010. Goizueta Basque. *Journal of the International Phonetic Association* 40:113–127.

- Hualde, José Ignacio, and Jon Ortiz de Urbina. 2003. *A grammar of Basque*. Berlin: Mouton de Gruyter.
- van der Hulst, Harry. 1984. *Syllable structure and stress in Dutch*. Dordrecht: Foris Publications.
- van der Hulst, Harry. 2012. Deconstructing stress. *Lingua* 122:1494–1521.
- van der Hulst, Harry. 2014a. Representing rhythm. In *Word stress: Theoretical and typological issues*, ed. Harry van der Hulst, 325–365. New York, NY: Cambridge University Press.
- van der Hulst, Harry. 2014b. The study of word accent: Past, present, and future. In *Word stress: Theoretical and typological issues*, ed. Harry van der Hulst, 3–55. New York, NY: Cambridge University Press.
- Hyman, Larry M. 1985. *A theory of phonological weight*. Dordrecht: Foris Publications.
- Hyman, Larry M. 1986. The representation of multiple tone heights. In *The phonological representation of suprasegmentals*, ed. Koen Bogers, Harry van der Hulst, and Maarten Mous, 109–152. Dordrecht: Foris Publications.
- Hyman, Larry M. 1993. Register tones and tonal geometry. In *The phonology of tone: The representation of tonal register*, ed. Harry van der Hulst and Keith L. Snider, 75–108. Berlin: Mouton de Gruyter.
- Hyman, Larry M. 2001. Privative tone in Bantu. In *Cross-linguistic studies of tonal phenomena: Tonogenesis, Japanese accentology, and other topics*, ed. Shigeki Kaji, 237–257. Tokyo: ILCAA.
- Hyman, Larry M. 2006. Word-prosodic typology. *Phonology* 23:225–257.
- Hyman, Larry M. 2009. How (not) to do phonological typology: The case of pitch-accent. *Language Sciences* 31:213–238.

- Hyman, Larry M. 2011. Do tones have features? In *Tones and features*, ed. John A. Goldsmith, Elizabeth Hume, and Leo Wetzels, 50–80. Berlin: Mouton de Gruyter.
- Idsardi, William. 1992. The computation of prosody. Doctoral Dissertation, MIT, Cambridge, MA.
- Inkelas, Sharon, and Draga Zec. 1988. Serbo-Croatian pitch accent: The interaction of tone, stress, and intonation. *Language* 64:227–248.
- Kager, René. 1995a. Consequences of catalexis. In *Leiden in last: HIL phonology papers I*, ed. Harry van der Hulst and Jeroen van de Weijer, 269–298. The Hague: Holland Academic Graphics.
- Kager, René. 1995b. The metrical theory of stress. In *The handbook of phonological theory*, ed. John A. Goldsmith, 367–402. Cambridge, Massachusetts: Blackwell.
- Karlin, Robin. 2014. Finnish dialectal epenthesis: Two distinct types of vowel insertion. Paper presented at the 22nd Manchester Phonology Meeting.
- Kaye, Jonathan, and Jean Lowenstamm. 1984. De la syllabicité. In *Forme sonore du langage: Structure des représentations en phonologie*, ed. F. Dell, D. Hirst, and J. R. Vergnaud, 123–159. Paris: Hermann.
- Kingston, John. 2011. Tonogenesis. In *The Blackwell companion to phonology*, ed. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 2304–2333. Cambridge: Blackwell.
- Kiparsky, Paul. 1991. Ms., Stanford University and Wissenschaftskolleg zu Berlin.
- Kouwenberg, Silvia, and Eric Murray. 1994. *Papiamentu*. Munich: Lincom Europa.
- Kraehenmann, Astrid. 2001. Swiss German stops: Geminate all over the world. *Phonology* 18:109–145.

- Krekoski, Ross. 2013. On tone and the nature of features. Generals Paper, University of Toronto.
- de Lacy, Paul. 2006. *Markedness: Reduction and preservation in phonology*. Cambridge: Cambridge University Press.
- de Lacy, Paul. 2014. Evaluating evidence for stress systems. In *Word stress: Theoretical and typological issues*, ed. Harry van der Hulst, 149–193. New York, NY: Cambridge University Press.
- Lahiri, Aditi, Allison Wetterlin, and Elisabet Jönsson-Steiner. 2005. Lexical specification of tone in North Germanic. *Nordic Journal of Linguistics* 28:61–96.
- Leben, William R. 1973. *Suprasegmental phonology*. Doctoral Dissertation, MIT, Cambridge, MA.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
- Lehiste, Ilse. 1977. Variability in the production of suprasegmental patterns. In *Studies in Finno-Ugric Linguistics in honor of Alo Raun*, ed. Denis Sinor, 131–139. Bloomington: Indiana University.
- Lehiste, Ilse. 1985. An Estonian word game and the phonematic status of long vowels. *Linguistic Inquiry* 16:490–492.
- Lehiste, Ilse. 2003. Prosodic changes in progress: From quantity language to accent language. In *Development in prosodic systems*, ed. Paula Fikkert and Haike Jacobs, 47–65. Berlin: Mouton de Gruyter.
- Lehtonen, Jaakko. 1970. *Aspects of quantity in Standard Finnish*. Jyväskylä: Jyväskylä University.
- Levin, Juliette. 1985. *A metrical theory of syllabicity*. Doctoral Dissertation, MIT, Cambridge, MA.

- Lieberman, Mark. 1975. The intonational system of English. Doctoral Dissertation, MIT, Cambridge, MA.
- Lieberman, Mark, and Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8:249–336.
- Lieber, Rochelle. 1987. *An integrated theory of autosegmental processes*. Albany, NY: State University of New York Press.
- Lippus, Pärtel, Karl Pajusalu, and Jüri Allik. 2011. The role of pitch cue in the perception of the Estonian long quantity. In *Prosodic categories: Production, perception and comprehension*, ed. Sónia Frota, Gorka Elordieta, and Pilar Prieto, 231–242. Dordrecht: Springer Netherlands.
- Lippus, Pärtel, and Jaan Ross. 2011. Has Estonian quantity system changed in a century? Comparison of historical and contemporary data. *ICPhS* 17:1262–1265.
- Lockwood, David. 1983. Tone in a non-substantive theory of language. *Lacus Forum* 10:131–140.
- Lombardi, Linda. 1990. The nonlinear organization of the affricate. *Natural Language and Linguistic Theory* 8:375–425.
- McCarthy, John J. 1979. Formal problems in semitic phonology and morphology. Doctoral Dissertation, MIT, Cambridge, MA.
- McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12:373–418.
- McCarthy, John J., and Alan Prince. 1986. Prosodic morphology. Ms., MIT and Brandeis University.
- McCawley, James D. 1978. What is a tone language? In *Tone: A linguistic survey*, ed. Victoria Fromkin, 113–131. New York: Academic Press.

- Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.
- Morén, Bruce. 2003. The Parallel Structures model of feature geometry. *Working Papers of the Cornell Phonetics Laboratory* 15:194–270.
- Mürk, Harry. 1981. The grade alternation system in Estonian morphophonology. Master's thesis, University of Toronto.
- Mürk, Harry. 1982. Paradigmatic length alternations in Estonian: An auto-segmental/metrical analysis. *Toronto Working Papers in Linguistics* 3:43–60.
- Myers, Scott. 1997. OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15:847–892.
- Nahkola, Kari. 1987. *Yleisgminaatio: Äänne­muutoksen synty ja vaihteet kielisysteemissä erityisesti Tampereen seudun hämäläismurteiden kannalta*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Odden, David. 1997. Some theoretical issues in Estonian prosody. In *Estonian prosody: Papers from a symposium*, ed. Ilse Lehiste and Jaan Ross, 165–195. Tallinn: IEL.
- Odden, David. 2011. Features impinging on tone. In *Tones and features*, ed. John A. Goldsmith, Elizabeth Hume, and Leo Wetzels, 81–107. Berlin: Mouton de Gruyter.
- Pater, Joe. 1994. Against the specification of an 'exceptional' English stress pattern. *Toronto Working Papers in Linguistics* 13:95–121.
- Pierrehumbert, Janet B., and Mary E. Beckman. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.
- Poser, William J. 1984. The phonetics and phonology of tone and intonation in Japanese. Doctoral Dissertation, MIT, Cambridge, MA.

- Prillop, Külli. 2013. Feet, syllables, moras and the Estonian quantity system. *Linguistica Uralica* 49:1–29.
- Prince, Alan. 1980. A metrical theory of Estonian quantity. *Linguistic Inquiry* 11:511–562.
- Prince, Alan. 1983. Relating to the grid. *Linguistic Inquiry* 14:19–100.
- Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell Publishing.
- Raadik, Maire. 2013. *Eesti õigekeelsussõnaraamat*. Tallinn: Eesti Keele Sihtasutus. Available online at <http://www.eki.ee/dict/qs/>.
- Remijsen, Bert, and Vincent J. van Heuven. 2005. Stress, tone and discourse prominence in the Curaçao dialect of papiamentu. *Phonology* 22:205–235.
- Revithiadou, Anthi. 1999. Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems. Doctoral Dissertation, Leiden University, Leiden.
- Rice, Keren. 2000. Tonal markedness in three Athapaskan languages. Colloquium handout, University of California Berkeley.
- Rivera-Castillo, Yolanda. 1998. Tone and stress in Papiamentu: The contribution of a constraint-based analysis to the problem of creole genesis. *Journal of Pidgin and Creole Languages* 13:297–334.
- Rivera-Castillo, Yolanda. 2009. Subsystem interface and tone typology in Papiamentu. *Studies in Linguistics* 33:437–458.
- Rivera-Castillo, Yolanda, and Lucy Pickering. 2004. Phonetic correlates of stress and tone in a mixed system. *Journal of Pidgin and Creole Languages* 19:261–284.
- Roca, Iggy. 2005. Saturation of parameter settings in Spanish stress. *Phonology* 22:345–394.

- Römer, Raúl. 1983. Papiamentu tones. In *Papiamentu: Problems and possibilities*, ed. Enrique Muller, 85–96. Zutphen: De Walburg Pers.
- Römer, Raúl. 1991. Studies in Papiamentu tonology. Amsterdam & Kingston: Amsterdam Centre for Caribbean Studies.
- Sagey, Elizabeth Caroline. 1986. The representation of features and relations in non-linear phonology. Doctoral Dissertation, MIT, Cambridge, MA.
- Selkirk, Elisabeth. 1984. *Phonology and syntax*. Cambridge, MA: MIT Press.
- Selkirk, Elisabeth. 1990. A two-root theory of length. *University of Massachusetts Occasional Papers in Linguistics* 14:123–171.
- Sloan, Kelly Dawn. 1991. Syllables and templates: Evidence from Southern Sierra Miwok. Doctoral Dissertation, MIT, Cambridge, MA.
- Snider, Keith L. 1999. *The geometry and features of tone*. Publications in Linguistics. Dallas and Arlington: Summer Institute of Linguistics and University of Texas at Arlington.
- Spahr, Christopher. 2011. Dialectal gemination in Finnish: Phonetics/phonology interplay in moraic theory. Master's thesis, University of Toronto.
- Spahr, Christopher. 2013. Rethinking the morphophonology of Estonian quantity. In *Nordic Prosody: Proceedings of the XIth Nordic Prosody Conference, Tartu 2012*, ed. Eva-Liina Asu and Pärtel Lippus, 362–371. Frankfurt: Peter Lang Edition.
- Spahr, Christopher. 2014a. A contrastive hierarchical account of positional neutralization. *The Linguistic Review* 31:551–585.
- Spahr, Christopher. 2014b. Floating vowels and prosodic structure in Estonian quantity. Ms., University of Toronto.



- Spahr, Christopher. 2014c. Prosodic epenthesis and floating vowels in Estonian quantity. Paper presented at the Eleventh Old World Conference in Phonology, LUCL and Meertens Instituut Amsterdam.
- Suomi, Kari, Juhani Toivonen, and Riikka Ylitalo. 2008. *Finnish sound structure: Phonetics, phonology, phonotactics, and prosody*. Studia Humoniora Ouluensia 9. Oulu: University of Oulu.
- Suomi, Kari, and Riikka Ylitalo. 2004. Durational and tonal correlates of accent in Finnish. *Journal of Phonetics* 32:35–63.
- Szpyra, Jolanta. 1992. Ghost segments in nonlinear phonology: Polish yers. *Language* 68:277–312.
- Topintzi, Nina. 2010. *Onsets: Suprasegmental and prosodic behaviour*. Cambridge: Cambridge University Press.
- Tranel, Bernard. 1991. CVC light syllables, geminates and Moraic Theory. *Phonology* 8:291–302.
- Tranel, Bernard. 1995. French final consonants and nonlinear phonology. *Lingua* 95:131–167.
- Viitso, Tiit-Rein. 2003. Structure of the Estonian language. In *Estonian language*, ed. Mati Ereht, 9–129. Linguistica Uralica Supplementary Series Volume 1.
- Walker, Rachel. 2011. *Vowel patterns in language*. Cambridge: Cambridge University Press.
- Wiik, Kalevi, and Ilse Lehiste. 1968. Vowel quantity in Finnish disyllabic words. In *Congressus Secundus Internationalis Fenno-Ugristarum, Pars I*, ed. M. Kahla and A. Räisänen, 569–574. Helsinki: Societas Fenno-Ugrica.
- Yip, Moira. 1989. Contour tones. *Phonology* 6:149–174.
- Ylitalo, Riikka. 2009. The realization of prominence in three varieties of Standard Spoken Finnish. Doctoral Dissertation, University of Oulu, Oulu.

- Zec, Draga. 1999. Footed tones and tonal feet: rhythmic constituency in a pitch-accent language. *Phonology* 16:225–264.
- Zec, Draga. 2009. Tone and stress in prominence based prosodic systems. In *Annual workshop on formal approaches to Slavic linguistics: The Yale meeting 2008*, ed. Jodi Reich, Maria Babyonyshev, and Daria Kavitskaya, 269–288. Ann Arbor: Michigan Slavic Publications.
- Zhang, Jie. 2007. Contour tone distribution is not an artifact of tonal melody mapping. *Studies in the Linguistic Sciences* 33:1–61.
- Zimmermann, Eva. 2013. Vowel deletion as mora usurpation: The case of Yine. *Phonology* 30:125–163.