CONSTRAINTS ON PHONOLOGICAL INTERACTIONS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF LINGUISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Lev A. Blumenfeld

September 2006

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Paul Kiparsky) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Arto Anttila)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(William Leben)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Armin Mester)

Approved for the University Committee on Graduate Studies.

Abstract

Optimality Theory (OT) is committed to a view of phonology where significant generalizations are placed in the character of output structures. Markedness constraints state these surface preferences, and grammars are free to choose from a variety of paths (repair strategies) that enforce the output structures. OT faces systematic difficulties with cases where a given markedness constraint is observed to cause fewer repairs than predicted by the theory. This dissertation examines several cases of this type, termed 'too-many-solutions' problems. I argue that the difficulty faced by OT is due to the significant phonological generalizations being most insightfully stated not as output preferences, but preferences for input-output mappings. I argue for a new type of OT markedness constraint to handle such 'procedural' generalizations. Unlike traditional markedness constraints, these new constraints penalize dispreferred processes rather than output forms. The typology of interactions between prosodic and segmental properties provides the empirical evidence for the proposals. The asymmetry between, on the one hand, those phonological categories to which stress can be sensitive, and on the other, those properties which it can condition has posed an intractable too-many-solutions problem for standard OT. At the root of the difficulties, I argue, is the fact that the important generalization in the domain of prosody-segmental interactions is in the processes, not the outputs. The formal proposals are first developed with reference to this empirical domain. I propose a theory of procedural markedness constraints which refer to the direction of interaction between the relevant categories. These constraints penalize the candidates that involve typologically unobserved repair strategies in such a way that those candidates cannot be optimal. In the final part of the dissertation I apply the new theory of constraints to another problematic too-many-solutions case, the typology of vowel syncope and epenthesis. Here, too, the procedural markedness constraints become necessary to account for systematic gaps in the conditioning environments of the two processes.

Acknowledgments

I wish to express my gratitude, first, to the members of my committee. Paul Kiparsky has contributed to every aspect of my work at Stanford, from the first days to the dissertation. His inspired approach to linguistics, appreciation of complex arguments and artful analyses, and personal generosity have left a deep impression on many generations of students, and I am not an exception.

Arto Anttila has supplied a healthy dose of discipline, skepticism, and rigor, when any of the three were lacking. His penetrating questions have greatly improved the quality of both the ideas and the writing of this thesis.

I am grateful to Will Leben's careful and detailed criticism of many aspects of my work, as well as warm encouragement in the matters of research and pedagogy.

Armin Mester, the external member of my committee, brought a fresh perspective, new questions, and new ideas to this project. I thank him for his meticulous and immensely helpful comments on the drafts of my thesis, and for the fruitful discussions on many topics in phonology.

The importance of the issues addressed in this thesis first became clear to me in a 'directed reading' with Edward Flemming, for whose intelligent criticism and guidance at the earliest stages of my graduate work I am immensely grateful.

Many people outside of my thesis committee have helped me in various ways. I am grateful to John Beavers, Paul de Lacy, Nilson Gabas, Maria Gouskova, Bruce Hayes, Philip Hofmeister, Larry Hyman, Junko Itô, for her encouragement and insightful criticism of this project, Florian Jaeger, Andrew Koontz-Garboden, Jaye Padgett, A.K.Polivanova, for her recent and not so recent wise advice and generosity, Donca Steriade, Judith Tonhauser, Nina Topintzi, Paul Smolensky, and audiences at CLS, Berkeley, Stanford, and Santa Cruz.

I also wish to thank my department for its support over the past five years. Stanford has provided an optimal setting, with the right amount of the right kind of supervision at the right time, an absolute minimum of bureaucratic hassles, and an atmosphere that

mixed independence with cooperation in the right proportion. This means that I, and not my teachers, must take full responsibility for my not knowing many things about linguistics, and for the errors and shortcomings of this thesis.

Finally, I owe a private debt of gratitude to my family and friends in the U.S., Russia, and Latvia. As Wisława Szymborska once said, the simplest words are the most serious. *Thank you. Спасибо. Paldies.*

Table of Contents

INTRODUCTION

The central argument for Optimality Theory (OT) was an argument of explanatory adequacy. Kisseberth 1970 was the first to observe that rule-based phonology had no mechanism for explaning relationships between outputs of phonological rules. SPE-style bracketing conventions, he argued, did not always pick out the sets of rules that had a natural relationship with each other, a relationship that was based on the similarity of outputs these rules produced. Kisseberth showed that languages tend to contain clusters of apparently different rules, all of which conspire to create a particular pattern of syllable structure, seemingly catering to a preference to produce a certain kind of output. Over the years, Kisseberth's original insight was developed, culminating in the realization that the weight and importance of output-based conspiracies was too great for standard rule-based theory to handle.

These arguments became the cornerstone of OT, the most radical version of the set of theories that recognized the importance of output generalizations.

It has become apparent in recent years, however, that OT faces a number of systematic challenges that appear to be related to its radical commitment to locating ALL phonological generalizations in the output. These challenges fall into two broad categories. The first set, which I briefly mention here but will not deal with in the remainder of the thesis, have to do with SERIALITY and OPACITY. Cases where the conditioning environment of a phonological process is not met in the surface representation, or where a conditioning environment is present but the process has not applied, have required extrinsic rule ordering in derivational theory. Because this mechanism is not available to OT, the theory faces a systematic challenge in handling opaque interactions. Much work has been done in this area, using two general strategies: first, introducing a serial or stratal architecture into OT, and second, modifying the set of constraints and their interpretation to allow opacity to be handled solely by output-oriented constraints.

The second systematic challenge faced by OT, which will be the subject of this thesis, also has to do with generalizations not being surface-based. It has become clear in recent years that Kisseberth's original argument in favor of surface conspiracies can be turned around as an argument against the radical surface orientation of OT. Perhaps most clearly this can be seen in Steriade's (2001) work on the typology of repairs for the constraint against final voiced stops. Given a markedness constraint penalizing final obstruents from being voiced, *[+voi, –son]#, the only attested repair for violations of such a constraint is final voicing. Other potential changes that also could obviate violations of this constraint are not attested as responses to final voicing. Among them are nasalization, deletion, metathesis, and epenthesis, as illustrated below.

(1)	/tab/	→	[tab]	(faithful output)	
		→	[tap]	(devoicing)	
		↛	[tam]	(nasalization)	*unattested*
		↛	[ta]	(deletion)	
		↛	[tba]	(metathesis)	
		↛	[tabə]	(epenthesis)	

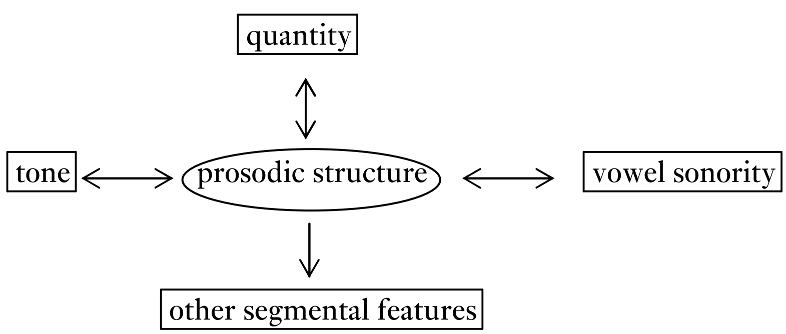
Steriade stresses that while final epenthesis, deletion, etc. are common across languages, these processes are unattested AS REPAIRS for final obstruent voicing, i.e. they never selectively target voiced stops. However, standard OT has no direct way of ruling out such repairs, as long as the surface condition expressed by the markedness constraint *[+voi, –son]# is satisfied in their outputs. It appears that surface-oriented constraints conspire to produce some input-output mappings but not others. This general challenge for OT has come to be known as the TOO-MANY-SOLUTIONS problem. Steriade compares this difficulty to Kisseberth's classic conspiracy argument.

> "Kisseberth's (1970) insight that conspiracies arise when the sound system aims at a specific target structure via multiple means can lead one to ask the same question, in the context of rule-based phonology: if the rule of final devoicing aims to eliminate final voiced obstruents, why aren't there rules of final obstruent nasalization, deletion, metathesis or post-voiced obstruent epenthesis?" (Steriade 2001: 6).

Such constraint conspiracies will be the focus of my dissertation. I will argue that the origin of the difficulty faced by OT in this and other similar cases has to do with the locus of the phonologically significant generalizations. At least some generalizations in phonology are most insightfully formulated not as output statements, but as statements about input-output mappings, and about environments of processes. The thrust of the argument is thus analogous to the original conspiracy argument in favor of OT, albeit in the opposite direction.

The thesis is organized as follows. In Chapter 1 I take a look at a too-many-solutions problem that has not received systematic attention in the literature, the interaction of prosodic structure with segmental features. The typology of the interactions is asymmetrical. Broadly speaking, metrical structure can condition the distribution of a greater range of phonological properties than it can be sensitive to. Stress can interact bidirectionally with only three properties: quantity, tone, and vowel sonority, while the set of segmental features that can be sensitive to prosodic structure is broader and includes such consonantal properties as aspiration, continuancy, and voicing. While the distribution of these features is commonly sensitive to stress, they are not observed to condition the placement of stress in any language. This asymmetry in interaction between prosody and segment is schematically illustrated below.

(2)



I will argue that such a pattern of interaction between two phonological properties presents a challenge for standard OT, because the generalization is best stated not in terms of output structures, but as a constraint on input-output mappings. OT markedness constraints are output conditions. However, a constraint that calls for two properties such as stress and aspiration to cooccur on the same syllable in the output cannot account for aspiration attraction to stressed syllables (stress-driven aspiration) without also predicting interaction in the reverse direction, i.e. aspiration-driven stress. This too-many-solutions problems turns out to be a general one in the case of stress-segmental interactions. As I will argue in Chapter 2, the problem cannot be handled by three existing general proposals on too-many-solutions problems: Targeted Constraint theory, the P-map, and fixed rankings between classes of constraints.

In Chapter 3 I move on to my proposal. The diagnosis of the difficulty faced by standard OT in too-many-solutions problems has its origin in the locus of the phonologically significant generalization. While standard OT has claimed that all generalizations lie in the output structures, and while OT markedness constraints are statements about outputs, I argue that in the case of prosody-segmental interactions it is in terms of the input-output mapping that the generalization is most insightfully stated.

This diagnosis leads to the solution. I introduce a new class of markedness constraints that directly penalize the unwanted input-output mappings by stating the asymmetrical direction of interaction between two phonological properties. These PROCEDURAL constraints are given in the form of implicational 'If-then' statements, such that the asymmetry of interaction between the two properties mentioned in the constraint corresponds to its asymmetrical statement. For example, the procedural stress-aspiration constraint will be stated in the form 'If a syllable is stressed, then its onset is aspirated. The formal machinery introduced in Chapter 3 ensures that the property mentioned in the antecedent part of the constraint can force the unfaithful mapping of the property mentioned in the consequent, but not vice versa. Thus, the procedural stress-aspiration constraint will be able to force stress-driven aspiration, but not aspiration-driven stress.

This proposal provides a general way of handling asymmetrical interactions in phonology, and a general way of dealing with too-many-solutions problems. In Chapter 4 I will apply the proposals from Chapter 3 to a new set of cases, vowel epenthesis and deletion. I will argue that the typology of both of these processes is not as rich as

Standard OT would predict, and show how my proposal can be used to constrain the interaction of the phonological properties. I will argue that vowel epenthesis is used exclusively as a response to pressures of syllable structure, sonority sequencing, syllable contact, and word minimality, but cannot be used to avoid violations of other metrical constraints. The typology of vowel syncope is constrained in terms of its environment: syncope targets only weak vowels, i.e. those vowels that are unstressed, posttonic, in the weak branch of feet, and so forth. Crucially, I will argue that metrical constraints cannot force the deletion of STRESSED vowels, a claim that goes against current thinking in OT.

In general, both syncope and epenthesis show generalizations that are best stated not in terms of output preferences, but in terms of the environments of the processes. This, once again, diagnoses the difficulty that standard OT has with the typology of the processes, and makes them amenable to the procedural constraints proposed in this dissertation.

CHAPTER 1

PROSODY-SEGMENTAL INTERACTIONS: THE TYPOLOGY

1.0 Introduction

In the previous introductory chapter I discussed some of the problems facing OT as a radically output-oriented theory of phonology. This chapter turns to a typological survey of a phenomenon that presents a too-many-solutions problem for OT: the interaction of the prosodic and segmental components of phonology. This set of issues, not previously given a systematic treatment in OT theorizing, will prove especially insidious for the current setup of the theory, because it defies any attempts to attack it using the techniques developed for other too-many-solutions problems. In this chapter I will argue that the problem of prosody-segmental interactions diagnoses a general flaw of standard OT, because the generalization behind the typological facts is best formulated not in terms of output structures, but in terms of input-output mappings. In Chapter 2 I will show that none of the available proposals intended to deal with too-many-solutions problems is able to handle prosody-segmental interactions in a satisfactory way. I will turn to my proposal on how such generalizations can be allowed their place in the theory of phonology in Chapter 3.

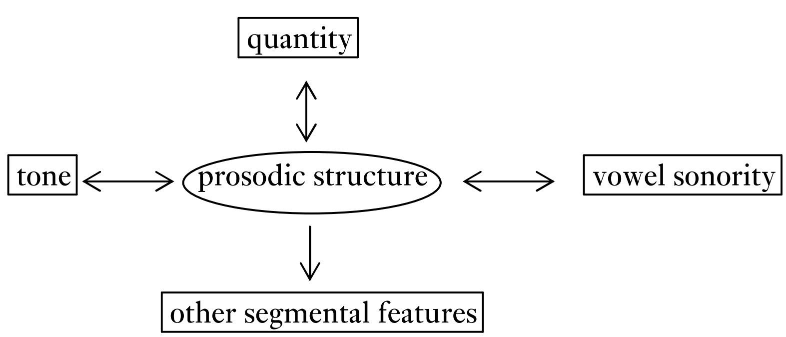
1.1 The facts

In this section I will survey the typology of the interaction between prosodic structure and segments. The crosslinguistic data will show that stress assignment rules can refer to only a very restricted set of properties in a representation. Phonological factors that may figure in stress assignment rules are limited to three general categories: quantity (vowel length and syllable weight), tone, and vowel sonority.[1] The claim I will defend in this dissertation is that all phonological stress systems can be described with reference solely to these three categories.

The reverse direction of interaction between prosody and segments is much less constrained. Prosodic structure can influence the realization of a host of segmental properties. Quantity, tone, and vowel sonority are among them, but stress can also condition the distribution of many other segmental features. In many languages, fortition and lenition of consonants is sensitive to their prosodic context. Metrical structure frequently affects the distribution of laryngeal features. Prosodic constituents can serve as domains of segmental processes such as harmony.

Apart from quantity, tone, and vowel sonority, none of the varied effects of stress on segments are observed as conditioning factors in stress assignment. Thus, the key typological claim in this chapter is that stress can influence more properties than it can be sensitive to: prosody-segmental interactions are asymmetrical. This asymmetry is given schematically below.

(1)



After surveying the typology, I will show that this asymmetry in interaction poses a general problem for OT as an output-oriented theory. I will conclude this chapter by discussing potential counterexamples to the typological generalization – cases where stress assignment appears to be sensitive to properties other that quantity, tone, and

[1] Of course, morphology can also play a role in stress assignment. I am concerned here with those factors of stress assignment which are independent of the morphological structure of words.

vowel sonority – arguing that alternative analyses that preserve the typological claim are available.

1.1.1 Bidirectional interactions

In this section I survey the three factors with which stress interacts bidirectionally: quantity, tone, and vowel sonority. In each of these cases, there are languages where stress assignment requires access to information about these categories, and there are other languages where these categories are distributed in a way that is sensitive to the location of stress (which, in turn, is assigned according to other principles).

1.1.1.1 Quantity

Metrical structure has an intimate relationship with syllable QUANTITY. The prominence of heavier syllables plays a role in stress systems in two complementary ways. On the one hand, there are languages where stress gravitates toward heavier syllables. On the other hand, there are languages where stress is assigned based on criteria other than weight, but the location of stress influences the weight distribution in the resulting form. Many languages combine features of both weight-driven stress and stress-driven weight.

Interactions of stress and weight are extremely common; perhaps a majority of the world's languages have some relationship between the two domains. Such interactions have been extensively documented and analyzed in the major theoretical works on metrical stress (Liberman 1975, Liberman and Prince 1977, Prince 1983, Halle and Vergnaud 1987, Halle and Idsardi 1995, Hayes 1981, Hayes 1995, Gordon 1999, among many others).

Let me give a few examples of the familiar pattern. The Classical Latin stress rule is the prototypical case of weight-driven stress. The main stress falls on the penultimate syllable if it is heavy, and otherwise on the antepenultimate syllable. The basic Latin pattern can be seen, with variations, in many languages. For example, pre-classical (Plautine) Latin had a stress rule identical to Latin except that words ending in LLLH or LLLL sequences were stressed on the preantepenult. Fijian (Dixon 1988) stress rule is identical to the Classical Latin except shifted one syllable rightward: stress falls on the final syllable if it is heavy, and on the penultimate syllable otherwise. Lebanese Arabic (section 4.2.3.3) has a stress rule that is identical to Latin, with the added requirement that final superheavy (trimoraic) syllables are stressed. Languages with stress that gravitates to the beginning of the form also often have weight sensitivity: Tümpisa Shosone, for example, stresses the second syllable if it is heavy and the initial syllable otherwise (Dayley 1989). Hopi has the converse system: it stresses the initial syllable if it is heavy and the second syllable otherwise (Jeanne 1982). The list of languages with weight-sensitive stress can proceed almost indefinitely; the reader is referred to Hayes 1995 and Gordon 1999 for extensive surveys.

Stress-driven weight is also a rather common phenomenon. In many languages, stressed vowels lengthen, e.g. in Hixkaryana, Carib, and various dialects of Yupik (some of these systems are discussed in Chapter 4). Gemination is also frequently used to make stressed syllables heavy. In Latvian, which has regular initial stress, voiceless obstruents following initial stressed light syllables surface as geminates, making that syllable heavy (Kariņš 1996).

1.1.1.2 Tone

The first systematic theoretical discussion of tone-stress interactions can be found in de Lacy (2002). De Lacy's analysis, framed in Optimality Theory, posits the tonal markedness scale H ⟩⟩ M ⟩⟩ L, and a set of constraints on the relationship between stress and tone. The constraints force tones higher on the markedness scale to be preferentially aligned with metrically prominent syllables and tones on the lower end of the scale to align with non-prominent syllables. These output conditions can be satisfied in two ways: by attracting stress to tone, or by attracting tone to stress. These two patterns represent the two types of languages, those with tone-driven stress and those with stress-driven tone. De Lacy provides examples of both.

Ayutla Mixtec has a tone-driven stress generalization: stress falls on the leftmost H-toned syllable immediately followed by a L-toned syllable (de Lacy 2002: 5). Failing this condition, the leftmost H toned syllable is stressed, else the leftmost M-toned syllable followed by a syllable with a L tone, and else, by default, the leftmost syllable. De Lacy analyzes this system as a case of stress attraction to underlying tones. The alignment of the tone markedness scale with the metrical prominence of syllables results in a preference to build binary feet in such a way that the stressed syllable has a H or a M tone, and the unstressed syllable a L tone. De Lacy proposes a constraint system that generates the following harmony scale of metrical feet-to-tone alignment (2). The stress generalization can then be stated with reference to this scale: for each given form, the most harmonic possible binary foot is formed.

(2) (H́L) 〉〉 (H́H) 〉〉 (ḾL) 〉〉 (ḾM), (ĹL)

De Lacy's example of a stress-driven tone language is Lamba. In this language, H tones do not surface on the morpheme with which they are underlyingly associated, but gravitate toward metrically prominent positions, as determined by a trochaic stress system.

I have argued elsewhere that both tone-to-stress and stress-to-tone systems can coexist in the same language (Blumenfeld 2004). In the lexical phonology, Ancient Greek had a tone-to-stress system. Metrical structure was assigned by a generalization similar to the Latin stress rule, and the tonal melody HL* was associated to the metrical head of the word. Conversely, the metrical constituents built in the phrasal phonology gravitated toward the tones inherited from the lexical phonology in a stress-to-tone fashion.

1.1.1.3 Vowel sonority

Vowel quality can also interact with stress in a bidirectional fashion, an interaction that is limited to a highly constrained set of cases. Broadly speaking, the higher a vowels stands on the sonority scale (3), the more likely it is to attract stress. Likewise, the interaction can proceed in the opposite direction: unstressed vowels are shifted lower on the sonority scale, via raising or vowel reduction.

(3) ə, ɨ 〈〈 i, u 〈〈 e, o 〈〈 a

I postpone further discussion of sonority-driven stress until Section 3.6, which will be devoted to examining the typology in detail.

1.1.2 Unidirectional interactions

The three types of cases outlined above comprise only a small fraction of the typology of prosody-segmental interactions. Apart from quantity, tone, and vowel sonority, the distribution of a large number of other properties can be sensitive to metrical structure. In all such cases, the interaction takes place only in one direction: from prosody to segments.

In a recent survey, González (2003) lists more than seventy languages that have segmental processes conditioned by stress or metrical structure. A summary of this survey is given below in (4), with a non-exhaustive list of languages under each type of interaction. The interactions fall into two main categories: strengthening in prominent syllables and weakening in non-prominent syllables. The former type includes increased duration of consonants and VOT-related differences, such as onset gemination in stressed syllables (Senoufo) or fortition in foot-initial syllables (Yupik). Onset epenthesis (usually [ʔ]) also falls into this category. Weakening processes in non-prominent positions include the familiar English flapping case, as well as other types of lenition such as Spanish fricativization. Some languages have consonant deletion (usually ʔ or *h*) in weak positions. The most common type of strengthening/weakening alternations conditioned by prosody involves laryngeal features (voicing, aspiration, glottalization) (4)c.

(4) a. **Strengthening in prominent syllables**

Increased duration of consonants in stressed syllables
Spanish, Senoufo, Urubu-Kaapor
Fortition
Yupik, Maori, Squamish
Onset epenthesis in stressed syllables
English, Dutch, Paipai, Huariapano

b. **Weakening in non-prominent syllables**

Flapping
English, Djabugay, Senoufo
Other lenition
Spanish, Nganasan, Paamese, Guyabero
Consonant deletion
English, Squamish, Popoloca, Capanahua

c. **Laryngeal feature alternations**

Voicing
Senoufo, Wembawemba, Wergaia, Trique
Post-aspiration
English, Maori, Farsi
Pre-aspiration
Icelandic, Faroese, Ingush, Toreva Hopi
Glottalization
Saanich, Gitksan

In the table below I give a representative list of languages that have stress-sensitive strengthening processes. Unless otherwise noted, the source is González 2003.

(5)

English	voiceless stop aspiration	onset of σ́	
Silacayopan Mixtec	t → t^h	onset of σ́	
Toreva Hopi	C preaspiration	after σ́	
Icelandic	C preaspiration	after σ́ with a short V	
Senoufo	C→C:	onset of σ́	Mills 1984
Popoloca	C→C:	after σ́	
Latvian	voiceless stop gemination	after σ́ with short V	Kariņš 1996
Maori	voiceless stop affrication	onset of σ́	
Norton Sound Yupik	{w, j, l} → {v, z, ɮ}	foot-initial	
Guyabero	d → θ	after σ́	
Nganasan	C voicing	foot-initial	
Twana	ʔ attraction	to σ́	
Bagnere-de-Luchon French	liquid attraction	to σ́	Blevins & Garrett 1998
Dutch	ʔ epenthesis	σ́	
Paipai	ʔ epenthesis	word-initial σ́	

The table below lists a representative sample of stress-sensitive weakening processes.

(6)

West Tarangan	{g,dʒ} → {w,j}	medial unstressed σ	Nivens 1992
Djabugay	{r,d} → ɾ	V'__V	Patz 1991
Copala Trique	{d, g} → {ð, ɣ}	V'__V	
Kupia	spirantiz., flapping	V__unstrV	Christmas & id. 1975
Pattani	deaspiration	unstressed syl	Sarma 1982
Senoufo	Stop voicing	V__V	
Wergaia	Stop devoicing	unstressed σ	
Oneida	h, ʔ deletion	posttonic σ	
Southern Tati	h→∅	except onset of σ́	Yar-Shater 1969
Capanahua	ʔ deletion	weak-footed σ	
Chilean Spanish	s-deletion	unstressed σ	

As a matter of illustration, one language, Senoufo, has an unusually rich array of segmental effects of stress (Niger Congo, Gur, Ivory Coast; Mills 1984, González 2003). Stress is usually initial in this language; no secondary stress is reported. Senoufo has at least five separate processes sensitive to stress. Onset consonants are lenited in unstressed syllables, with voiceless consonants becoming voiced, and voiced consonants becoming spirantized. Conversely, onset consonants in stressed syllables are lengthened. Senoufo also has flapping of /d/ in onsets of unstressed syllables. Finally, both secondary articulation and glottalization are contrastive only in stressed syllables. These processes are summarized below.

- Onset consonants lenited in unstressed syllables
  Voiceless → voiced, voiced → spirantized (Mills 1984: 131)
- Onset consonants are longer in stressed syllables (p. 119)
- /d/ is flapped in onset of unstressed syllables (p. 96)
- Secondary articulation contrastive only in stressed syllable (p.143)
- Glottalization contrastive only in stressed syllable (p.148)

The second major type of interaction between metrical structure and segmentism is less direct. Segmental processes may require reference to metrical constituents as domains of application (Flemming 1994). Nasal harmony, laxness harmony, and height harmony have all been attested with metrical constituents serving as domains of application. For example, in Guaraní, (Gregores and Suárez 1967, Flemming 1994), [nas] spreads from a stressed vowel leftward up to the next stressed vowel. In Tudanca Spanish, final high vowels are lax, and laxness spreads leftward until it reaches the stressed syllable (Flemming 1994, Walker 2004). Examples, taken from Flemming 1994, are shown below, with capitalization indicating laxness.

(7)	a.	(pÍntU)	'male calf'	b.	(pínta)	'female calf'
		(čÍkU)	'boy'		(číka)	'girl'
		se(kÁlU)	'to dry him'		se(kálo)	'to dry it'
	c.	o(rÉgAnU)	'oregano'			
		(pÓrtIkU)	'portico'			
		ra(kÍtIkU)	'rachitic'			

Flemming's analysis of this case posits left-headed stress feet serving as the domain of laxness harmony. The rationale for laxness spreading can be understood to result from a prohibition on disagreement in laxness within the stress foot.

I will discuss the Tudanca Spanish case in more detail in Chapter 3.

1.2 Too many solutions

The previous two sections established the asymmetrical nature of prosody-segmental interactions. Metrical structure interacts bidirectionally only with three properties: quantity, tone, and vowel sonority. All the other effects of prosodic structure – the various consonantal fortitions and lenitions, and the use of prosodic constituents as domains of processes – show a unidirectional influence of prosody on segments, but not the reverse.

Interaction between two phonological categories is modeled in OT with constraints that state the relationship between those two categories in the output structure. OT constraints employed to account for prosody-segmental interactions must mention the two interacting categories, the prosodic and a segmental one. A standard way of stating such constraints is to relativize a markedness constraint to a prosodic domain (Smith 2005). Such relativized constraints contain a featural markedness component that is understood to apply within a particular prosodic category. The constraint is violated whenever the material that violates the segmental part is found in the prosodic domain mentioned by the constraint.

Examples of such constraints are given below. The constraint (8)a penalizes aspirated segments in non-head syllables within a foot. I will write this constraint as ASPIRATE/σ́ below for brevity. The harmony constraint (8)b enforces agreement in the feature [nas] within a stress foot. The reduction constraint (8)c prohibits mid vowels from occupying unstressed positions.

(8) a. *[spread gl]/NONHEAD
'No aspirated segments in the weak syllable of a foot'
b. AGREE[nas]$_{\varphi}$
'All segments within a foot have the same value of the feature [nas]'
c. *MidV/NONHEAD
'No mid vowels in the weak position of a foot'

The theoretical underpinnings of such complex constraints are subject to debate. De Lacy (2002) proposes to derive constraints that regulate the relationship between stress, sonority, and tone from the formal principles of the theory of harmonic scales. Under his proposal, markedness hierarchies such as the sonority hierarchy or the tonal scale (see above) interact with prosodic markedness constraints to produce sets of composite constraints that regulate the relationship between the two domains.

Smith (2005) proposes a two-step SCHEMA-FILTER theory of constraints, where a general process of constraint building creates constraints out of elementary units according to a general constraint schema. This constraint construction mechanism is a purely formal device, without any sensitivity to the substantive content of the constraints. The set of constraints it produces are then subject to functionally-motivated filters that determine which constraints are phonetically and psycholinguistically grounded. Only the constraints that pass the functional filters are available in the grammars of natural languages.

However, no matter what the theoretical basis used to argue for the constraints, any constraint intended to express the relationship between a prosodic and a segmental property must at the very least MENTION those two properties. As a result, in the general case, for each of the prosody-segmental constraints, there are at least two logically possible repairs available, one having to do with segmental phonology, and the other with prosodic phonology. If a given form violates a segmental markedness constraint relativized to a prosodic domain, then there are at least two ways of obviating such a violation. First, one can modify the segmental structure of the form, and second, one can modify the prosodic structure. These two repairs are available for all constraints mentioning a prosodic and a segmental category.

(9) $M_{(\text{prosodic domain})}$ → (1) Fix segments
→ (2) Fix prosody

The upshot of the typological discussion above is that, in the general case, these two solutions are too many. In those cases where stress is observed to interact unidirectionally with segmentism, only the first repair in (9) – the segmental repair – is typologically observed. Segmental features such as aspiration and consonant quality can cater to prosodic preferences, but not vice versa. This asymmetry is the crux of the problem from the point of view of OT.

Let me illustrate the too-many-solutions problem with the example of stress-driven aspiration. The constraint ASPIRATE/σ́, no matter how it is formulated or grounded, mentions a prosodic category (stressed syllable), and a segmental category (aspiration). Therefore, any violation of such a constraint can lead to at least two different repairs: one that modifies the segmental property and one that modifies the prosodic property.

(10) ASPIRATE/σ́ → (1) Aspirate onsets of stressed syllables
→ (2) Put stress on syllables with aspirated onsets

Both repairs result in a surface structure that satisfies the constraint ASPIRATE/σ́, and thus both are predicted by OT to exist in languages. The segmental repair – aspiration attraction to stressed syllables – is observed in languages like English and Mixtec. Some of these languages are mentioned in the preceding section, and an exhaustive list is supplied in González 2003. In the tableau below in (11) I give a hypothetical example of a language with default initial stress and a high-ranking ASPIRATE/σ́, causing both the attraction of the aspiration to the initial stressed syllable from its different location in the input /pit^{h}a/, and the insertion of aspiration on the stressed syllable when it is not in the input /pita/.

(11)

		STRESS INITIAL	ASPIRATE/σ́	DEP-h	MAX-h
/pitʰa/	pítʰa		*!		
	pitʰá	*!			
☞	pʰíta			*	*
/pita/	píta		*!		
	pítʰa		*!	*	
	pitʰá	*!		*	
☞	pʰíta			*	

Tableau (11) illustrates the segmental repair of a ASPIRATE/σ́ violation – the familiar and typologically well-attested pattern of stress-driven aspiration.

The constraint can also lead to the opposite interaction – aspiration-driven stress. If ranked high enough, ASPIRATE/σ́ can cause stress to be attracted to the syllable that has onset aspiration in the underlying form. Consider again a language with default initial stress. Reranking ASPIRATE/σ́ above the stress constraint and above the faithfulness constraint militating against inserting aspiration results in a pattern where stress is attracted away from its default initial position to the syllable that has an aspirated onset in the underlying form. The following tableau (12) illustrates this case.

(12)

		ASPIRATE/σ́	DEP-h	STRESS INITIAL	MAX-h
/pitʰa/	pítʰa	*!			
☞	pitʰá			*	
	pʰíta		*!		*
	píta	*!			*

ASPIRATE/σ́ kills any candidate where a stressed syllable has no aspiration: the fully faithful candidate *pítʰa* and the aspiration-less candidate *píta*. The high-ranking faithfulness constraint DEP-h prevents a shift of aspiration from its underlying form.[2]

[2] The choice made here in favor of the MAX-h and DEP-h constraints over constraints like IDENT[spread gl] does not affect the argument. There must be SOME faithfulness constraint regulating the input-output

The only viable candidate left is *pitʰá*, the form with a non-default stress that has been attracted to the syllable with an aspirated onset.

To illustrate that the default stress in this hypothetical language is indeed on the initial syllable it is enough to see what happens to an input without any aspiration.

(13)

			ASPIRATE/σ́	DEP-h	STRESS INITIAL	MAX-h
/pita/		pítʰa	*!			
		pitʰá		*	*!	
	☞	pʰíta		*		
		píta	*!			

Here, there is no faithfulness difference between the finally stressed *pitʰá* and the initially stressed *pítʰa* – both involve the insertion of an aspiration – and thus the stress constraints decide in favor of the default initial stress.

To summarize, the hypothetical language illustrated in tableaux (12)-(13) has the following pattern.

(14) /pita/ → *pʰíta*
/pitʰa/ → *pitʰá*

This pattern amounts to aspiration-driven stress: a system which is typologically unattested. The unwanted prediction results from the fact that the constraint ASPIRATE/σ́ only states a preferred output pattern, viz. that stress and aspiration should coincide in the same syllable, but not the derivation route by which the pattern is achieved. Both the segmental and the prosodic repairs (9) are theoretically available, but only the former is attested.

The trade-off between the two types of repair is a general property of the constraints in (8) and other similar constraints. The constraint driving nasal harmony within the

mapping of aspiration, and whenever that constraint is ranked high enough, the unwanted pattern stress foot, AGREE[nas]$_{\varphi}$, can be satisfied either by violating the segmental faithfulness constraint – i.e. by applying harmony within the foot – or by violating a prosodic constraint (faithfulness or markedness, depending on whether stress is predictable in the system or not) by moving the prosodic domain boundary to accommodate the segments. For example, a language that makes use of such a repair strategy would have a stress system that assigns, say, penultimate default stress except in words that have a different [nas] value in the final and penultimate vowels – exactly the forms where harmony would be applicable. This hypothetical situation is illustrated below.

(15) Stress in a language with high-ranking AGREE[nas]$_{\varphi}$ and a prosodic repair

a. Default penultimate

/ara/ → ára

/ãrã/ → ã́rã

b. Final in forms where harmony would otherwise apply

/arã/ → arã́

/ãra/ → ãrá

The bizarre system shown in (15), where stress assignment depends on whether the vowels in the last two syllables have the same or different [nas] specifications, is not attested.

In the general case, standard OT is plagued by a systematic too-many-solutions problem in the case of prosody-segmental interaction. The theory cannot account for any case of asymmetry between the ability of stress to condition the distribution of some segmental feature and its inability to be sensitive to that feature. OT predicts there to be systems with aspiration-driven stress, flap-driven stress, fricative-driven stress, glottalization-driven stress, and so forth. As I claimed above, none of these is attested. This unwanted typological prediction will be the subject of my proposed modification of OT in chapter 3.

Finally, let me illustrate this unwanted interaction with a syllable-based constraint, *[+voi]/CODA. Although the problem falls outside of the scope of stress-segmental interactions, it is analogous to the problems discussed so far and, I think, is an especially

---

illustrated in tableau (12) results.

conspicuous example of an acute too-many-solutions problem in OT. The coda voicing constraints can be satisfied in the same two ways as any prosody-segmental constraint: by fixing the segments, or by fixing the prosody.

$*\begin{bmatrix}+\text{voi}\\ -\text{son}\end{bmatrix}$ in Coda → (1) Fix segments
→ (2) Fix prosody

In other words, a constraint like *[+voi]/CODA can trigger two general types of repairs, one affecting the [+voi] part of the constraint and the other affecting the CODA part. More specifically, a prohibition against voiced codas could trigger any of the repairs in (16) below.

(16)	a. Fix segments			
	Devoicing:	/abra/	→	apra
	Epenthesis:	/abra/	→	abəra
	Nasalization:	/abra/	→	amra
	Deletion	/abra/	→	ara
	Metathesis:	/abra/	→	arba
	b. Fix prosody			
	Resyllabification:	/abra/	→	a.bra

The range of segmental repairs predicted by the theory is far greater than the attested repairs: only devoicing is attested as a response to *[+voi]/CODA. This problem need not concern us for now: it has been addressed by Steriade 2001 and used to argue for the P-map theory, to which I will come back in a later section.

What matters in this section, and what Steriade 2001 does not address, is that Standard OT predicts that for constraints like *[+voi]/CODA and other constraints of this type, languages should systematically have a choice between prosodic and segmental repairs. In the case of coda voicing neutralization, the choice is between keeping the offending syllable in the coda but satisfying the constraint by modifying the feature specification, turning the voiced *b* to a voiceless *p*. The prosodic repair would involve satisfying the prohibition of coda voicing by moving the segment out of the coda into the onset of the following syllable, which allows the language to be faithful to underlying voicing.

In the case of devoicing, the language involving prosodic repair would have the following generalization: voiceless segments that are first members of consonant clusters syllabify as codas of the preceding syllables, while voiced segments form complex onsets with following syllables.

(17) a. /apra/ → ap.ra
b. /abra/ → a.bra

This happens because the voiced but not the voiceless segments are prohibited from codas. However, if there is a difference in the syllabification of consonants based on their voicing, it goes the other way: in Icelandic and Ancient Greek, the voiced but not the voiceless segments make the preceding syllable heavy. The standard analysis of such voicing asymmetries in syllabification appeals to syllable contact: *ab.ra* is preferred to *ap.ra* because it has a smaller sonority rise across the syllable boundary.

1.3 Counterexamples

In this section I turn to some alleged counterexamples to the generalization about prosody-segmental interactions, arguing for alternative analyses. Karo stress (section 1.3.1) has been claimed by Gabas (1998) to be sensitive to the location of flaps. A small set of other onset-sensitive stress systems will be briefly discussed in section 1.3.2.

1.3.1 Stress and lenition in Karo

Karo, a Tupí language of Brazil, has been claimed to possess a stress system that directly contradicts the generalization proposed in this section. According to the description of Gabas (1998), stress assignment in Karo is sensitive to the voicing of obstruent stops, in such a way that certain types of stops are preferably located in weak positions of feet. This contradicts the proposal that only stress can cater to segments, not the other way around. In this section I will show that Gabas's analysis of Karo stress is not the only possible one, and that a different analysis that does not involve voicing-sensitive stress is not only possible but preferable.

Karo has contrastive tone (with at most one H or M tone syllable per word), contrastive nasalization, and contrastive voicing. The Karo stress rule as formulated by Gabas is as follows.

(18) a. Assign stress to the syllable with H tone;
b. Else to the syllable with the nasalized vowel;
c. Else to the penultimate syllable if the final syllable begins with [b], [g], or [ɾ];
d. Else to the final syllable.

The most puzzling part of the rule is (18)c: stress appears to be sensitive to the nature of the consonant in the following. There is a general cross-linguistic preference for leniting or voicing posttonic onsets (English, Swahili, Tohono O'odham, Welsh), and Karo stress as presented by Gabas is the reverse of this posttonic lenition: stress is assigned to syllables preceding voiced consonants.

The normal situation is either like in English, where lenition takes place, or like in Russian, where lenition does not take place and the posttonic voiceless obstruents surface faithfully (19)-(20).

(19)

English			STRESS	LENITE	FAITH
/atom/	☞	á[ɾ]om			*
		á[t]om		*!	
		atóm	*!		

(20)

Russian			STRESS	FAITH	LENITE
/atom/		á[ɾ]om		*!	
	☞	á[t]om			*
		atóm	*!		

In both cases the STRESS family of constraints outrank the segmental faithfulness and markedness constraints, the ranking between which determines whether the posttonic *t* is flapped or not. The high-ranking STRESS ensures that the stress assignment rule is not affected by the segmental properties of adjacent consonants.

In Karo, however, the stress constraint seems to be ranked below the markedness constraint. As the following examples show, stress is assigned to the final syllable unless the onset of the final syllable is a flap, in which case stress is penultimate. As the tableau below illustrates, the segmental faithfulness and markedness constraints must outrank the stress assignment constraints in order to ensure that stress is sensitive to the nature of the final syllable onset.

(21) a. /parat/ → [páɾat] 'curimba'
b. /pako/ → [pakó] 'pacu'

(22) a. IDENT-C 'No feature change on Cs'
b. *D/HEAD 'No weak Cs in stressed syllables'
c. FINALSTRESS 'Stress is word-final'

(23)

			*D/HEAD	IDENT-C	FINALSTRESS
parat	☞	páɾat			*
		paɾát	*!		
		patát		*!	
		pátat		*!	*
pako		págo		*!	*
		páko			*!
	☞	pakó			
		pagó	*!	*	

Is this a counterexample to the central generalization that stress does not cater to segmental phonology like stress-sensitive lenition?

In the remainder of this section I will argue against Gabas's assumption that voicing, not stress, is contrastive. A better analysis takes stress to be partially unpredictable, and derives lenition as conditioned by stress. As a result, Karo is not a counterexample to the generalization on stress-segmental interaction, but simply a language like English, Welsh, etc., with a process of posttonic lenition.

The inventory of Karo, as Gabas presents it, is as follows.[3]

(24)	stops:	p	t	c	k	ʔ	
		b			g		
	liquids		ɾ				
	fricatives						h
	approximants	w		y			

Gabas's description unambiguously implies that voiced and voiceless stops do not contrast in onsets of unstressed syllables; however, Gabas does not use this crucial fact for the analysis of voicing and stress. The list of environments where voiced and voiceless stops occur is given below.

[3] Because there is no [d], [ɾ] can be grouped with the voiced stop series. I will assume that [ɾ] derives from *d* under pressure of an independent constraint.

(25) a. Voiceless stops occur everywhere except intervocalic onsets of unstressed σs

V__V́ (in this position voiceless stops are lengthened)

ʔ__V

#__V

V__#

b. Voiced stops occur intervocalically

V__V

In onsets of stressed syllables, voiced and voiceless stops contrast. In all other positions, their distribution is predictable: voiced stops occur when intervocalic, and voiceless stops occur elsewhere. The contrast in voicing in onsets of stressed syllables can lead to EXCEPTIONS to the stress rule (18)c. In the following examples, voiced stops in onsets of final syllables do not cause the stress to be penultimate: it is irregularly final, in contradiction to (18)c.

(26)	acibéʔ	'raiz'
	koɾé^{b}m	'também'
	pagó^{d}n	'amigo'

Some of the occurrences of voiced stops in onsets of stressed syllables do not generate exceptions to (18)c, because the (18)a or (18)b clauses preclude it.

(27)	a.	koɾét
		   |
		   H
			'jacu'	[stress assigned to V with H tone]

	b.	wakə̃́ỹa	'cutia'	[stress assigned to nasal V]
		morĩ́ỹa	'miçanga'	

Furthermore, Karo has a "morphophonemic" process of stop lenition that is sensitive to stress, which Gabas claims to apply across morpheme boundaries. It applies intervocalically before unstressed vowels, and lenites word-initial voiceless stops to voiced stops when not immediately followed by stress.

(28) /p, t, k/ → [b/w, ɾ, g] / V__$V_{[-accent]}$

(29) a. /e-penaoy/ → [ebenaóy] 'você dançou'
/e-tati/ → [eɾatí] 'te trouxe'
/e-kuru?cu/ → [eguru?cú] 'minha saliva'
b. /e-pəgat/ → [epə́gat] 'te queimou'
/i-toy/ → [itóy] 'viu alguém'
/e-kɨgat/ → [ekɨ́gat] 'te pegou'

Gabas treats stops as contrastive in voicing throughout, even thought they only contrast in onsets of stressed syllables. This misses the connection between their distribution, the morphophonemic alternations, and the placement of stress: these three facts are treated as independent of each other on Gabas's analysis, missing the generalization that they all reflect the same preference for leniting posttonic stops. Instead of assuming that the process (28) is limited to derived environments, I propose that it applies across the board, and that stress is partially unpredictable. My reanalysis is summarized below.

(30) a. Stress rules (18)a or (18)b are still in place.[4]
b. If (18)a or (18)b are not applicable, stress is unpredictably either penultimate or final.
c. The lenition rule (28) applies across the board.

The underlying forms and derivations according to this reanalysis are shown below. The tableau is given in (33).

(31) a. /pátat/ → [páɾat] 'curimba'
b. /pakó/ → [pakó] 'pacu'
c. /acibéʔ/ → [acibéʔ] 'raiz'

The analysis requires the following lenition constraint.

(32) *T/NonHead 'lenite in weak position'

[4] (18)b could be improved if nasalized Vs are treated as underlying VN sequences (there are no surface codas except [ʔ]). Stress is then assigned to Vs with H tone, else to closed syllables.

(33)

			FAITH STRESS	MAX[voi]	*T/NONHEAD	DEP[voi]	FINALSTRESS
pátat	☞	páɾat					*
		paɾát	*!				
		patát	*!	*			
		pátat		*!	*		*
pako		págo				*!	*
		páko			*!		*
	☞	pakó					
		pagó				*!	
acibéʔ	☞	acibéʔ					
		acipéʔ		*!			
		acíbeʔ	*!				*
		acípeʔ	*!	*	*		*

There is a problem, however: the unpredictably penultimate/final distinction in stress only exist in forms whose final syllable has a stop onset (Gabas p.c.). In all cases where there is no stop involved, the final syllable is accented.

(34) a. kɔyɔ́ 'juriti'
b. yaʔɔ́ 'calango'
c. maní 'macaxeira'

This is not predicted by my analysis, because FAITHSTRESS is high-ranked, and there is no constraint that would prevent stress from mapping faithfully in *acibéʔ* but not in #*máni*.

However, this is not a problem that is unique to my reanalysis: on Gabas's analysis, the existence of exceptions like *acibéʔ* but not #*máni* is also accidental. In an OT implementation of Gabas's proposal, FAITHSTRESS must also be high-ranked in order to ensure that *acibéʔ* surfaces faithfully.

(35) Gabas's system

			FAITH STRESS	MAX[voi]	*D/HEAD	DEP[voi]	FINALSTRESS
acibéʔ	☞	acibéʔ			*		
		acipéʔ		*!			
		acíbeʔ	*!				*
		acípeʔ	*!	*			*
mani	☞	maní					
		máni					*!
máni		maní	*!				
	(☞)	máni					

Despite the problem of not accounting for the absence of exceptions like #*máni*, the reanalysis of the Karo stress system shows that it is not necessary to treat it as segment-sensitive. The reanalysis that takes stress rather than voicing to be unpredictable fares better in that it expresses the connection between three facts that were unrelated in Gabas's analysis: the distribution of voiced/voiceless stops, the 'morphophonemic' rule of stop voicing, and the stress placement rule.

1.3.2 Onset-sensitive stress

In a small class of languages, stress is sensitive to the presence or absense of onsets on the initial syllable (Davis 1988, Goedemans 1997, González 2003). Such systems include Aranda, Banawá, Iowa-Oto, and typically have the following stress generalization: stress is assigned to the initial syllable unless it is onsetless, in which case stress falls on the second syllable. In other words, stress is repelled from the initial onsetless syllables.

Second, in at least one language, Alyawarra, there is a variation on this pattern: stress is repelled not only from onsetless syllables, but also from glide-initial syllables (Goedemans 1997: 4). The Alyawarra type may be treated on a par with the Aranda type if the glides can be argued to be in the syllable nucleus rather than in the onset.

This small and constrained set of cases does not present a serious counterexample to the generalization that stress can only be sensitive to quantity, tone, and sonority. As

Goedemans (1997) has shown, these systems can be analyzed using the constraint ALIGN(FT, ONS), calling for feet to be left-aligned with a syllable onset. I will not further discuss these onset-repelling systems in this thesis.

CHAPTER 2

PREVIOUS APPROACHES TO TOO-MANY-SOLUTIONS PROBLEMS

2.0 Introduction

In the previous chapter I surveyed the typology of prosody-segmental interactions and argued that the systematic asymmetry in the relationship between prosodic structure and segmental features poses a too-many-solutions problem for OT. This problem has not been systematically addressed in the literature. Apart from an indirect reference to it in de Lacy 2003 (see section 2.2 below), prosody-segmental interactions have escaped the notice of phonological theory.

However, the challenge that the typological data poses for the theory is not entirely new. Several sets of cases where a significant generalization about input-output mappings produces a too-many-solutions problem have been uncovered in recent literature. The problems are, in some ways, analogous to the ones discussed here. In each case, a markedness constraint that states a SURFACE preference is observed to provoke fewer repair strategies than the theory would predict. As Steriade (2001) noted, the problem is akin to a phonological conspiracy in the sense of Kisseberth 1970, although in a new guise. Kisseberth observed that from the point of view of derivational theories, phonological rules conspire to create less marked output structures. Conversely, from the point of view of OT, constraints conspire to avoid certain dispreferred input-output mappings.[5] This way of understanding too-many-solutions problems diagnoses their true origin: OT's radical claim that ALL significant phonological generalizations are located in output structures is too strong. At least some generalizations lie in input-output mappings, and phonological theory must be equipped to handle them.

[5] I will come back to a more detailed discussion of such 'conspiracies' in Chs. 3 and 4.

In this chapter I will discuss three important attempts to handle too-many-solutions problems in OT: Wilson's Targeted Constraints theory (Section 2.1), de Lacy's fixed ranking proposal (Section 2.2), Steriade's P-map hypothesis (Section 2.3), and. In each case, my focus will be on attempting to apply the proposal to prosody-segmental interactions. I will show that the proposed approaches cannot deal with the data in a general way: Wilson's theory fails empirically; Steriade's theory cannot handle covert prosodic structure because of its overly phonetic grounding; and de Lacy's proposal, empirically the best of the three, runs into problems with fixed rankings that are in some cases contradictory and in other cases too parochial.

2.1 Targeted constraints

Wilson (2001) addresses the too-many-solutions problem found in consonant cluster simplification. Apart from the effects of sonority and of morphological structure, Wilson claims that there are no other factors that influence which of two consonants in a cluster will be deleted under simplification. The only influencing factor is a consonant's position in the cluster, and in the configuration VC_1C_2V, it is always the first consonant C_1 that is deleted.

(1) FIRST CONSONANT DELETION (Wilson 2001: 148)
Across languages, deletion processes that apply to intervocalic biconsonantal clusters consistently delete the FIRST consonant (schematically, $VC_1C_2V \rightarrow VC_2V$).

In OT, consonant deletion in clusters is driven by a markedness constraint that militates against sequences of consonants, call it CLUSTERCOND. This constraint is antagonistic to the faithfulness constraint MAX-C that penalizes any deletion of an input consonant. Given the input /VC_1C_2V/, these two constraints cannot in principle choose between the desired output [VC_2V] and the non-occurring output [VC_1V]. Other constraints must make that decision. Generalization (1) cannot be accounted for so long as there is at least one constraint that favors C_1 over C_2, because then the factorial typology will contain the pattern of deletion where C_1 rather than C_2 surfaces. Clearly, such a constraint exists: it could be any one of many markedness constraints that favor C_1 over C_2. The tableau illustrating the unwanted deletion is shown in (2) below.

(2) **M** incorrectly causes deletion of C_2 (Wilson 2001: 148)

	VC_1C_2V	CLUSTERCOND	MAX-C	**M**
	VC_1C_2V	*!		
☞	VC_2V		*	*!
☹	VC_1V		*	

Wilson's point is that markedness IN PRINCIPLE must be made incapable of making the decision about which consonant is deleted and which is preserved in cluster simplification.

> "According to generalization (1), a consonant is deleted or preserved based solely on the POSITION that it would occupy in the cluster. But previous OT approaches to consonant deletion predict that the decision about which consonant deletes will instead be made – either universally or as a typological option – on the grounds of MARKEDNESS". (Wilson 2001: 148).

This is a too-many-solutions problem: for a given marked configuration, $/VC_1C_2V/$, OT predicts more repairs than are attested in the typology. The task of Wilson's Targeted Constraints is to render markedness constraints incapable of affecting the cluster simplification pattern.

Wilson's solutions, couched in Steriade's Licensing by Cue theory, is to target constraints to a specific repair. In the configuration VC_1C_2V, the first consonant is deficient as compared to the second in terms of perceptual cues. In general, prevocalic consonants have better cues due to the release burst and formant transitions. Wilson's targeted markedness constraint specifically calls for the deletion of the perceptually weaker non-prevocalic consonant C_1, while deleting the stronger consonant C_2 does not count as a repair of the targeted constraint.

The standard markedness constraint CLUSTERCOND penalizes any candidate that possesses an offending cluster, and it is satisfied by any candidate that does not, no matter what the repair for the cluster is, be it deletion of either consonant, deletion of the whole cluster, or epenthesis. In the harmonic ordering that the constraint imposes, any candidate that contains no cluster is more harmonic than any candidate with a cluster.

(3) a. CLUSTERCOND (Standard)
For any two candidate *x* and *y*, *x* is more harmonic than *y* if *x* contains no consonant cluster and *y* contains a consonant cluster.

b.

VC_1V	$\succ$	VC_1C_2V
VC_2V	$\succ$	VC_1C_2V
VV	$\succ$	VC_1C_2V
VC_1VC_2V	$\succ$	VC_1C_2V

The harmonic ordering asserted by the targeted version of the constraint is weaker: Wilson's NOWEAKCONSONANT (NOWKC) constraint prefers only one specific repair, *viz.* the deletion of the first (weak) consonant in a cluster, while not imposing any harmonic ordering on candidates with other repairs of C_1C_2.

(4) a. NOWKC (Targeted; Wilson 2001: 160)
Let *x* be any candidate and α be any consonant in *x* that is not released by a vowel. If candidate *y* is exactly like *x* except that α has been removed, then *y* is more harmonic than *x* (i.e. $y \succ x$).

b. $VC_2V \quad \succ \quad VC_1C_2V$

The constraint does not impose any ordering between VC_1C_2V and the other candidates.

The harmonic orderings asserted by targeted constraints may not be expressed with standard violation marks, as Wilson points out in the footnote on pp. 162-163.[6] Wilson's notation therefore departs from the traditional way of using asterisks as violation marks. Instead, in each cell of the tableau, starting from the cells for the highest-ranked constraint, for each candidate the harmonic orderings imposed by the constraint against that candidate are recorded, and cumulative orderings are tallied at the bottom of the tableau. Lower-ranked constraints impose their own orderings, and only those orderings count that are compatible with the cumulative ordering passed down from the higher-ranked constraints. Incompatible orderings are parenthesized.

(5)

	$/VC_1C_2V/$	NOWKC	MAX-C	**M**
	VC_1C_2V	$VC_2V \succ VC_1C_2V$!		$(VC_1V \succ VC_1C_2V)$
☞	VC_2V		$(VC_1C_2V \succ VC_2V)$	$(VC_1V \succ VC_2V)$
	VC_1V		$VC_1C_2V \succ VC_1V$!	
	cumul	$VC_2V \succ VC_1C_2V$	$VC_2V \succ VC_1C_2V \succ VC_1V$	

The right-ranked constraint NOWKC only asserts the ordering $VC_2V \succ VC_1C_2V$; it is silent on all other candidates. The faithfulness constraint MAX orders VC_1C_2V above both of the deletion candidates, but one of these orderings, $VC_1C_2V \succ VC_2V$, contradicts the ordering passed down by the higher-ranked constraints. Thus the candidates VC_1V incurs a fatal violation. The markedness constraint's contribution is irrelevant.

Wilson illustrates this with a specific example from Diola Fogny, which I repeat here (Wilson 2001: 165). In Diola, complex clusters are simplified by deleting the first member, e.g. /let+ku+jaw/ → *lekujaw* 'they won't go'. The relevant markedness constraints are the place constraints, *DOR ⟩⟩ *COR. In Standard OT, he higher-ranked *DOR is capable of forcing the deletion of the more marked second member of the cluster /tk/. However, this does not take place if the targeted version of the cluster constraint NOWKC is used.

[6] The reason is that harmonic orderings expressible in terms of violation marks are stratified orders, *i.e.* $a \succ b$ and $a \cong c$ imply $c \succ b$: non-comparable members share order relations. This does not hold for

(6)

	/let+ku+jaw/	NOWKC	MAX	*DOR	*COR
	letkujaw	leku ≻ letku!		(letu ≻ letku)	(leku ≻ letku)
☞	lekujaw		(letku ≻ leku)	(letu ≻ leku)	
	letujaw		letku ≻ letu!		leku ≻ letu
	cumul	leku ≻ letku	leku ≻ letku ≻ letu		

Wilson demonstrates that the addition of the targeted version of the cluster constraint accounts for the typological generalization in (1): the factorial typology of NOWKC, IDENT, MAX, and markedness does not contain any rankings which lead to the first consonant in a cluster being deleted. Furthermore, the typology is not only restrictive, but sufficient, as it allows vowel epenthesis as a repair for consonant clusters.

In the remainder of the section I will attempt to apply Wilson's targeted constraints theory to the problem outlined in the previous chapter. I will show that targeted constraints cannot account for the relevant typological generalizations, and, following McCarthy's (2002) criticism of Wilson's theory, will diagnose the problem with TC.

Consider the example of stress-aspiration interactions. Given a marked configuration where the stress and aspiration are not located on the same syllable, the Standard OT stress-to-aspiration constraint is unable to rule out the repair where the stress shifts to the syllable with aspiration, while typologically, only the reverse repair, aspiration shifting to the stressed syllable, is attested. In the general case, given a marked structure in the input, and two possible repairs, A and X, where A is typologically attested and X is not, the targeted version of markedness constraint that militates against the marked structure of the input needs to assert the ordering of A above the candidate that violates the non-targeted version of same constraint, saying nothing about the other candidates, including X (McCarthy 2002). Thus, in the stress/aspiration case, the TC would assert that any candidate that repairs the violation by aspirating the stressed syllable is more harmonic than the FFC.

---

targeted constraints, *e.g.* $VC_2V \succ VC_1C_2V$ and $VC_2V \cong VC_1V$ but not $VC_1V \succ VC_1C_2V$.

(7) Targeted STRESS-TO-ASPIRATION (T-ASPIRATE)
Let x be any candidate and α a stressed syllable in x. If candidate y is exactly the same as x except that aspiration is located on α's onset, then $y \succ x$.

Assume that the stress placement constraint is the cover constraint PENULT, and the relevant aspiration faithfulness constraint is NOFLOP-*h*.

For the hypothetical input /patʰa/ there are three relevant candidates: the candidate (a) that violates STRESS-TO-ASPIRATION, [pátʰa], with stress in the default position and aspiration realized faithfully, at the cost of misalignment between the two; (b) the candidate [pʰáta], with stress in the default position and aspiration attracted to it at the cost of violating NOFLOP; and (c) [patʰá], with aspiration surfacing faithfully and stress attracted to it at the cost of violating PENULT.

The theory's task is to ensure that (c) [patʰá] is a perpetual loser. Standard OT fails at this task, as there is a ranking ASPIRATE ⟩⟩ NOFLOP ⟩⟩ PENULT where (c) wins. It turns out, however, that replacing ASPIRATE with its targeted counterpart T-ASPIRATE makes no difference: (c) can still win. This is shown in the following tableau.

(8)

		/patʰa/	T-ASPIRATE	NOFLOP	PENULT
	a.	pátʰa	b ≻ a!		
	b.	pʰáta		(a ≻ b) c ≻ b!	
☠	c	patʰá			(a ≻ c) (b ≻ c)
		cumul	b ≻ a	c ≻ b ≻ a	

T-ASPIRATE fatally penalizes the candidate (a), where aspiration and stress are disassociated. The remaining two candidates represent the prosodic and the segmental repairs (b) and (c). The segmental repair violates the segmental faithfulness constraint NOFLOP, because it involves aspiration being attracted from its input position to the onset of the stressed syllable. This causes the unwanted candidate (c) to emerge as optimal. The TC theory thus fares no better than Standard OT in failing to rule out the pathological interactions.

This problem is a general one in the domain of stress-segmental interactions where the typological generalization is that only stress influences segments, but not vice versa. McCarthy's (2002) criticism of the Targeted Constraint theory allows us to see why. Observing that the addition of some reasonable constraints to Wilson's analyses undermines the typological predictions of the Targeted Constraint theory, McCarthy formulates the general conditions when replacing a Standard OT constraint with a targeted constraint will account for a typological generalization.

Suppose there is a markedness constraint M, an input /I/, and three candidates, C which violates M, and two candidates that are repairs for M, A and X. Let us assume there is a too-many-solutions problem: candidate A but not candidate X is attested typologically as a repair for violations of M, as illustrated schematically below.

(9)

			M
/I/		C	*
	☞	M	
	☠	X	

The strategy of the targeted constraint approach would be to replace M with a targeted version T-M that imposes the ordering {A ≻ C|X} on the candidate set. McCarthy shows that this strategy can only ensure that X is a perpetual loser if two conditions hold:

(10) a. A and X are equal in faithfulness, and
b. A is equal to X in markedness on constraints other than T-M.
(McCarthy 2002: 287).

If at least of the two conditions above fails to hold, then there are rankings where X is optimal.

In the case of prosody-segmental interactions under discussion, it is condition (10)a that systematically does not obtain. Consider a markedness constraint M that mentions a prosodic category Π and a segmental property Σ. There are three relevant candidates: candidate C which violates M, candidate A which repairs the violation of M by modifying Σ, and candidate X which repairs the violation of M by modifying Π. Candidate A is the typologically attested segmental repair, and candidate X is the unwanted prosodic repair.

Also, assume that modifying Σ, whatever Σ is, violates the segmental faithfulness constraint FAITH (there must be at least one such constraint). The segmental repair candidate A would violate this FAITH constraint. Neither C nor X would incur violations of faithfulness: the former because it is the fully faithful candidate, and the latter because, by assumption, it involves an unfaithful mapping of prosodic and not segmental structure.

There also must be a prosodic constraint PROS (markedness or faithfulness) which X violates but neither A nor C do: this violation results from 'modifying Π'.[7] Once again, C does not violate it because it is the fully faithful candidate or a candidate with default prosody, while A, by assumption, modifies the segmental but not the prosodic structure in order to satisfy the markedness constraint M.

Finally, the targeted markedness constraint prefers A over C and does not impose any ordering on X. The orderings imposed by the constraints are summarized below.

(11) a. T-M: $\{A \succ C | X\}$
b. FAITH: $\{X \succ A | C \succ A\}$
c. PROS: $\{A \succ X | C \succ X\}$

These orderings indicate that McCarthy's necessary conditions for the targeted constraint to rule out the unwanted candidate do not obtain. Candidates A and X are

[7] If 'modifying Π' means a non-default prosodic structure assignment, then PROS is a markedness constraint. If 'modifying Π' means unfaithfully mapping underlying prosodic structure, then PROS is a faithfulness constraint.

not equal in faithfulness: X satisfies FAITH but A violates it. This means that, ranked high enough, the constraint FAITH can trump the effect of the targeted markedness constraint. The tableau for the general case is given below. It shows that if FAITH outranks PROS, then the unwanted prosodic repair candidate X wins.

(12)

	/I/	T-M	FAITH	PROS
	C	A ≻ C!		
	A		(C ≻ A) X ≻ A!	
☠	X			(C ≻ X) (A ≻ X)
	cumul	A ≻ C	X ≻ A ≻ C	

In other words, replacing the relevant markedness constraint with a targeted version has no effect on the typological predictions of the theory, for the reason that McCarthy made clear: there are constraints that favor the unwanted candidate, and these constraints can trump the effect of the targeted constraint. A and X are not equal in faithfulness, failing to satisfy (10)a: A violates FAITH and X does not.

I conclude that Targeted Constraints do not present a viable strategy for dealing with prosody-segmental interactions.

2.2 Fixed ranking

The most comprehensive attack on too-many-solutions problems as a general challenge for OT comes in Paul de Lacy's work (2003). Recognizing the ubiquitous nature of the problem in phonological interactions, de Lacy proposes a general mechanism of handling cases where a given markedness constraint is observed to condition fewer repairs than OT predicts. He identifies several classes of phonological properties that can and cannot affect each other, mentioning, among other problems, the case of prosody-segmental interactions, although it is not the main focus of the work. De Lacy's strategy is to separate OT constraints into classes based on what type of phonological category they refer to, and to impose fixed rankings between those classes of constraints in order to derive asymmetrical interactions between them. Of all the proposals, de Lacy's constitutes the least radical departure from OT in its formal setup.

In this section I will illustrate de Lacy's approach and attempt to apply it to prosody-segmental interactions. I will argue that the proposal, while empirically sound for many cases, suffers from two weaknesses: first, some cases of prosody-segmental interactions appear to require contradictory fixed rankings, and second, it relies on an arbitrary separation of constraints into categories.

The classes of properties ('representational categories') into which de Lacy divides the phonological world are listed below, together with the classes of OT constraints to which the categories correspond (de Lacy 2003: 2).

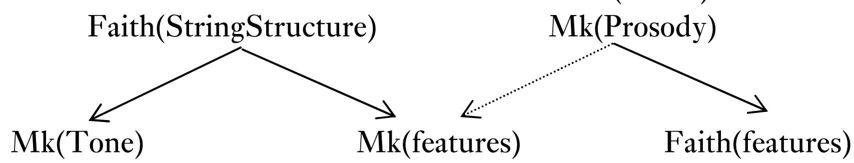
(13) a. String structure (the number of segments in a candidate and their order);
  MAX, DEP, CONTIGUITY, INTEGRITY
 b. Sonority
  Sonority-sequencing constraints
 c. Prosodic structure
  ONSET, NOCODA, FTBIN, etc.
 d. Tone
  OCP
 e. Features
  IDENT; Featural markedness

De Lacy addresses three central generalizations about which categories can and cannot influence which other categories.

(14) a. Feature conditions cannot affect string structure or prosody;
  Prosody-segmental interactions;
  Epenthesis/syncope cannot be sensitive to feature cooccurrence conditions;
 b. Tone conditions cannot affect string structure (can affect prosody);
  Tonal conditions cannot force epenthesis/syncope
 c. Sonority and prosodic conditions can affect string structure and prosody.
  Epenthesis/syncope can be conditioned by cluster phonotactics.

De Lacy's proposal is to separate constraints into classes that correspond to the representational categories and to impose fixed rankings between members of those classes. This is shown in (15) below (De Lacy 2003: 5).

(15) THE PROSODIC PRIMACY FIXED RANKINGS (PPFR)



One of the fixed rankings in (15) directly relevant to the problems at hand is Mk(Prosody) ⟩⟩ Faith(Features). De Lacy's supporting example for this ranking is analogous to one of the cases discussed in the previous chapter: the interaction of stress placement with vowel reduction (2003: 9). In a language with reduction of /o/ to [ə] and a default trochaic stress system, a high-ranking IDENT[round] could cause a non-default stress assignment just in case it would circumvent vowel reduction. Below I repeat de Lacy's hypothetical example. Three constraints are necessary: TROCHEE, enforcing default foot structure, REDUCE, penalizing full vowels in unstressed syllables, and IDENT[round], a faithfulness constraint that makes it more costly to reduce to schwa a rounded vowel like /o/ than an unrounded vowel like /a/.

(16) a. TROCHEE
b. REDUCE 'no full vowels in unstressed syllables'
c. IDENT[round]

Because IDENT[round] prefers the reduction of unrounded vowels, it can produce a pattern where stress is attracted to a rounded vowel IN ORDER to prevent its reduction. The following tableau illustrates this hypothetical mapping of /pato/ to *pətó*, with non-default stress assignment in violation of TROCHEE that caters to IDENT[round].

(17) de Lacy (2003: 9)

	/pato/	IDENT[round]	REDUCE	TROCHEE
	páto		*!	
	pátə	*!		
☠	pətó			*

Recall the analogous examples from the previous chapter in a language with reduction of mid vowels, e.g. a language with a five-vowel inventory in stressed syllables and a three-vowel inventory in unstressed syllables. High-ranking faithfulness could cause stress to be attracted to mid vowels to allow them to escape reduction. De Lacy argues that this unwanted interaction can be ruled out by imposing a fixed ranking between prosodic and featural constraints. If all prosodic constraints like TROCHEE are required to rank above all featural constraints like IDENT[round], this type of interaction is ruled out, as shown in the tableau below.

(18) de Lacy (2003: 9)

	/pato/	TROCHEE	IDENT[round]	REDUCE
☞	páto			*
☞	pátə		*	
	pətó	*!		

De Lacy's strategy is a general one. The problematic interactions between consonantal features and stress discussed in the previous chapter are ruled out as long as the faithfulness constraints for those features are required to rank below the relevant stress placement constraints. For example, in the case of stress-flapping interaction, the prosodic constraints are require to rank above the faithfulness constraint violated by the /t/ ↔ [ɾ] mapping, i.e. IDENT[son].

Along the same lines, de Lacy proposes to rank the prosodic markedness constraints above featural markedness. His examples include prosody-segmental interactions of the type where a prosodic category provides the domain for some segmental process, such as agreement in nasality within a foot or in voice within a syllable. The fixed ranking predicts, in de Lacy's words, that "feature conditions cannot force a change in prosodic structure" (2003: 10). The supporting examples are not unlike the examples presented in the previous chapter. AGREE[voice]$_\sigma$, the voicing agreement constraint relativized to the syllable, could force a voice-driven difference in syllabification if sufficiently high-ranked. An input like /adra/ would surface with a complex onset [a.dra], while /atra/ would come out as [at.ra] to satisfy the agreement constraint.

(19)

			AGREE	IDENT[voi]	NOCODA	*COMPLEX
/adra/	☞	a.dra				*
		ad.ra			*!	
/atra/		a.dra		*!		*
		a.tra	*!			*
	☞	at.ra			*	

As mentioned in the previous chapter, attested examples of voice-dependent syllabification show the opposite effect: it is the voiced, not the voiceless consonants that prefer to syllabify as codas, under pressure of syllable contact. What is problematic in the grammar in (19) is the mapping /atra/ → [at.ra], where it is the constraint AGREE which forces the segment [t] into the coda of the preceding syllable, in violation of NOCODA.

De Lacy argues that, once the ranking Mk(Prosody) (NOCODA, *COMPLEX) and Mk(Features) (AGREE) is fixed, the unwanted prediction does not arise.

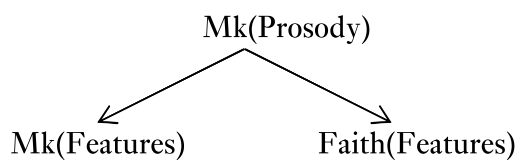
(20)

			NOCODA	*COMPLEX	AGREE	IDENT
/adra/	☞	a.dra		*		
		ad.ra	*!			
/atra/	☞	a.dra		*		*
	☞	a.tra		*	*	
		at.ra	*!			

If, as in (20), the constraint NOCODA ranks above *COMPLEX, the outputs for both /adra/ and /atra/ have a complex onset. If the reverse ranking holds, both inputs produce forms with a coda. In either case, there is no problematic relationship between syllabification and voicing as in (19).

The following diagram summarizes de Lacy's fixed ranking proposal relevant to the discussion here.

(21)



In the remainder of this section I will discuss the consequences of de Lacy's proposal for stress-segmental interactions. I will conclude that not all unwanted interactions are ruled out once sonority-stress effects are brought into the picture. I will also discuss some technical problems with de Lacy's proposal.

Along with prosodic constraints, de Lacy calls sonority constraints the 'heavyweights': the phonological categories they refer to are able to influence all other phonological categories. A summary of the sonority constraints is given below.

(22) Sonority constraints
  a. MARGIN{sonority level}, NUC{sonority level}
  b. Syllable contact
  c. Sonority sequencing

Sonority can affect string structure and segments, and therefore all sonority constraints can rank above MAX, DEP, and IDENT constraints. Evidence for Sonority ⟩⟩ StringStructure comes from cases like the deletion of high (i.e. less sonorous) vowels in Arabic, showing that *NUC {i,u} ⟩⟩ MAX. Likewise, there are many examples where sonority factors cause feature changes.

Prosodic constraints, which include syllable structure and foot structure, are likewise able to cause deletion/epenthesis (and thus outrank MAX and DEP), and feature change (and thus outrank IDENT). Sonority constraints must also be able to outrank prosodic constraints, because sonority factors can block prosodic processes, and because sonority considerations can trump prosodic markedness. For example, the constraint *NUC{i,u} must outrank the prosodic constraints NOCODA and *COMPLEX in order to allow processes like the Arabic deletions in /fihimna/ → *fhimna*, /kibirat/ → *kibrat*. The reverse ranking is needed for languages where vowel syncope is blocked by prosodic factors. However, if *NUC{i,u} is allowed to outrank stress constraints, then stress may be attracted to high (less sonorous) vowels in order to avoid syncope – an unattested system. Consider the hypothetical input /pilata/ in a language with default penultimate stress and high-ranking *NUC{i,u}/WEAK (essentially the Arabic constraint).

(23)

		/pilata/	*NUC{i,u}	MAX	PENULT
	a.	piláta	*!		
	b.	pláta		*!	
	c.	pílta		*!	
☠	d.	pílata			*

The candidate (a) is killed right away because it has a high vowel in the unstressed syllable. The next two candidates which try two different syncope patterns are duly eliminated by MAX, and what remains is the pathological candidate where stress is attracted to the high vowel, *pílata*. Note that if there were no high vowel in the input, stress would surface in its default penultimate position, since *NUC{i,u} would not be active.

The reader might wonder whether the bad predictions in (23) arise not because of the ranking of the sonority constraint above stress, but because MAX outranks the stress constraint. Indeed, moving MAX down the ranking at first blush appears to eliminate the problem.

(24)

		/pilata/	*NUC{i,u}	PENULT	MAX
	a.	piláta	*!		
☞	b.	pláta			*
☞	c.	pílta			*
	d.	pílata		*!	

Here, PENULT is ranked high enough to eliminate the unwanted candidate at a point when the reasonable candidates (b) and (c) are still in the running. Other constraints would then decide between them.

De Lacy is silent on the relative ranking of StringStructure constraints like MAX and the prosodic constraints. Let us assume for the sake of argument that there is a fixed ranking Mk(Prosody) ⟩⟩ StringStructure.[8] As I will show, the unwanted candidate (23)d can still win even with this assumption, so the bad predictions must arise from the ranking *NUC{i,u} ⟩⟩ PENULT.

In order for (23)d to win given that PENULT ⟩⟩ MAX, some high-ranked constraint(s) must eliminate the candidates (b) and (c), and these constraints must be of a type that is allowed to outrank MkProsody. It is not difficult to find ways in which (b) and (c) are worse than (a) and (d): for one thing, they have more complex syllable structure and thus violate NOCODA and *COMPLEX, two prosodic constraints that can outrank PENULT. This allows the candidate (23)d to emerge as optimal even if MAX is low-ranked, as illustrated below.

(25)

		/pilata/	NOCODA	*COMPLEX	*NUC{i,u}	STRESS	MAX
	a.	piláta			*!		
	b.	pláta		*!			*
	c.	pílta	*!				*
☠	d.	pílata				*	

[8] This assumption is a stretch: MAX must be able to outrank prosodic constraints like NOCODA – otherwise codas would be universally deleted. DEP must be able to outrank stress constraints like FTBIN, otherwise degenerate feet would not exist.

Just as in the case (23), the input /palata/ would surface with default penultimate stress *paláta* because this candidate does not violate *NUC{i,u}, and we have in effect a stress system where less sonorous vowels attract stress.

In sum, the hypothetical example of {i,u} syncope shows that if sonority constraints are allowed to outrank stress constraints, bad predictions are made. At the same time, sonority constraints must be freely ranked with respect to syllable structure constraints, for reasons outlined above. This calls for a more fine-grained classification if de Lacy's approach is to be maintained: prosodic constraints must be divided into stress (foot structure) and syllable structure, and the two sub-classes come with different fixed ranking stipulations: one but not the other is required to outrank sonority constraints.

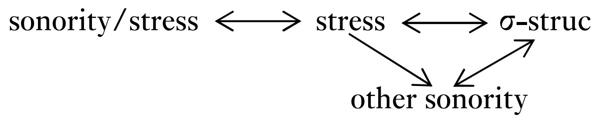
Let me now turn to the constraints not mentioned in de Lacy 2003, but discussed by him at length elsewhere (de Lacy 2002): the stress-sonority constraints. These are responsible for quality-driven stress systems. De Lacy proposed the following hierarchies:

(26) a. *HEAD/{ə}, *HEAD/{ə,i,u}, etc.
b. *NONHEAD/{a}, *NONHEAD/{a,e,o}, etc.

There is no a priori way to decide whether these stress-sonority constraints belong to the stress class or to the sonority class (I will come back below to the general problem of finding criteria for sorting constraints into classes). Counting these constraints as prosodic would make for a simpler overall system, for, unlike sonority constraints, *HEAD[ə] and its ilk must be freely ranked with respect to foot structure constraints, in order to account for both sonority-driven and non-sonority-driven stress. It is not crucial that the decision on where to place sonority-stress constraints be made now.

To summarize the discussion so far, below in (27) is the diagram of the more fine-grained structure with foot and syllable prosody separated into two categories.

(27)



However, it turns out that even (27) is not enough. The *NUC{i,u} constraint discussed above is really a sonority-stress constraint from de Lacy's hierarchy, namely *NONHEAD{ə,i,u} (assuming there is no *ə* in the system). This constraint must clearly rank freely with respect to stress constraints, and yet it can still drive the unfortunate stress system where stress is attracted to less sonorous vowels.

So far I have used a cover constraint for stress to refer to several constraints at once, e.g. PENULT should be properly decomposed into AL-FT-RIGHT and FTBIN. Could it be the case that once the more primitive components of constraints like PENULT are considered, it would become possible to establish more parochial, but less contradictory fixed rankings? Once again, the answer is no: consider the ranking between *NONHEAD{ə,i,u} and AL-FT-RIGHT. The sonority-stress constraint must be able to outrank AL-FT-RIGHT in order to account for stress systems where sonority trumps alignment; e.g. in Kiriwina, stress shifts to the antepenultimate syllable whenever the penult contains a low-sonority vowel, showing that *NONHEAD{ə,i,u} ⟩⟩ AL-FT-RIGHT (de Lacy 2003). Conversely, the reverse ranking is needed for a host of languages where sonority plays no role in main stress assignment, such as English or Latin. And yet, despite the fact that AL-FT-RIGHT and *NONHEAD{ə,i,u} must be freely ranked with each other, the pathological stress system of (25) is predicted to exist.

This shows that the problem of asymmetrical prosody-segmental interactions cannot be solved by fixed rankings in the general case: de Lacy's proposal accounts for some cases, but leaves out others that are exactly analogous and call for a unified analysis.

Apart from the empirical incompleteness of de Lacy's account, it faces a difficulty that ultimately makes it a useful summary of the data rather than an explanatory account. One has to be able to decide *a priori* which of the several classes of constraints a particular constraint belongs to. This classification is easy enough with REPRESENTATIONS: we can all agree on what is a feature, what is a tone, what is a segment, and what is prosodic structure. Elementary constraints can be classified straightforwardly as well: the IDENT constraints all refer to features, the MAX and DEP constraints, at least in some versions of OT, refer to segments and string structure.

However, in practice, most useful OT constraints straddle representational boundaries. Such as the constraints of interest for prosody segmental interactions: they all mention two elements, one prosodic and one segmental. Such are the *CODA/[+voice], AGREE[nas]$_{\varphi}$, and many others of the same type. Likewise, the sonority-stress constraints mention both a stress category (headedness) and a featural or sonority category (location on the sonority scale). The same goes for tone-stress constraints.

How to decide where these hybrid constraints should belong? De Lacy's way, which he does not justify, is to count as featural (i.e. belonging to MkFeatures) those constraints that mention features, possibly together with some other property. So, *[+voi] and AGREE[nas]$_{\varphi}$ count as MkFeatures, even though the latter also mentions prosody; AL-FT-RIGHT and NOCODA count as prosodic. Also non-featural are the stress-sonority and stress-tone constraints like *NONHEAD{ə,i,u} and *HEAD-L.

So there seems to be a double standard: if you mention features at least once, you are a featural constraint, but you are a prosodic constraint if you mention NOTHING BUT prosody. 'Features' here must be construed narrowly to exclude tone and sonority. This strategy is required for empirical reasons, but its basis is unclear. De Lacy in effect replaced one stipulation ('features cannot influence prosody') with another ('feature-prosody constraints are MkFeatures').

In sum, de Lacy's fixed rankings proposal is a first serious attempt to handle too-many-solutions problems as a general challenge for OT. It helps to draw a more clear picture of what the generalizations are, but is neither empirically complete nor truly explanatory.

Let me now briefly discuss how de Lacy's fixed ranking proposal might handle the subject of Chapter 4 of this dissertation, the typology of vowel epenthesis and vowel syncope. As I will argue in Chapter 4, syncope and epenthesis are restricted in terms of environments in which they apply. There is a significant procedural generalization about these two processes: epenthesis universally serves to resolve consonant cluster phonotactics, while syncope targets weak (unstressed, unparsed, etc.) vowels. Conversely, epenthesis is not used as a response for violations of metrical constraints such as *CLASH, *LAPSE, FTBIN, and NONFIN. The only case where epenthesis is used for something other than relieving a marked consonant cluster is minimality-driven epenthesis: in some languages, subminimal words are supplied with an epenthetic vowel in order to bring them up to the minimal size. On the other hand, metrically driven syncope can never target stressed vowels. In particular, it cannot be used as a repair for violations of SWP. I leave a detailed discussion of the empirical issues for Chapter 4; for now, I assume the generalizations as stated here.

Let us see what kinds of fixed rankings must be imposed in order to capture these generalizations. The fact that epenthesis is used as a repair for cluster phonotactics means that DEP-V is freely ranked with respect to the constraints against marked clusters. These constraints, as we will see in Chapter 4, include syllable structure constraints such as NOCODA and *COMPLEX, and syllable contact constraints. On the other hand, DEP-V must outrank *CLASH, *LAPSE, FTBIN, and NONFIN, in order to prevent epenthesis as a repair for metrical markedness. The fact that epenthesis is a possible repair for minimality suggests that DEP-V is freely ranked with respect to GRW=PRW.

As for MAX-V, the fixed rankings are quite different. On the one hand, MAX-V must universally outrank SWP in order to prevent stressed vowel syncope. On the other hand, this constraint must rank freely with respect to the other metrical constraints, such as PARSE-σ, in order to produce attested cases of metrically driven syncope. This picture is summarized below: the lines indicate fixed rankings, and the absence of a line connecting two classes of constraints means that the two classes must be freely ranked.

(28)

	DEP-V	MAX-V
	|	|
SYLLSTRUC	METRICAL	SWP

Although empirically sound, fixed rankings with this amount of detail become difficult to justify. More seriously, the important generalization on the ENVIRONMENTS of syncope and epenthesis that will be the subject of Chapter 4 is lost in a picture like (28).

2.3 The P-map

The two solutions presented in the previous sections, Wilson's Targeted Constraints and de Lacy's system of fixed ranking were formalist proposals insofar as they attempted a modification of either how the constraints are interpreted (Wilson), or the ways constraints are allowed to interact (de Lacy), with a limited or no attempt to ground the proposals in anything outside the theory.

Another general attempt to limit the power of OT to exclude systematic overprediction of repair strategies comes in Steriade's (2001) P-Map theory. Steriade takes a more directly functionalist view: the formal modifications of OT that she proposes rather mechanically follow from perceptual factors. This section will be devoted to the application of Steriade's ideas to the problems at hand, with special attention to the general conditions when a perceptually-based theory like the P-Map can handle too-many-solutions problems.

Since Lombardi (2001[1995]), it has been known that many laryngeal processes like coda neutralization and final devoicing involve a too-many-solutions problem when analyzed in Standard OT. If a language disprefers final voiced obstruent stops, then THE ONLY way that such stops can be eliminated in the output is by voicing neutralization. However, there are many other imaginable repairs that would also remove the violation of the markedness constraint against final voiced stops, *[voi]#: making the stop into a sonorant, epenthesizing a vowel after it, or deleting it altogether would surely do. And yet, none of these is attested. To be sure, final nasalization, epenthesis, and deletion are common enough, but NOT AS A RESPONSE to final voiced obstruents: if a language deletes final voiced stops, it will also delete voiceless stops as well.

And yet, Standard OT with freely ranked MAX, DEP, and IDENT constraints does not account for the privileged status of devoicing relative to other changes. The rankings that give both the attested pattern of final devoicing and the three impossible patterns are given in (29)a and (29)c-b, respectively. The constraints involved are the standard DEP, MAX, and IDENT constraints, in conflict with the markedness constraint *[voi]#.

(29) a. /tab/ → [tap]
DEP-V, MAX-C, IDENT[nas] ⟩⟩ *[voi]# ⟩⟩ IDENT[voi]
b. /tab/ → ☠ [tam]
DEP-V, MAX-C, IDENT[voi] ⟩⟩ *[voi]# ⟩⟩ IDENT[nas]
c. /tab/ → ☠ [tabə]
MAX-C, IDENT[voi], IDENT[nas] ⟩⟩ *[voi]# ⟩⟩ DEP-V
d. /tab/ → ☠ [ta]
DEP-V, IDENT[voi], IDENT[nas] ⟩⟩ *[voi]# ⟩⟩ MAX-C

Steriade's observation is that preventing IDENT[voi] from ranking above any of the other relevant faithfulness constraints rules out the unattested patterns. The research program is then to couch this fixed ranking in external perceptual factors.

The general strategy is to link the ranking of faithfulness constraints in a grammar to speakers' knowledge of perceptual similarity of potential outputs. This knowledge is encoded in the P-Map, which serves as an interface between speakers' phonetic knowledge and phonological grammars. The novel hypothesis of Steriade's is that given a choice of repair strategies for a marked structure, speakers pick those strategies that involve the least perceptible deviation from the input. In other words, the more perceptibly unfaithful a mapping is, the more cost it incurs, and other things being equal, the least costly modification of the input should be chosen. "Other things" here refers to the markedness violations: another way to state Steriade's hypothesis is that if two outputs fare equally on some markedness constraint, the one that involves a less perceptible deviation from the input is more harmonic.

Applying the proposal to laryngeal interactions, Steriade first demonstrates what the perceptual similarity of the relevant forms is – in other words, constructs a P-Map, arguing from confusion matrices and data on imperfect rhyme. The speakers' knowledge of similarity, in turn, comes from "daily experience with confusability", as well as, potentially, other sources. Steriade arrives at a scale of perceived similarity, with voicing distinctions being less perceptible compared to other feature changes, epenthesis, and deletion. Now there is an explanation for why deletion and other processes like those in (29)c-b are not viable repairs for final voicing: speakers do not pick those repairs because there is always a less perceptible alternative available, namely devoicing.

This explanation depends on [tab] being more similar to [tap] than to [tam], [tabə], or [ta]. The phonological grammar reflects the speakers' knowledge that [tab] is more similar to [tap] than to [tam] though fixed ranking of positional faithfulness constraints. Namely, the constraint against final devoicing, IDENT[voi]/__#, is fixed to rank lower than IDENT[nas]/__#, the constraint against changing the nasality of the final segment, and both constraints rank lower than MAX and DEP.

Is the P-Map applicable to segment-prosody interactions? The structure of the problem is similar to the too-many-solutions situation that Lombardi and Steriade faced: there is a markedness constraint which militates against some configuration, and two ways to avoid a violation of that constraint. Of these two repairs, only one is typologically attested.

The markedness constraint is the familiar prosody-segmental constraint, and the two repairs are the prosodic and segmental modification. In order for the P-map to account for such a too many solutions problem, we would need to establish that the unwanted prosodic repair systematically incurs a greater perceptual cost than the segmental repair.

First, however, a technical issue must be addressed. In the exclusively segmental domain of laryngeal interactions, Steriade could talk about the perceptibility of deviations from INPUTS, because the relevant properties are always present in the input.

The fixed ranking derived from the P-map then held between faithfulness constraints and determined which modifications of the input were better (i.e. less perceptually costly) than others.

This setup sometimes, but not always works for prosody-segmental interactions. Consider first the systems where prosodic structure is contrastive: footing or headedness is encoded in the underlying forms. The generalization then is that input prosodic structure is not modified in order to cater to segmental features. This situation is easily expressible in P-map terms. Take the familiar example, the hypothetical input /pát^{h}a/ with lexical initial stress and aspiration on the second syllable. The fully faithful candidate for this input, [pát^{h}a], would violate the markedness constraint that calls for aspiration to coincide with stress; the two repairs are shifting aspiration or stress from their input locations. A P-map solution to why only the former but not the latter repair is attested would be to establish that the form *pát^{h}a* is more perceptually similar to *p^{h}áta* than to *pat^{h}á*. Here the fully faithful candidate *pát^{h}a* serves as a kind of baseline with which comparisons for perceptual similarity are made. If the perceptual distance facts can be established, then the fixed ranking between faithfulness to stress and faithfulness to aspiration would hold, NOFLOPSTRESS ⟩⟩ NOFLOP-*h*, making it universally more costly to shift stress than to shift aspiration, insofar as both achieve the same markedness results.

The same argument will not quite work with a nearly identical case, where stress is not lexical but predictable, and the conflict is between a stress-segmental constraint and other stress markedness constraints. To modify our example slightly, take the input /pat^{h}a/ in a system with default penultimate stress (rather than lexically specified stress). In order for the P-map to work, we must somehow decide which repairs forced by ASPIRATE incur the least perceptual cost. However, unlike in the faithfulness example, is not obvious which candidate is the baseline of comparison for perceptual similarity, because not all relevant properties are present in the input: stress is not supplied from the underlying form but assigned by constraint raking.

For now let me simply make a workable stipulation about what the baseline should be, so that we can move on to the substance of the P-map proposal rather than the technical details.

(30) To determine the best (i.e. least perceptually costly repairs) forced by a markedness constraint M, the baseline of comparison is the most harmonic candidate that violates M.[9] The candidates which satisfy M are then compared with the baseline, and the most perceptually similar one is the best possible repair.

In the example at hand, we are trying to determine the best way to repair violations of ASPIRATE. The most harmonic candidate that violates this constraint is *pát^h^a*: it is fully faithful to segments and perfect on stress markedness. This candidate is the baseline of comparison. The next step would be to compare candidates that satisfy ASPIRATE (*p^h^áta* and *pat^h^á*) and determine which of the two is more perceptually similar to *pát^h^a*. The answer would then determine the relative fixed ranking of the constraints violated in the two candidates, PENULT and NOFLOP-*h*, making one universally more harmonic than the other.

Let us assume for now that there is a way to technically extend the P-map approach to pure-markedness cases like most stress-segmental interactions. In the remainder of this section I will discuss the substantive side of Steriade's proposal with respect to the issues at hand.

The general ideology of the P-map theory is surface-oriented, in that surface perceptibility of features is the only factor that determines the choice of repair processes to apply. The conflict that speakers aim to resolve is between their own interests, i.e. articulatory pressures expressed in markedness constraints, and the interests of the hearer, i.e. avoidance of perceptible deviations from established lexical norms.

[9] This definition is superficially similar to McCarthy's definition of the Sympathetic candidate.

> "The view presented here is that speakers are actively concerned with avoiding *perceptible* deviations from established lexical norms, but they are otherwise not averse to linguistic innovation, insofar as it remains covert. […] The P-map serves as the instrument differentiating more from less perceptible innovations". (Steriade 2001: 18)

If Steriade's general proposal is to be extended to the interactions between segments and prosody, the task would be to establish that the attested repairs involve less perceptible deviations than the non-attested ones. Conversely, if it can be established that a certain prosodic repair involves a perceptually greater unfaithfulness than a segmental repair, then the P-Map theory predicts that my generalization should not hold: precisely in those cases, the prosodic repair should be chosen.

Let me start with the limiting case. Unlike most segmental processes, any type of prosodic restructuring can be *covert*, i.e. involve no perceptible consequence at all. On Steriade's preview expressed in the quote above, only *perceptible* deviations are disfavored, while speakers "are not averse to linguistic innovation, insofar as it remains covert". The upshot of this claim is that covert prosodic restructuring should always be preferred to overt segmental changes.

To be more specific, consider coda devoicing again. Suppose that a language has a general dispreference of complex margins, so that inputs like /apra/ are mapped to outputs like [ap.ra]. This indicates the ranking *COMPLEX ⟩⟩ NOCODA. Now suppose that there is a constraint against coda voiced stops, *[+voi]/CODA. There are two logically possible ways of resolving inputs like /abra/: either not have a voiced stop or not have a coda, i.e. either to fix segments or to fix prosody. This would give the outputs [ap.ra] and [a.bra], respectively.

/abra/ → (1) [ap.ra] (Fix segments)
/abra/ → (2) [a.bra] (Fix prosody)

The choice between (1) and (2) depends on the relative ranking of IDENT[voi] and *COMPLEX: if the faithfulness constraint is low-ranked, then the output contains a devoiced stop, but if the faithfulness constraint is high ranked, the output contains a complex syllable margin. This ranking will not affect the mapping of /apra/, as illustrated in the following tableau.

(31)

			*[+voi]/CODA	IDENT[voi]	*COMPLEX	NOCODA
/abra/	☞	[ap.ra]		*(!)		*
	☞	[a.bra]			*(!)	
		[ab.ra]	*!			*
		[a.pra]		*(!)	*(!)	
/apra/	☞	[ap.ra]				*
		[a.pra]			*!	

Once again, the relevant property, syllabification, is not present in the input. In order to use to P-map to determine which repair for the constraint *[+voi]/CODA and input /abra/ is preferable, we need a baseline of comparison. In accordance with the stipulation (30) above, the baseline of comparison is the most harmonic candidate that violates the constraint, namely *ab.ra*. The question then becomes, which of the two repairs, *ap.ra* or *a.bra*, is perceptually more similar to *ab.ra*. Crucially, resyllabification does not necessarily involve any perceptible differences: [ab.ra] and [a.bra] can be pronounced identically. Of the two repairs, the prosodic repair *a.bra* involves no perceptible deviation from the baseline, and the theory predicts it should be universally preferred. The P-map theory not only fails to rule out the unwanted output in this example, but actively ENFORCES it.

There is a way out in examples like these for a proponent of surface-oriented theories like the P-map, which lies in denying the existence of syllable-final devoicing and even syllable structure altogether (Steriade 1999). On such views, the devoicing process that maps /abra/ to [apra] involves preconsonantal rather than coda neutralization, so shifting the syllable boundary to put the offending voiced stop into the onset of the following syllable would not constitute a repair of the markedness constraint violation.

In a more radical theory that denies the existence of syllable structure, the very option of prosodic restructuring does not exist.

However, as I will show below, the P-map solution does not work in the general case, even in situations where it cannot be saved by denying the role of syllable structure in laryngeal neutralization. The general prediction of the P-map is that covert prosodic structure, if it exists at all, should NEVER condition segmental unfaithfulness, because modifying covert structure is by definition perceptually cheaper than modifying anything overt. This is the limiting case of a situation where the choice of the prosodic repair involves a demonstrably smaller perceptual deviation from the input than the segmental repair, and thus, on the P-map view, should be preferred to the segmental repair.

There are many cases in languages when the prosodic boundaries of constituents larger than a syllable are inaudible. For example, the right boundaries of trochaic feet, so long as they are not followed immediately by another trochaic foot, cannot be heard. A form with initial stress and no secondary stresses has at least four possibilities of footing, which are identical on the surface but represent different metrical constituent structure (Kenstowicz 1994).

(32)	a.	(σ́ σ) σ σ σ	Single trochee
	b.	(σ́ σ) (σ σ) σ	Trochees with no secondary stress realized
	c.	(σ́ σ σ) σ σ	Ternary foot
	d.	(σ́ σ σ σ σ)	Unbounded foot

Although there is no difference in the realization of stress between the four possibilities of (32)a–d, the covert placement of the right foot boundary could have potentially overt consequences for segmental processes sensitive to footing.[10] Such consequences have been documented for a number of Panoan languages by González (2004), where various vowel and consonant alternations are sensitive to covert feet. One such case is the

[10] Further evidence for the reality of covert constituent boundaries comes from metrics: the meter of Plautus makes reference to covert right foot boundaries of trochaic feet (Blumenfeld 2004).

process of glottal stop deletion in Capanahua (Loos 1967, Safir 1979, Loos 1986, González 2002). Primary stress in Capanahua falls on the second syllable if it is heavy, otherwise on the first syllable. The only reported realization of stress is high pitch, which falls on the stressed syllable and spreads rightward up to the penultimate syllable. Thus Capanahua words can take one of the two shapes schematized in (33).

(33) a. Initial stress: (σ̲́ σ́) (σ́ σ́) σ
b. Peninital stress: σ (σ̲́ σ́) (σ́ σ)

There are no H/L pitches alternating in a binary fashion that would overtly signal the presence of secondary stress feet posited in (33), nor is there any other correlate of secondary stress (Loos 1967, González p.c.). However, the covert feet must be present, because they condition the deletion of the glottal stop. As shown in the following examples in (34), glottal stops are deleted foot-finally, so that the morpheme sequence /ɾaʔ-taʔ/ surfaces either as [ɾaʔ-ta] or as [ɾa-taʔ], depending on whether it is preceded by an even or an odd number of syllables.

(34)	a. /ʔotʃiti-ɾaʔ-taʔ-ki/ 'dog-prob-decl-cert'	(ʔotʃi)(ti-ɾa-)(taʔ-ki)	'it is probably a dog'
	b. /ʔotʃiti-ma-ɾaʔ-taʔ-ki/ 'dog-neg-prob-decl-cert'	(ʔotʃi)(ti-ma-)(ɾaʔ-ta-)ki	'it is probably not a dog'
	c. /ta-mani-ʔt-wi/	(tama)(niʔwi)	'take a step'
	/ka-riʔbi-wi/	(kari)(biwi)	'go again'

According to the predictions of the P-map, examples such as the Capanahua alternation should not exist, because Capanahua sacrifices segmental faithfulness to covert prosody. The prosody-segmental interaction in Capanahua is in the same direction as the prosody-segmental interactions in English and other languages where the conditioning prosodic structure is audible and overt: it is the segments that cater to prosody, not the other way around.

One possible answer to this argument against the P-map would appeal to the possibility that covert features structure is, in general, dispreferred. Languages strive for ways of overtly realizing the covert metrical structure, which is normally done through pitch, intensity, and duration. Any such realization of metrical prominence would involve some unfaithfulness, i.e. some perceptual cost. For example, if a language lengthens its stressed vowels, it makes it more difficult to maintain a length contrast in those positions. There is a tradeoff between how clearly the metrical prominence is realized and how much faithfulness cost is involved in its realization. Consonantal processes sensitive to foot structure are then simply another way of signaling the position of the foot boundaries, and Capanahua, in this line of thinking, can be said to employ segmental cues to metrical structure rather than prosodic ones. The glottal stop deletion process, on this view, would not be an unmotivated instance of unfaithfulness, but rather a functionally motivated means to supplying the listener with cues to the covert metrical structure.

However, the Capanahua example illustrates that in the general case this counterargument is invalid. Glottal stops are contrastive in the language, while secondary (as well as primary) stress is not. Thus there should not be any functional motivation to signal non-contrastive boundaries of feet. And yet, the covert and predictable feet are preserved, while the unpredictable glottal stops are deleted, resulting in their neutralization with ∅.

Furthermore, the argument that segmental alternations like the glottal stop deletion in Capanahua have a function in signaling covert metrical structure makes a more general prediction that such alternations should not occur, or at least occur less frequently, in languages with clearly audible dynamic stress. This is, however, not the case; according to the survey in Bybee et al. (1998), the reverse generalization is true: the more perceptible the metrical structure, the more likely the language is to have segmental alternations sensitive to it.[11]

[11] This generalization indicate that at least some of the segmental alternations conditioned by prosody are in fact conditioned not by the abstract constituent structure, but by its overt correlates such as intensity,

The P-map makes an even stronger and more clearly false prediction. Because innovations in covert structure have no perceptual cost, they should be used whenever possible if the covert innovation leads to a faithful mapping of overt segmental information. For example, if a language has a process of vowel reduction in unfooted syllables, then there is always the possibility of constructing a covert secondary stress foot over those reducible syllables in order to save them from reduction. According to the letter of the P-map proposal, this is what ought to happen: covert innovations are free, while the danger of reducing vowels has perceptual consequences. Likewise, if a language (e.g. Mam, England 1983) has a constraint against multiple long vowels in the same word – i.e. a constraint against long vowels in any position except in the head of a stress foot – nothing prevents the language from constructing covert feet over long vowels in other positions of the word in order to allow them to faithfully surface. Again, on the P-map proposal this should be the solution of choice, because there is no cost in the covert innovation, while the benefit of building a secondary stress foot would be the faithful rendition of a vowel that would otherwise shorten.

Looking at less extreme examples where prosody is covert, the exact predictions of the P-map view are difficult to tease out: there is nothing GENERAL about how prosody and segments ought to interact that follows from the setup of the theory. Rather, the predictions are to be evaluated on a case-by-case basis, determining in each type of interaction whether being unfaithful to prosody or unfaithful to segments incurs the least perceptual cost. One general expectation that comes out of the P-map is that the behavior of a prosodic system may depend on how it is phonetically realized: high-intensity dynamic stress may be more salient and thus more resilient to unfaithful mappings than a pitch accent. This prediction is not borne out. In general, the

---

duration, and high pitch. This is clearly true of consonantal alternations like Verner's Law, where the voicing of obstruents depends on the accentuation of neighboring syllables. Phonetically plausible accounts of Verner's Law appeal to high pitch, rather than purely abstract prominence, as the conditioning factor. Likewise, aerodynamic factors are likely explanations for English-style alternations like aspiration and flapping. At the same time, as argued in detail in González 2003, there are many alternations that are conditioned by foot structure not reducible to phonetic correlates of prominence. Capanahua, with no phonetic correlates at all, is one such example.

directionality of prosody-segments interaction does not depend on whether the system is one of dynamic stress or pitch accent. Pitch accents cannot be sensitive to purely segmental properties in the same way that dynamic stress cannot be.

One ready diagnostic for perceptual similarity appealed to by Steriade is imperfect rhyme. On the assumption that more perceptually similar strings make better rhymes, it is easy to deduce relative similarity of features and segments from their (non-)use in rhyme. This method can also shed some light on the predictions of the P-map for prosody-segmental interactions. In dynamic stress systems imperfect rhyme practically never involves a stress mismatch, even if very liberal with segmental differences between members of a rhyming pair. In both English and Russian, voicing, place of articulation, and vowel quality may all mismatch in imperfect rhyme, but stress almost never can. The fact that prosodic agreement within the rhyme is an absolute requirement suggests that differences in dynamic stress are more perceptible that differences in segments, leading to the P-map style argument that segmental repairs are preferred to prosodic repairs.

While this prediction is correct for dynamic languages, the argument is not so easy to make if we look at the prosodic systems where prominence is realized in some other way. In Serbo-Croatian the metrical heads surface as pitch accents, and, unlike in English and Russian, imperfect rhyme CAN involve accent mismatch (Eekman 1974).

(35) tvôm ~ mȉlom
kȁmen ~ plên
starína ~ sȑbina (Kostić, Eekman 1974)

And yet, the generalizations about the interaction of pitch accent systems with segments is the same as for dynamic stress systems. This shows that while the imperfect rhyme facts may go with the phonetic details of a prosodic system, the role of that system in the grammar and its interface with segmental phonology does not.

The reason why these incorrect predictions of the P-map arise is due to the surface-oriented nature of the theory: the only factor that plays a role in phonological structure, according to the P-map, is the phonetic realization of a feature. However, metrical constituent structure appears to possess a level of abstractness with which the P-map cannot deal.

CHAPTER 3

ASYMMETRICAL INTERACTIONS AND ASYMMETRICAL CONSTRAINTS

3.0 Introduction

The three approaches to too-many-solutions problems discussed in the previous chapter do not treat them as symptomatic of significant generalizations about input-output mappings rather than outputs. Because this systematic challenge to OT is not recognized, previous attacks on these problems, apart from de Lacy's general proposal, see them as an assortment of individual difficult cases and analyses. Surface-oriented theories such as the P-map, whose predictions are closely tied to the phonetic detail of each particular case, face this difficulty in an especially acute form, unable to provide a unified account for a systematic pattern across various examples with vastly different phonetics. The aim of this chapter is to remedy the situation by offering a general analysis of non-surface-based generalizations that treats them as a systematic phonological problem.

The chapter is organized as follows. I begin by arguing against a potential objection that the generalization about input-output mappings in the domain of prosody-segmental interactions is not a synchronically active one and should not be dealt with by a theory of UG. Concluding that it is a significant phonological generalization, I then discuss the place of non-surface generalizations in a surface-oriented theory. I then propose a new class of markedness constraints that refer directly to input-output mappings and penalize candidates that involve typologically undesirable processes. I then discuss a new proposal for dealing with sonority-driven stress, and conclude the chapter with a brief discussion of some applications of my theory, setting the stage for a detailed investigation of syncope and epenthesis in chapter 4.

3.1 Is the generalization synchronic?

The shortcomings of the previous approaches discussed in the last chapter, I suggest, are due to a general property of OT: its radical claim that all significant phonological generalizations are located in the character of output forms. The data presented in this dissertation suggest that there are phonological patterns that are not attributable to purely surface markedness pressures. I take it as an argument against the radicalism of OT and in favor of a theory of phonology that affords both kinds of generalizations their place. However, such a claim does not automatically follow from the observed existence of typologically observable procedural generalizations, because any phonological generalization can potentially be synchronically inert but arise through external factors such as diachronic tendencies. If this is the case, a synchronic theory of linguistic competence is not responsible for dealing with such a generalization. Thus, before I turn to developing the mechanism of OT constraint interpretation that permits the theory to account for non-surface-based generalizations, it is necessary to show that the relevant claims about typology constitute real synchronic generalizations that the theory should be able to address.

Theoretical linguistics has long entertained the idea that not all observed generalizations can be due to innate UG-based properties of grammar (Blevins 2004). It is not the case that every typological observation about language reflects innate principles; gaps in the typology do not automatically call for UG-based explanation. There are systematic classes of patterns in languages that are absent for other reasons: because they are difficult to learn or to process, or because there is no conceivable diachronic path that leads to them. Just how much of grammar can be explained externally is, of course, subject to vigorous debate in linguistics. I will steer clear of this controversy and adopt the view that is mainstream at least in current phonological theory – that AT LEAST SOME properties of grammars have an external explanation, and that there are independent ways of telling whether a particular generalization is a true universal or an accidental fact that can be explained without invoking innate properties of grammar. This dichotomy between two kinds of generalizations is most clearly laid out in Kiparsky 2004, who proposes several ways of deciding whether a given systematic observation about grammars is a UG-based 'universal' or a 'typological generalization' that can be explained without reference to UG.

If it can be shown that generalizations such as the asymmetries in prosody-segmental interactions are not true universals but arise due to external factors like diachrony, they would fall outside of the responsibility of a synchronic theory of phonology like OT. The question, thus, is whether the relevant facts fall in the category of 'universals' or 'typological generalizations' in the sense of Kiparsky 2004.

Applying several of Kiparsky's criteria to prosody-segmental interactions shows that the generalization cannot be dismissed as spurious. The claim made here is that it is real in the sense that it must be expressed in a theory of the knowledge of language.

Kiparsky argues that true universals are IRREVERSIBLE in the sense that language change cannot subvert them. For example, the sonority hierarchy can be argued to be part of UG on the grounds that sonority-based generalizations are not reversed by potential sound changes. There is a common tendency in languages for the high-sonority vowel *a* to attract stress more strongly than other vowels. This fact follows from the assumption that the sonority hierarchy plays a role in metrical prominence. There is a common sound change by which *a* raises to *ə*, a vowel that is lowest on the sonority scale. If this sound change occurs in a language where stress is attracted to every *a*, then the resulting language would have a rule by which it is the schwa – the least sonorous vowel – that is the most prone to be stressed. The claim is more subtle than a simple assertion that such a sequence of events cannot occur. Rather, Kiparsky argues, whenever it does occur, the result will be, from a synchronic point of view, an arbitrary stress system with stress encoded in the lexicon, rather than a system based on the productive generalization that the vowel *ə* is obligatorily stressed. In other words, the sonority hierarchy is a true universal enshrined in UG because it constrains language change.

This criterion of IRREVERSIBILITY indicates that the asymmetry between prosody-segmental interactions discussed above in this dissertation is a true universal rather than a typological generalization. The argument is similar in structure to Kiparsky's argument for the irreversibility of the sonority hierarchy. There is a reasonably likely path of change that can lead to a system that violates the generalization about prosody-segmental interactions, but this path does not seem to be taken by languages in a way that leads to a synchronic system contradicting the proposed generalization. As explained in detail in Chapter 1, there is a common tendency for certain consonantal features to be sensitive to stress. This tendency leads to a synchronic state where the location of the consonantal feature can be predicted in terms of the location of stress. From this starting point, the change that is required for a language to develop a system where stress is attracted to that consonantal property is a RULE REVERSAL reanalysis of the system. If the learners misinterpret the coincidence of stress and some other feature like aspiration as a result of stress attraction rather than aspiration attraction, the result would be a system with aspiration-driven stress. However, speakers appear not to mislearn stress systems in such a way.

Let me illustrate this point with Karo, the language discussed in Chapter 1 as a potential counterexample to the generalization. Recall that Gabas's (1998) analysis of the stress system, outlined in (1) below, made use of stress repulsion from a syllable with a voiced or lenited segment (18)c.

(1) a. Assign stress to the syllable with H tone;
b. Else to the syllable with the nasalized vowel;
c. Else to the penultimate syllable if the final syllable begins with [b], [g], or [ɾ];
d. Else to the final syllable.

Although stress is normally final, it shifts to the penult just when the final syllable's onset is a voiced stop or a [ɾ], as illustrated by the two representative items below.

(2)	a. /pako/	→	[pakó]	'pacu'
	b. /parat/	→	[páɾat]	'curimba'

However, I argued in Chapter 1 that the analysis that requires stress repulsion from syllables with voiced onsets is mistaken. A better analysis makes use of lenition in onsets of unstressed syllables. The main argument was that such a process is independently necessary to account for the alternations at the prefix-stem boundary. On my analysis, forms like those in (2) have unpredictable stress, penultimate or final, and lenition applies to onsets of unstressed syllables to derive flaps and voiced segments from voiceless ones in the ordinary way.

This example illustrates that the relationship between stress and consonant quality in Karo is potentially ambiguous, and thus might have caused the speakers to mislearn it as lenition-driven stress rather than as stress-driven lenition. And yet, the rule reversal path was not taken. This appears to hold in general: even when linguists might be tempted to posit a reverse interaction between segments and stress, speakers do not do so, because a synchronic principle of UG blocks a likely diachronic path. I take this as an argument in favor of the reality of the generalizations I argued for in Chapter 1.

The argument, once again, must be more subtle than simply saying that the particular diachronic path does not occur. Rather, the diachronic sequence of events is precluded by UG from creating a SYNCHRONIC state that flouts the proposed generalization. A rule reversal would, by definition, create a synchronic system that violates the claim of unidirectionality in prosody-segmental interactions. There are other conceivable paths by which a system like that in Karo can fall apart; the claim is that none of them lead to a situation where stress is driven by flapping.

The second relevant criterion for distinguishing universals from typological generalizations is CONVERGENCE. If a generalization arises as an accidental effect of diachronic tendency, it has only one diachronic source. True universals, on the other hand, should be observed to emerge from a variety of converging diachronic pathways. This criterion also suggests that the generalization about stress-segmental asymmetries is a true synchronic universal. The synchronic situation where stress conditions the location of a consonantal property can arise through at least two ways diachronically. First, consonantal features such as aspiration can develop in prosodically strong positions, as happened in English. Second, prosodically strong positions can protect aspiration from deleting (Vijayakrishnan 1999).

The third argument in favor of the stress-segmental asymmetries being true universals is the fact that they emerge outside of the grammar proper. I refer the reader to Section 3.5.2 below, where psycholinguistic evidence for the special status of prosody in grammars is discussed.

I take the arguments above to indicate that phonological theory cannot be released from the responsibility to deal with the process-based generalization about the interaction of prosodic structure and segments.

3.2 Locus of the generalization

Assuming that there is a generalization to explain, let me now detail the place of asymmetric phonological interactions in an output-oriented theory. This discussion will set the stage for the formal analysis in subsequent sections.

A generalization about processes is a claim that a given marked structure is repaired in fewer ways than the freely interacting, output-oriented constraints that OT would make one expect. In the context of discussing a too-many-solution problem in the domain of word-final devoicing, Steriade points out that the problem can be viewed as the mirror-image of the conspiracy argument first suggested by Kisseberth (1970), and later given prominence in the OT literature. Just as, from the point of view of a derivational theory, rules "conspire" to create a particular output structure, from the point of view of OT constraints "conspire" to produce only some input-output mappings but not others. Steriade's suggestion, reproduced in the quote below, will be further developed in Chapter 4, in the discussion of the contexts of vowel syncope.

> "Kisseberth's (1970) insight that conspiracies arise when the sound system aims at a specific target structure via multiple means can lead one to ask the same question, in the context of rule-based phonology: if the rule of final devoicing aims to eliminate final voiced obstruents, why aren't there rules of final obstruent nasalization, deletion, metathesis or post-voiced obstruent epenthesis?" (Steriade 2001: 6).

As discussed above in Chapter 1, the too-many-solutions problem in the case of prosody-segmental interactions arises because misalignment of certain prosodic and segmental categories, such as stress and aspiration, is repaired only by modifying the segments but never by modifying the prosody. Following Steriade's suggestion, then, from the point of view of OT this can be thought of as a conspiracy: OT constraints conspire to allow one but not the other among the potential repairs for the marked structure.

Steriade's approach to the too-many-solutions problem (Section 2.3) was to elaborate the faithfulness apparatus by introducing fixed rankings, grounded in the theory of similarity. The standard OT setup of input-output faithfulness constraints freely interacting with each other and with the markedness constraints is incapable of dealing with the process-based conspiracies. "The P-map's broadest claim", she writes, "is that the range of systematic, cross-linguistically invariant differences [in perceptual similarity that give rise to the too-many-solutions problem go] beyond the capabilities of current theories of correspondence" (2001: 6). We can add to this outlook on faithfulness a similar pessimistic view of standard OT markedness.

In general, in order to determine whether a candidate violates an OT markedness constraint, one is not required to look at any information outside of the candidate itself. Conventional markedness constraints are statements that hold of particular output forms, regardless of their derivational provenance, and of any other property of the grammar in question. The basis of the innovation proposed in this dissertation is the argument that procedural generalizations, if they are to be expressed in markedness constraints, require those constraints to access information outside of the surface form.

To take a specific example, given the situation where stress and aspiration are misaligned, an output constraint stating simply that aspiration should be located on the same syllable as stress is not enough. In order for a markedness constraint to determine whether a given surface form is the result of a licit repair strategy, the constraint must have access not only to the structure of that surface form, but also to the derivational provenance of the relevant properties in it. This means that, when faced with an output where stress and aspiration cooccur on the same syllable, such as a form like *pit^há*, the constraint must be able to distinguish whether both properties have been inherited from the input, or whether one has been attracted to other. Obviously, this information is not available solely from the structure of the candidate itself, which contains no clues about whether or not stress has been attracted to aspiration. That information can be obtained only by looking at the rest of the grammar, i.e. at the ranking of the other constraints governing the distribution of stress. Specifically, in order to see if stress has been attracted to aspiration in a form like *pit^há*, we need to know how stress would have behaved had aspiration not been present. In other words, if an OT markedness constraint is to be endowed with the power to penalize unwanted repairs, its violation profile must have access to the rest of the grammar.

Let me illustrate this crucial point in more detail. If there is a markedness constraint that is violated just in case an undesirable process like aspiration-driven stress has taken place, identical candidates in different grammars would have different violation profiles of this constraint. Suppose there are two languages, *A* and *B*, which have different stress systems. Language *A* stresses the final syllable by default, while language *B* stresses the initial syllable. An input without aspiration like /pita/ would surface with final and initial stress in the two languages, respectively. Now suppose an input with aspiration, /pit^ha/, surfaces in both languages with final stress, *pit^há*. In this case stress can be said to be attracted to aspiration in language *B*, because the presence of aspiration makes a difference for the location of stress. On the other hand, language *A* has no stress-driven aspiration, because the location of stress is the same in the forms with and without aspiration.

(3) Language *A* Language *B*

/pita/	→	*pitá*	/pita/	→	*píta*
/pit^{h}a/	→	*pit^{h}á*	/pit^{h}a/	→	*pit^{h}á*
(no attraction)			(aspiration-driven stress)		

This example illustrates that any markedness constraint that penalizes unwanted repairs such as stress shifts driven by aspiration must have access to information outside of the candidate under evaluation. The identical form *pit^{h}á* would incur a violation of such a constraint in language *B* but not in language *A*; what determines this assignment of a violation is the general character of the stress system in the language, which depends on the ranking of the remaining constraints.

This need to access the rest of the grammar in order to compute the violations of a constraint is a general property of process-based generalizations. Any theory that attempts to capture such generalizations by means of markedness constraints must endow those constraints with the ability to see outside of the candidate under evaluation. I will now propose a formal mechanism of constraint evaluation that allows markedness constraints access to such information, and thereby affords generalizations about processes their place in the theory.

3.3 Procedural constraints and the typology of repairs

Given the pessimism about the ability of standard OT constraints to directly express non-output generalizations, we might seek alternative routes to standard OT theorizing. The path I will take up in this dissertation attempts to preserve as many properties of OT as possible, but at the same time to endow the theory with the capacity to handle generalizations about properties that are not exclusively located in output structures. In this chapter I will develop one proposal that introduces a new type of constraint into the theory. These new constraints come with a mechanism that interprets their violation profile in a novel way, by accessing information outside of the candidate under evaluation. This move is intended as an exploration of what it would take for a theory to both preserve the parallel structure of evaluation and deal seriously with non-surface generalizations by elaborating markedness constraints.

A few words of caution are in order before I proceed with the analysis. A central claim of this dissertation is that there are certain phonological generalizations which are not in the output structures but in the input-output mappings. In order to afford such generalizations their place in the theory, I will take the most straightforward approach possible: I will introduce a new class of markedness constraints that directly refer to phonological processes and assign extra violation marks to undesirable ones. This is not the only possible strategy. Another approach would be to modify not the markedness but the faithfulness system. The P-map theory, for example, is, from the formal point of view, a theory of faithfulness constraints.

I will argue below that current views of input-output faithfulness cannot deal with the kind of asymmetrical interactions that have been the focus of my discussion. It appears that a very serious modification of either, or both, of the current theories of faithfulness and markedness is in order if procedural generalizations are to be accounted for. I consider the work in this dissertation to be an exploration of what it would take for a 'pure markedness' theory to achieve the stated goals; the possibility of a tenable 'pure faithfulness' approach remains open.

A procedural generalization of the type explored in this dissertation amounts to a prohibition of a certain way of getting from an input to an output. In OT terms, it means that there must be a mechanism for making certain candidates incapable of winning, i.e. making them perpetual losers. I will take the simplest possible approach: I will introduce constraints that directly penalize those candidates that involve undesired input-output mappings. The formal mechanism introduced in the remainder of the chapter will provide an algorithm for finding such candidates, and for ensuring that the violation pattern of the new constraints entails those candidates' perpetual loserdom. The strategy will be to attack the problem head on. The new constraints will state directly the directionality of interaction between two phonological categories, and assign a violation mark to each candidate that involves an undesirable interaction. The formal machinery developed below will provide a mechanism for identifying such candidates from among the candidate set.

Thus, the discussion begins by making more strict the notion of what a 'process' is in the context of OT. I will then formalize the intuitive understanding that constraints cause processes to apply by providing a way to determine which constraint causes which process. This discussion will serve to single out candidates that violate a given generalization about input-output mappings. Once these formal underpinnings are in place, I will introduce a new mechanism of assigning violation marks to candidates, penalizing those that involve undesired processes, and demonstrate how my proposal can account for the typological generalizations under discussion.

3.3.1 The notion of 'process' in OT

In rule-based phonology, the notion of 'process' is self-explanatory. Rewrite rules take one form as an input and return another form as an output. Each rule that applies in the course of a derivation corresponds to a process. The change that the rule inflicts upon a form is transparent in its statement. While the phonological grammar is thought of as a complex machine that turns inputs into outputs, the path between the underlying form and the surface form can be straightforwardly decomposed into elementary steps, each of which results from the application of a rule.

In OT, there is no primitive concept that is a direct counterpart to the stepwise derivations of the earlier theories.[12] There is only one well-defined process in OT, the input-output mapping, which is achieved in a single step. There is no notion of "modifying" an underlying form incrementally to arrive at a surface form. While in rule-based theory the question of whether a particular process has or has not applied in the

[12] McCarthy's theory of candidate chains (McCarthy 2006), an important recent development in OT, is a notable exception. Most relevant for the discussion here is that theory's requirement that every output be reachable from the input via a sequency of discrete markedness-improving steps. I leave for future research the question of how structure-building operations such as stress assignment, and prosodic

course of a given derivation can be straightforwardly answered, it is not directly answerable in an OT derivation. Let me return to the hypothetical example of the two languages in (3). The intuition is clear: stress is attracted to aspiration in language *B* but not in language *A*. In derivational terms, it is clear what would be responsible for this attraction: at some point in the derivational history of *pit^{h}á* in language *B*, there must be a rule that affects metrical structure in a way that is sensitive to the aspiration of some segment in the form. On the other hand, the metrical structure assignment rules of language *A* make no reference to such segmental information. Once again, there is no straightforward translation of this explanation into OT terms, because OT does not have any counterpart of stress-assignment rules. The task of this section is, then, to formalize the intuition about the different interactions between stress and segments in language *A* and language *B*.

Let me begin defining the notion of 'process' in OT with a concrete example of the Latin stress rule, gradually making the discussion more general and eventually arriving at an abstract definition. The Latin stress rule assigns stress to the penultimate syllable if it is heavy and to the antepenult otherwise. Descriptively, a moraic trochee is built as close to the right edge of the word as possible, without incorporating the final syllable, unless leaving out the final syllable would leave too little material to make a non-degenerate foot. This statement in terms of conflicting pressures translates into an OT analysis of this system involves the following constraints and ranking.

(4) a.	AL-FT-R	'All feet right'
b.	FT-BIN	'Feet are binary at some level of analysis'
c.	*μμμ	'Feet are not trimoraic'
d.	TROCHEE	'Feet are trochaic'
e.	NON-FIN	'The final syllable is not footed'
f.	CULMIN	'There is at least one foot in every PrWd'

structure in general, fit into the theory, and whether it can be used to constrain the candidate space sufficiently to rule out the unwanted interactions between prosody and segments.

FT-BIN, CULMIN, TROCHEE
|
*μμμ
|
NON-FIN
|
AL-FT-R

All of these constraints must rank above the prosodic faithfulness constraint MAX-HEAD, because stress in Latin is (almost) entirely predictable, so no matter whether some other syllable is marked as the head in the input, the output corresponds to the Latin stress rule.

Now suppose we add a constraint to the system that calls on stressed syllables to have aspirated onsets, ASPIRATE/σ́. What does it mean formally for this constraint to force or not to force a stress shift?

If, in Standard OT, such a constraint is ranked above the stress constraints in (4), and provided that the aspiration faithfulness constraint is also high-ranked, then stress will surface on any syllable whose onset is aspirated in the input. For example, the input /p^hilippus/ would surface with initial stress, *p^h*íli*ppus*, rather than with the output of Latin stress rule, *p^hilíppus*. Intuitively, the constraint ASPIRATE/σ́ has caused a stress shift because adding it to the grammar has changed the output pattern. If the constraint were not present, stress would have surfaced in its 'normal' position.

I take this intuition as the basis for the formal notion of stress shift. To determine whether a constraint *C* affects stress placement, we will compare the location of stress in the actual output to its location in the alternative grammar with the constraint *C* removed from the ranking. A stress shift will then be said to occur whenever stress is in a different place in the outputs of the two grammars.

In the remainder of this section, I will built up the notion of 'process' incrementally in four steps, as follows. I will begin by defining a correspondence relation between candidates that share an input, making precise the notion of being 'in a different place' in two outputs of two grammars (3.3.1.1). This will lead to a definition of STRESS SHIFT

(3.3.1.2). Having illustrated this particular case of the more general notion, I will then define the concept of DESIGNATED STATE of a phonological object (3.3.1.3), which will allow for a general definition of PROCESS (3.3.1.4).

OUTLINE FOR THE REST OF THE SECTION

- (a) Define correspondence between candidates that share an input
- (b) Define STRESS SHIFT
- (c) Define DESIGNATED STATE
- (d) Define PROCESS

3.3.1.1 Correspondence between candidates that share an input

Once again, the intuition to be formalized is that a constraint *C* causes a stress shift for a particular input if removing *C* from the ranking results in a different stress pattern.

We need to first pin down the concept of 'being in the same place' when we are talking about two candidates in two different grammars. The material in the two candidates must stand in some correspondence relation. Because the candidates share an input, and because each of them stands in a correspondence relation with that input, we can define cross-candidate correspondence transitively, via the input. I rely here on the concept of T-CORRESPONDENCE from McCarthy 2003, where it was defined for two candidates in the same grammar. I reproduce McCarthy's definition below.

(5) **T-CORRESPONDENCE** (Definition) (McCarthy 2003:8)
Let *cand1* and *cand2* be two candidates from input *inp*. Let *s1* be a segment (or other corresponding element) in *cand1* and *s2* be a segment in *cand2*. Then *s1* t-corresponds to *s2* iff *s1* corresponds to some segment *s-inp* in *inp* and *s2* also corresponds to *s-inp*.

This idea can be extended in the obvious way to two candidates from two different grammars, as long as they share an input: two pieces of phonological structure in two candidates stand in a 'cross-grammar' t-correspondence relation as long as they share a correspondent in the input.

This definition is straightforward enough when applied to those aspects of structure that have an input corresopndent. But, of course, epenthetic segments correspond to nothing in the input and hence, by the definition in (5), do not t-correspond to anything. Also outside of the scope of this definition are elements like syllable structure, which may not have input correspondents (Prince and Smolensky 1993, McCarthy and Prince 1997; McCarthy 1999).[13]

At the same time, there is a clear intuitive sense in which syllables in two output forms can correspond: the first syllable of *pʰílippus* from the example above is in some sense 'the same' as the first syllable of *pʰilíppus*, and ditto for the second and third syllables of those forms. This 'sameness', however, does not follow from the definition in (5), because, arguably, those syllables have no correspondents in the input and thus they do not t-correspond to anything. I will work around this problem by stipulating that two syllables t-correspond to each other if the segments that fill their nuclei t-correspond.

(6) T-CORRESPONDENCE (Extension)
Let *cand1* and *cand2* be two candidates from input *inp*. Let σ_1 be a syllable in *cand1* and σ_2 be a syllable in *cand2*. Then σ_1 t-corresponds to σ_2 iff the segments filling the nuclei of σ_1 and σ_2 t-correspond according to the definition in (5).

By this definition, clearly, the first syllables of *pʰílippus* and *pʰilíppus* from the same example stand in a t-correspondence relation, because the vowels in their nuclei come from the same input.

This definition entails that vowel epenthesis and deletion, as well as other syllable-adding operations, do not 'disrupt' t-correspondence between syllables: in the candidate with initial epenthesis *əpʰílippus*, the second syllable *pʰí* t-corresponds to the initial syllable of *pʰílippus* because their nuclei share an input and thus t-correspond. To see

[13] A strong argument can be made that there are in fact faithfulness constraints referring to syllable structure (see Kiparsky to appear). If this is so, then syllable structure does fall under the scope of (5), and requires no special treatment.

how vowel syncope affects t-correspondence, consider the following two candidates sharing an input (7). Assume that the vowels named V_i in the two candidates t-correspond. Then, by the definition in (6), the syllables σ_1 and σ_3 in the two candidates t-correspond, while σ_2 in (7)a t-corresponds to nothing in (7)b, because its nucleus V_2 has no t-correspondent there.

(7) a. b.

3.3.1.2 Stress shift

Recall that we are trying to pin down the intuition that an OT constraint causes a process affecting a piece of phonological structure if removing the constraint from the ranking produces a DIFFERENT result with respect to that structure. The notion of t-correspondence allows us to talk about phonological objects being 'same' and 'different' in two different candidates from two different grammars, so long as those candidates share an input. The words 'same/different' will be used here as shorthand equivalents of 'standing/not standing in t-correspondence'. In particular, we can say that the stress in the form *pʰílippus* is not in the 'same' place as the stress in *pʰilíppus*, because the stress-bearing syllables *pʰi* and *li* do not t-correspond.

This leads me to the notion of STRESS SHIFT, defined for some input and some constraint.

(8) **STRESS SHIFT** (Definition)
Given a grammar *G*, a constraint *C*, and an input /i/, *C* is said to force a STRESS SHIFT for /i/ if the location of stress is different in the optimal candidate for /i/ in *G* and the optimal candidate for /i/ in *G*' that is identical to *G* except that *C* has been removed.

In this sense, the constraint ASPIRATE/σ́ forces a stress shift for the input /pʰilippus/, because the output in the grammar with this constraint is *pʰílippus*, and the output in the grammar without this constraint is *pʰilíppus*, and the stress-bearing syllables in the two forms are different, as explained above. The constraint *[+voi], on the other hand, does not force a stress shift, because its presence or absence in the system has no effect on the location of stress. Likewise, ASPIRATE/σ́ does not force a stress shift for the input /pʰaretra/ 'quiver', because the aspiration is on the same syllable that receives stress by the Latin stress rule (*pʰá.re.tra*), so the output in the grammar with ASPIRATE/σ́ is identical to the output in the grammar without ASPIRATE/σ́.

More generally, the constraint ASPIRATE/σ́, if high-ranked, forces a stress shift in an easily definable set of cases: those where the location of the aspiration is not on the syllable that is the head according to the Latin stress rule.

The statement in (8) defines the process of stress shift caused by a constraint, a particular case of the more general notion of 'process', to which I turn in the next two sections. Note that it is hopeless to define 'process' in terms of the relation between inputs and outputs, because of the principle of Richness of the Base. An input may have the stress on any syllable, e.g. on the last syllable in /pʰaretrá/, but in the output stress will still surface on the antepenult, whether or not the constraint ASPIRATE/σ́ is present in the system. We would not want to say that a stress shift has occurred in this case. What matters is the location of stress in the output relative to its DEFAULT location, not relative to some arbitrary input.

3.3.1.3 Designated state

Let me now generalize this notion of 'process' beyond stress. Informally, a markedness constraint forces some process if this constraint contributes something to the input-output mapping, i.e. if the result would have been different had the constraint not been present.

In the stress shift example, I have referred informally to the 'default' place of stress – its location in the output for a grammar with some constraint taken out. This 'default location' determines whether a constraint has forced a stress shift or not. In order to generalize the definition of 'process' beyond stress, it is first necessary to be more precise about this notion of 'default', and to extend it to phonological categories other than stress.

Any given phonological object, be it a feature, a prosodic constituent, a tone, and so forth, can be affected by a markedness constraint. An object has been affected by a constraint if that object would in some way behave differently if the constraint were not there – i.e. if its t-correspondent in the grammar with the constraint removed is not identical to it. "Differently" here can mean a number of things, depending on the nature of the object in question: it could refer to the value of a feature, or its location, or the association lines of a tone, and so forth.

In order to determine whether a markedness constraint *C* forces a process that affects a phonological object *p*, we need to compare *p* in the grammar with *C* to *p* in the grammar without *C*. I first define this baseline of comparison, the DESIGNATED STATE (or LOCUS) of a phonological object.

(9) **THE DESIGNATED STATE (LOCUS)** (Definition)
Let there be a grammar *G*, a constraint *C*, an input /i/and its output [o] in *G*, and a phonological object *p* in [o]. Then the DESIGNATED STATE (or DESIGNATED LOCUS) of *p* for /i/ and *C* in *G* is *p*'s t-correspondent in the optimal candidate in the grammar *G*' that is identical to *G* except that *C* has been removed.

Because it is often useful to think of the designated state as some priviledged location in a form, I will use the term 'designated state' interchangeably with 'designated locus'. There is no theoretical difference betweeen the two; both depend on the definition (9).

Put simply, the designated state of some object for a given constraint is how that object would behave if the constraint were not present. I postpone the discussion of whether the designated state is uniquely determined by the definition given in (9) until Section 3.3.4.

Let me illustrate this notion of designated state with some examples. First, repeating the pseudo-Latin example from above, the designate state of stress for the input /pʰilippus/ and the constraint ASPIRATE/σ́ is the stress in the optimal candidate in the grammar without ASPIRATE/σ́, i.e. the penultimate stress in *pʰilíppus*, the default location according to the Latin stress rule.

In Latin there is also a constraint against long vowels followed by other vowels, *VV.V. In general, long vowels shorten in this environment. The input /docē-ō/ surfaces as [doceō] 'I teach'; this is the output in the grammar *G* where the constraint *VV.V is present. If the constraint *VV.V were not present in the system (grammar *G*'), nothing would prevent /docē-ō/ from surfacing faithfully as [docēō]. The segment *e* in the output of grammar *G* t-corresponds to the segment *ē* in the output of the grammar *G*', because they share the input /ē/. Thus, the designated state of the segment *e* for the input /docē-ō/ and the constraint *VV.V is the long vowel *ē* in [docēō].

To take another example, there is a constraint against final voiced stops *[+voi]#, active in Russian, which has final devoicing, and inactive in English, which does not. The input /rod/ surfaces as *ro*[d] in English, and the similar input /rod/ 'gender' in Russian comes out as *ro*[t]. The feature [voi] has a designated state for the constraint *[+voi]# and the input /rod/ in both languages, which is the correspondent of that feature in the output of /rod/ in the grammars that differ from the actual grammars of English and Russian in that *[+voi]# has been removed from the system. In both cases, the designated state is the feature [+voi] on the final segment of the word, because in the absence of *[+voi]# nothing prevents the final consonant from surfacing faithfully as *ro*[d].

3.3.1.4 Process

The concept of designated state defines a baseline of comparison that allows us to see whether a given constraint has caused some process to apply to some form. It bears emphasizing that this baseline cannot be determined from the input alone, but must be supplied, as formalized in the definition above, from the output in the alternative grammar with the constraint in question removed. The input fails to provide the crucial clues about a form's default behavior due to ROB, as explained above.

Now with the notion of the designated state of a phonological object we can define what a process is in OT.

(10) **PROCESS** (Definition)
For a grammar *G*, a constraint *C*, an input /i/ and its output [o] in *G*, and a phonological object *p*, the constraint *C* is said to affect *p* if *p* is not identical to its designated state as to location or value. *C* is then said to force a PROCESS affecting *p*.

Under this more general definition, the stress shift forced by the constraint ASPIRATE/σ́ in the pseudo-Latin example above is the process forced by that constraint. An example of a segmental process would be flapping in English, forced by the constraint *[t/d]/WEAK. If this constraint were not present in the system, inputs like /atom/ would surface faithfully without flaps. So, the designated state of the feature [son] of the English word [aɾom] is the *t* in [atom], and thus there is a process of flapping forced by the constraint.

To summarize this section: I suggested a formal way of answering the question of whether a given constraint has forced some process for a given form. This concept rests on a comparison between the actual output with the output in the alternative grammar where the constraint in question is removed. Unlike the eponymous concept in rule-based theories, process in OT is a matter of the whole grammar. In order to determine what processes a constraint forces, we need access to the entire ranking. In the case of stress shifts, for example, we need the entire ranking of the constraints in the language to determine what the designated state of stress is – i.e. its 'default' location. Given this OT-based notion of process, we can begin talking about typological generalizations having to do with input-output mappings an OT markedness constraint can cause. In the next section I introduce a new type of OT constraint whose job is to penalize candidates that involve unwanted processes.

3.3.2 The Implicational Constraint Principle

In this section I introduce the new constraints that will bear the burden of explaining typological generalizations like those discussed earlier in this dissertation. The new constraints will state what is a preferred phonological process rather than what is the preferred output. I will refer to these new constraints as PROCEDURAL markedness constraints. Throughout the discussion, the names of the new type of constraints will be preceded by the sign ✦: procedural markedness constraints will have names like ✦M, while traditional output-oriented constraints will retain names like M.

The task of the new procedural constraints is to penalize unwanted interactions between phonological properties, and thereby to enforce generalizations about input-output mappings. These constraints encode processes, i.e. interactions between two phonological categories. Such constraints will be stated as IMPLICATIONS: they have the general form 'If *P* has property *x*, they *Q* has property *y*'. The formal mechanism I propose will link the asymmetrical statement of such constraints to the asymmetry in interaction between the two properties mentioned, in such a way that the property mentioned in the antecedent can influence the property mentioned in the consequent, but not vice versa. I will propose a way of interpreting such constraints that assigns extra violation marks to those candidates that involve a process affecting the property mentioned in the antecedent of the constraint. As I will show, the effect of my proposal will be that a procedural onstraint 'If *P* has property *x*, they *Q* has property *y*' will have the ability to force a process affecting *Q* but not *P*. This mechanism provides a general way of accounting for procedural generalizations using OT markedness constraints.

If a constraint of the form 'If *P* has property *x*, they *Q* has property *y*' is interpreted as a standard OT constraint, the the direction of implication in the statement makes no predictions about the direction of interaction between the relevant aspects of representation. The constraint is simply a statement about an output structure. In principle, any process that results in the implication being true can be a repair strategy for such a constraint. As I showed in Chapter 1, the constraint ASPIRATE/σ́ 'If a syllable is stressed, its onset is aspirated' can be satisfied by aspiration just as easily as by a stress shift – i.e. by a process affecting the antecedent property of the constraint just as easily as by a process affecting the consequent property. In a language with default initial stress, an input like /pit^{h}a/ can surface with stress on the second syllable to satisfy the constraint ASPIRATE/σ́. The following tableau, repeated from Chapter 1, illustrates the unwanted 'aspiration-driven stress' language. Assuming that stress in this hypothetical language is initial by default, the mapping /pit^{h}a/ → *pit^{h}á* shows that the default can be overridden just in case aspiration is present on a non-initial syllable.

(11)

			ASPIRATE/σ́	DEP-h	STRESS INITIAL	MAX-h
/pit^{h}a/		pít^{h}a	*!			
	☞	pit^{h}á			*	
		p^{h}íta		*!		*
		píta	*!			*

In the more general case, the constraint 'If *P* has property *x*, then *Q* has property *y*', where *P* is any structure (segment, syllable, foot, adjacent consonants, etc.), can be satisfied by shifting *x* as well as by shifting *y*. The following tableau (12) illustrates this too-many-solutions problem in the general case. In the input, the properties *x* and *y* are located on different syllables (as indicated by the subscripts). A procedural constraint states that if something has property *x*, then it must also have property *y*. Different faithfulness constraints are violated by unfaithfully mapping *x* and *y*. Then the factorial typology contains at least three patterns: the fully faithful candidate which violates the markedness constraints, and two unfaithful candidates which violate the two faithfulness constraints. In order to satisfy the procedural constraint, both a solution that modifies *x* and a solution that modifies *y* are available. As argued in Chapter 1, these two solutions are too many.

(12)

			If *P* has property *x*, then *Q* has property *y*	FAITH-*x*	FAITH-*y*
/$\sigma_x\sigma_y$/	☞	$\sigma_x\sigma_y$	*		
	☞	$\sigma\sigma_{x,y}$		*	
	☞	$\sigma_{x,y}\sigma$			*

My solution is to replace the constraints ASPIRATE/$\acute{\sigma}$ and others responsible for similar too-many-solutions problems with procedural versions. The violation pattern of these constraints will ensure that, for tableau (12), no candidate with non-default (non-initial) stress will be able to win, and, likewise, in the general case in tableau (12), the candidate $\sigma\sigma_{x,y}$ will not be a potential winner. Forms such as *pitʰá* in tableau (12) and $\sigma\sigma_{x,y}$ in tableau (12) will become perpetual losers.

The strategy I use is to attack the problem in the most straightforward way. The phonological generalization behind the losing status of candidates like *pitʰá* and $\sigma\sigma_{x,y}$ is best stated in terms of input-output mappings. Procedural constraints directly refer to those mappings. The Implicational Constraint Principle introduced below serves as the mechanism for assigning violation marks to procedural constraints. Constraints like '✦If *P* has property *x*, then *Q* has property *y*' will receive violation marks not only when the implicational statement is false, but also when the constraint affects the phonological property mentioned in its antecedent. The notion of 'process' in the following statement refers to the definition in (10) above.

(13) **The Implicational Constraint Principle** (henceforth **ICP**)

A candidate *c* violates a procedural constraint '✦If *P* has property *x*, then *Q* has property *y*' iff:

a. In *c*, *P* has property *x*, and *Q* does not have property *y*, **OR**

b. The constraint forces a process affecting *x* in *c*.

Let me unpack this definition. It relies on the notion of 'process' introduced in Section 3.3.1, in the definitions (9) and (10). By these definitions, when a constraint affects *x* in a candidate, the actual state of *x* differs from the designated state of *x*. The designated state, in turn, is the location (and value, if it is a feature) of *x*'s t-correspondent in the optimal candidate in the grammar with the implicational constraint taken out of the ranking. In other words, what (13) says is that an implicational constraint is violated not only by those candidates where the implicational statement is false, but also by candidates in which the antecedent property mentioned in the constraint is not in its designated state.

As implied by the definition in (13), all procedural constraints subject to the ICP are assumed to be binary – each candidate incurs either 0 or 1 violations of such constraints, even if both of the conditions in (13) are met. This is needed to ensure that the typological consequences of the proposal hold. As I will show below in Section 3.3.4, this way of assigning violation marks ensures that the unwanted candidates can never win.[14] Before I take up that formal argument, for the remainder of this Section and in Section 3.3.3 I illustrate how (13) applies to concrete examples.

Let us take the tableau (12), repeated below as (14), to illustrate the proposal.

[14] I leave for future research the question whether, in case where one of the conditions of the ICP is violated in two different loci in the same candidate, it still incurs only one violation of the constraint. What matters here is that violation of both of the conditions of (13) does not lead to two asterisks.

(14)

			Aspirate/σ́	Dep-h	Stress Initial	Max-h
/pitʰa/		pítʰa	*!			
	☞	pitʰá			*	
		pʰíta		*!		*
		píta	*!			*

As discussed at length above, the problematic predictions of OT arise due to the action of the constraint Aspirate/σ́, which trumps the stress constraint(s) and forces a non-default stress to be assigned to the form. We can now replace this ordinary OT constraint with its procedural counterpart, ✦Aspirate/σ́, that would penalize not only any mismatch between stress and aspiration, but also candidates where stress shifts are used to repair that mismatch.[15] In other words, the new procedural constraint ✦Aspirate/σ́ requires something more than just for stress and aspiration to be located on the same syllable. 'Stress' is the antecedent property of this constraint, the property that is subject to clause (13)b of the ICP. The designated state of stress is determined with reference to the winning candidate in the evaluation where ✦Aspirate/σ́ is not present. The tableau for this evaluation is given below in (15); it shows that the winner has stress on the first syllable. The constraint taken out of the ranking is shaded.

(15)

			✦Aspirate/σ́	Dep-h	Stress Initial	Max-h
/pitʰa/	☞	pítʰa				
		pitʰá			*!	
		pʰíta		*!		*
		píta				*!

Thus, the designated locus of stress for the input /pitʰa/ and the constraint ✦Aspirate/σ́ is on the first syllable. According to the ICP, any candidate where a

[15] I postpone until Section 3.5 on constraint grounding the general discussion of how to determine which constraints are ordinary OT constraints and which are procedural.

stress shift has occurred violates this procedural constraint. We have the tools to find such candidates: a stress shift has occurred in any item whose stressed syllable does not t-correspond to the initial syllable in (15). All such candidates now receive extra violation marks by the ICP. To make the tableaux easier to read, I mark any such violations, incurred by clause (13)b of the ICP, with the sign ✦. There is no theoretical difference between ✦ used as a violation mark and the regular * violations. The following tableau shows the ✦ASPIRATE/σ́ constraint with the new violation pattern.

(16) DS of stress: initial syllable

			✦ASPIRATE/σ́	DEP-h	STRESS INITIAL	MAX-h
/pit^{h}a/		pít^{h}a	*!			
		pit^{h}á	✦!		*	
	☞	p^{h}íta		*		*
		píta	*!			*

The candidates *pít^{h}a* and *píta* incur ordinary violation marks of the ✦ASPIRATE constraint, because they contain stressed syllables with an aspirated onset. The candidate *pit^{h}á* incurs a ✦ violation of the constraint ✦ASPIRATE/σ́ because the location of stress is different from its designated location. The winner now is *p^{h}íta*, a candidate with default initial stress.[16]

Let me emphasize that a procedural constraint that is subject to the ICP is not a typical OT constraint: its violation profile for a given candidate set depends on the ranking of other constraints in the grammar. The reason for this is that such constraints are meant to rule out certain processes that are unattested typologically, and the notion of 'process' in OT, as discussed at length above in Section 3.3.1, only makes sense with reference to the ranking of the constraints in the language.

3.3.3 Tudanca Spanish

In Chapter 1, I went over laxness harmony in Tudanca Spanish, a case where a process is sensitive to the boundaries of a prosodic constituent. Harmony-prosody interactions present the familiar too-many-solutions problem for OT. Here I return to those Spanish data to illustrate the operation of the ICP in excluding unwanted interactions. First, to recapitulate the facts: final high vowels are lax, and laxness spreads leftward until it reaches the stressed syllable. Examples, taken from Flemming 1994, are shown below, with capitalization indicating laxness. The forms in (17)a have a final high vowel [U], which is lax. All vowels preceding it, up to the stressed syllable, are also lax; that laxness does not spread beyond the stressed syllable is shown by the form [se(kÁlU)]. The forms (17)b show the corresponding alternants where the final vowel is not high, and therefore not lax, and (7)c shows laxness spreading in forms where the stress falls on a syllable other than the penult.

(17)	a.	(pÍntU)	'male calf'	b.	(pínta)	'female calf'
		(čÍkU)	'boy'		(číka)	'girl'
		se(kÁlU)	'to dry him'		se(kálo)	'to dry it'
	c.	o(rÉgAnU)	'oregano'			
		(pÓrtIkU)	'portico'			
		ra(kÍtIkU)	'rachitic'			

Recall that the OT analysis of these facts made use of the following constraints.

(18)	a.	AGREE[tense]$_{\varphi}$	'All vowels within a foot have same value of [tense]'
	b.	*[+high, +tense]#	'No final tense vowels'
	c.	STRESS	Cover constraint for penultimate stress
	d.	IDENT[tense]	

[16] Once again, the typological argument that no stress shift candidate can EVER win will be made in section

The constraint driving laxness harmony within the stress foot, AGREE[tense]$_{\varphi}$ (18)a, can be satisfied either by violating the segmental faithfulness constraint – i.e. by applying harmony – or by violating a prosodic constraint by moving the prosodic domain boundary to accommodate the segments. This is the usual situation where at least two repairs are predicted to exist, given a constraint that mentions two categories (a prosodic and a segmental one). In the actual language, the constraint (18)a forces a segmental repair: harmony applies within the stress foot. This is illustrated in the following tableau, repeated from Chapter 1. The form ending in a high vowel surfaces with laxness harmony, /sekalu/ → [sek(ÁlU)], while the form ending in a non-high vowel surfaces faithfully, /sekalo/ → [sek(álo)].

(19) Tudanca Spanish

			STRESS	AGREE[tense]$_{\varphi}$	*[+high, +tense]#	IDENT[tense]
/sekalu/		se(kálu)			*!	
		seka(lÚ)	*!			*
		se(kálU)		*!		*
	☞	se(kÁlU)				**
/sekalo/	☞	se(kálo)				
		seka(lÓ)	*!			*
		se(kálO)		*!		*
		se(kÁlO)				*!*

Just as in other prosodically sensitive harmony processes, in Tudanca Spanish reranking the STRESS constraints to the bottom of the hierarchy produces a stress system where the stress feet are built in such a way that harmony does not get a chance to apply. In such a grammar, the harmony constraint AGREE[tense]$_{\varphi}$ would be satisfied not by an unfaithful mapping of segments, as in the actual Tudanca Spanish, but by violating prosodic markedness. This hypothetical situation is illustrated in the tableau below. The form ending in a non-high vowel still surfaces with the default stress, /sekalo/ → [se(kálo)]. However, the form /sekalu/ has a final stress, [seka(lÚ)], because the default stress would violate either the harmony constraint if harmony does not apply

3.3.4. For now, the examples simply serve to illustrate the action of the ICP.

(*[se(kálU)]), or the constraint against final high tense vowels if tensing does not apply (*[se(kálu)]), or the faithfulness constraint militating against changes in the [tense] feature (*[se(kÁlU)]). This leaves the candidate with final stress as the only viable option, as illustrated below.

(20) Pseudo-Tudanca Spanish

			AGREE[tense]$_{\varphi}$	*[+high, +tense]#	IDENT[tense]	STRESS
/sekalu/		se(kálu)		*!		
	☞	seka(lÚ)			*	*
		se(kálU)	*!		*	
		se(kÁlU)			*!*	
/sekalo/	☞	se(kálo)				
		seka(lÓ)			*!	*
		se(kálO)	*!		*	
		se(kÁlO)			*!*	

Let us now see how my proposal rules out the system in (20). Intuitively, while the "job" of the constraint AGREE[tense]$_{\varphi}$ is to produce harmony, in the unwanted mapping /sekalu/ → [seka(lÚ)] it has caused something other than harmony, viz. a stress shift. The solution is to make AGREE a procedural constraint whose only job is to enforce harmony. In order to rule out the unwanted interaction, the constraint AGREE[tense]$_{\varphi}$ must be prevented from causing stress shifts, i.e. must mention the location of metrical structure in its antecedent. I propose the following form of the constraint.

(21) ✦AGREE[tense]$_{\varphi}$
'If V_1 and V_2 are in the same foot, then they have the same value of the [tense]'

The antecedent property mentioned in the constraint (21) is 'being in the same foot': this is the property which the constraint ✦AGREE cannot force to change. By the ICP, the constraint ✦AGREE is violated not only in candidates where there is no agreement within the stress foot, but also by any candidate where foot boundaries are not in their designated state.

In order to determine the designated locus of feet, we take the constraint ✦AGREE out of the ranking. This is shown in the tableau below. The winners for both of the inputs have penultimate stress, and thus a main-stress foot comprising the final two syllables.

(22) Pseudo-Tudanca Spanish

			✦AGREE[tense]$_\varphi$	*[+high, +tense]#	IDENT[tense]	STRESS
/sekalu/		se(kálu)		*!		
		seka(lÚ)			*	*!
	☞	se(kálU)			*	
		se(kÁlU)			*!*	
/sekalo/	☞	se(kálo)				
		seka(lÓ)			*!	*
		se(kálO)			*!	
		se(kÁlO)			*!*	

The vowels *a* and *U* are not in the same foot in [seka(lÚ)], while their t-correspondents *a* and *U* in [se(kálU)] are. The designated state of these two vowels is 'being in the same foot'; any candidate where these two vowels are separated by a foot boundary thus incurs a ✦ violation by the ICP (as would any other candidate whose footing pattern differs from the designated state). Now, using this information, we can assign violation marks to the constraint ✦AGREE. The following tableau illustrates this for the input /sekalu/; the unwanted candidate is fatally penalized.

(23) Designated locus of stress: penultimate

			✦AGREE[tense]$_\varphi$	*[+hi, +tense]#	IDENT[tense]	STRESS
/sekalu/		se(kálu)		*!		
		seka(lÚ)	✦!		*	*
		se(kálU)	*!		*	
	☞	se(kÁlU)			**	

The new violation pattern does not affect the winner for the input with a final non-high vowel, /sekalo/, as the following tableau illustrates. Once again, the pathological candidate [seka(lÓ)] gets an extra ✦ violation mark because its stress foot is different from its designated state; this has no effect on the outcome.

(24) Designated locus of stress: penultimate

			✦AGREE[tense]$_φ$	*[+hi, +tense]#	IDENT[tense]	STRESS
/sekalo/	☞	se(kálo)				
		seka(lÓ)	✦!		*	*
		se(kálO)	*!		*	
		se(kÁlO)			*!*	

The discussion so far focused on the hypothetical language, similar to Tudanca Spanish except that the harmony constraint was able to force a stress shift under the standard OT analysis. Returning now to the actual language, let us make sure that the new violation pattern of the ✦AGREE constraint does not change the outcome. The designated locus of stress for the ✦AGREE constraint for both of the inputs, /sekalu/ and /sekalo/, is penultimate. No tableau should be necessary: because the STRESS constraints are high ranked, no candidate with a stress other than in the penultimate position will be optimal, regardless of whether or not ✦AGREE is taken out of the ranking.

Now the ✦AGREE[tense]$_φ$ assigns extra ✦ violation marks to any candidate with stress other than in the penultimate position, for both of the inputs. Since any such candidate is defeated by the higher-ranked STRESS, the new violation marks will not affect the outcome. The following tableau illustrates.

(25) Tudanca Spanish. DS of stress: penultimate

			STRESS	✦AGREE[tense]$_φ$	*[+high, +tense]#	IDENT[tense]
/sekalu/		se(kálu)			*!	
		seka(lÚ)	*!	✦		*
		se(kálU)		*!		*
	☞	se(kÁlU)				**
/sekalo/	☞	se(kálo)				
		seka(lÓ)	*!	✦		*
		se(kálO)		*!		*
		se(kÁlO)				*!*

To summarize this section: I returned to the Tudanca Spanish data from Chapter 1, and showed that the new violation marks assigned by the ICP prevent the unwanted stress shift candidates from winning. So far, the claim is not typological; I have merely shown that a particular ranking of the constraints does not produce a pathological winner. In the next section I move on to the broader consequence of my proposal, and show that in fact ANY candidate that incurs a ✦ violation mark is a perpetual loser. This will guarantee that the typological generaliztaion about prosody-segmental interactions is accounted for, and will provide a general mechanism for handling procedural generalizations in OT.

3.3.4 Candidates that incur ✦ violations are perpetual losers

In this section I offer a formal demonstration that candidates that incur ✦ violations by clause (13)b of the ICP cannot be optimal. This argument is crucial to the typological claim of my theory. Recall that the candidates that incur ✦ violations are those that involve undesirable input-output mappings that procedural constraints are designed to penalize. Thus, if no such candididate can win, the constraints make a typological prediction about processes, and we have a general way of accounting for generalizations about input-output mappings in OT.

My argument here will be made in three steps, as follows.

- **Step 1**: The 'designated state' is determined uniquely for each constraint.
- **Step 2**: For systems with only one procedural constraint, candidates with ✦ violations are perpetual losers.
- **Step 3**: Adding more procedural constraints to the system does not compromise the typological results.

I begin by taking up a question from Section 3.3.1 that I did not address there: is the designated state uniquely determined by the definition (9)? Let us assume there is a grammar *G* with procedural constraint ✦C, and the phonological category mentioned in the antecedent of the constraint is *P*. Does the grammar *G*', identical to *G* except that ✦C has been removed, uniquely determine the behavior of *P*? In order for my theory to get off the ground, the answer to this question must be affirmative. I argue here that it is.

To show that the grammar *G*' uniquely determines the behavior of the property *P*, it is sufficient to demonstrate that any two candidates that differ in the location or value of *P* have different violation profiles in *G*'. Having a different profile of violations means that the grammar distinguishes these two candidates, and that therefore they cannot both be equally harmonic.

Note first that ✦C, by definition, is a markedness constraint. Grammars *G* and *G*' do not differ in their sets of faithfulness constraints. As long as there are faithfulness constraints referring to *P*, the violation profiles of any candidates differing in *P* can be distinguished by faithfulness alone – and, a fortiori, can be distinguished by *G*'. This means that if *P* is a segmental feature or a piece of prosodic structure to which faithfulness can refer (such as stress), its designated state is uniquely determined by *G*'.

However, there is one important prosodic property which arguably is not subject to faithfulness constraints, namely syllable structure. In these cases, I suggest, the designated state can be determined by markedness alone. I will argue below and in Chapter 4 that the syllable structure markedness constraints ONSET, NOCODA, and *COMPLEX are not procedural. Thus, they are present in any grammar that identifies a designated state of some property. In order for syllable structure to be identified uniquely by such a grammar, it is sufficient that any two candidates that differ in syllable structure incur a different set of violations of these three constraints. This appears to be true, though I have no proof.[17]

[17] It is in fact possible to design pairs of candidates that differ in syllable structure but incur the same set of violations of the three constraints, e.g. [pa.trat.ra] vs. [pat.ra.tra], each incurring one violation of NOCODA and *COMPLEX. However, such candidates have a bounding set (Samek Lodovici & Prince 1999) containing the 'consistent' candidates [pa.tra.tra] and [pat.rat.ra]. This issue needs further exploration.

Now I am ready to proceed to Step 2 of the argument, viz. that, for grammars with only one procedural constraint, candidates incurring ✦ violations of that constraint are perpetual losers. One more definition is necessary here. In the discussion of the definition of 'process' above I referred to the 'designated state' of some phonological property for an input and a constraint as the state of that property in the optimal candidate in a grammar where the constraint in question has been taken out. Now we will need a convenient way to refer to that candidate in the alternative grammar; I will simply call this the DESIGNATED CANDIDATE (henceforth DC).

(26) **THE DESIGNATED CANDIDATE** (Definition)
For a grammar *G*, a constraint *C*, and an input /i/, a candidate is called the DESIGNATED CANDIDATE (DC) of constraint *C* for /i/ if it is the optimal candidate for that input in the grammar *G′* that is identical to *G* except that *C* is removed from the constraint set.

I leave open the question of whether the designated candidate is uniquely determined by *G*'. What is crucial is that, by the argument given above, each designated state is uniquely given by *G*'. This means that if *a* and *b* are two different designated candidates, they do not differ with respect to the property mentioned in the antecedent of the constraint *C*.

Suppose there is a procedural constraint '✦If *P* has property *x*, then *Q* has property *y*' and an input /i/. There is a certain designated state for *P* and some (possibly non-unique) DC for that input. The set of candidates *CAND* can be divided into two non-overlapping subsets. First, there are those candidates that have the property *P* in the designated state. All DCs belong to this set; there are others as well. Let us call this set *D*. Let *N* be the complement set of *D* in *CAND*. The set *N* includes all candidates that do not have *P* in the designated state. All and only members of *N* incur ✦ violations by the procedural constraint. The task is to show that no candidate belonging to *N* can be optimal.

Let us assume that constraints are functions that take a set of candidates and return a non-empty subset of that set (Tesar & Smolensky 2000, Samek-Lodovici & Prince 1999). Let *C* be the function corresponding to our implicational constraint, *H* be the function corresponding to all constraints higher ranked than *C*, and *L* be the function corresponding to all the constraints ranked lower than *C*. (*H* and *L* are compositions of the functions corresponding to individual constraints). If *C* is highest-ranked (lowest-ranked), let *H* (*L*) be the identity function.

By definition, the DCs are the winners in the system which has *H* and *L* but not *C*: the procedural constraint has been removed from the set, and the winner is determined by the remaining constraints. Those winners are given by the expression

(27) $\text{DC} = L(H(\textsc{Cand}))$.

First we apply the function *H* corresponding to the constraints ranked higher than the implicational constraint, and then the function *L* corresponding to all lower-ranked constraints.

The winner of the entire evaluation, with the implicational constraint put back into the ranking, is given by the expression

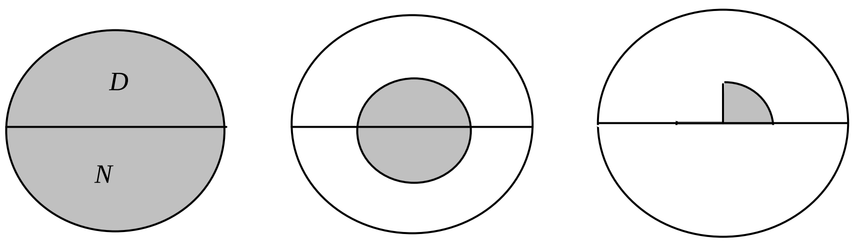
(28) $w = L(C(H(\textsc{Cand})))$.

The application of *H* to the full candidate set yields the set *H*(*Cand*); this is the set of candidates passed down to the lower ranked constraints. This set contains at least one member of *D*, namely the designated candidate *DC*, and possibly other members of *D*. It may also contain some members *N*.

The function *C* applies to this set. Because implicational constraints are, by definition, binary, there are two possibilities:

(29) a. Some remaining candidates incur violations of *C*, some do not, OR
b. All remaining candidates incur an equal number of violations of *C*.

Let us consider the case (29)a first. By the *ICP*, all candidates belonging to *N* incur ✦ violations. By assumption (29)a at least some candidates do not incur any violations of the constraint; therefore, all such non-violating candidates belong to *D*. These are the only candidates that are passed down to the lower-ranked constraints. Therefore, after the application of *C*, all remaining candidates belong to *D* and none to *N*. No matter what the lower-ranked constraints do, no candidate belonging to *N* can win, because no such candidate survives the application of *C*. The following diagrams illustrate this argument for case (29)a. Shading indicates that the candidates are still in the running.



Full candidate set   After application of *H*   After application of *C*

Now consider the case (29)b. Because, by assumption, all candidates in the set passed down to *C* incur an equal number of violations of the implicational constraint, the function *C* is an identity function: it does not eliminate any candidates but simply passes down to the lower-ranked constraints the full set of candidates that are still in the running. Because *C* can be ignored, the winner of the overall evaluation is *L*(*C*(*H*(*Cand*))) = *L*(*H*(*Cand*)). This winner is known: it is the possibly non-unique DC (27). In other words, in the case (29)b the constraint *C* has no effect, so the winner is the same as when the constraint is taken out of the set; by assumption, this winner is the DC, which belongs to the set *D*. Therefore, in the case (29)b as well no member of *N* can be optimal. This completes the demonstration.

In the next section I move to the more complicated situation where more than one implicational constraint is present in the system. There, too, I will argue that no candidate that incurs a ✦ violation may be optimal.

3.3.5 More than one procedural constraint

In the toy examples so far, only one procedural constraint was involved. Clearly, this is an unrealistic simplification. I have to address the question of how two or more such constraints in a grammar can interact, and show that the typological results demonstrated for one implicational constraint are not compromised by adding more constraints to the system. This is the third step of my argument for the perpetual loserdom of candidates that incur ✦ violations.

Because the ICP makes the violation profile of each procedural constraint dependent on the rest of the grammar, it is not immediately clear how the ICP can be applied when there is more than one constraint. In fact, at first blush, the situation appears rather difficult. It is easy to see that, if there are two constraints, applying the evaluation procedure to the two constraints one after another can create unwanted results. Suppose a language has, like English, both an aspiration constraint that attracts [ʰ] to onsets of stressed syllables, and a constraint that flaps posttonic coronal consonants. Assume, as usual in our hypothetical examples, default initial stress covered by the STRESS constraints.

(30) a. ✦ASPIRATE/σ́: If a syllable is stressed, then its onset is aspirated;
b. ✦FLAP: If a syllable is stressed, it is followed by a flap.

Let us see the consequences of computing the designated state of stress separately for each constraint. First we will apply the procedure to the constraint ✦ASPIRATE, and then to ✦FLAP. Suppose the language has the ranking IDENT[son] ⟩⟩ ✦ASPIRATE ⟩⟩ DEP-*h* ⟩⟩ ✦FLAP ⟩⟩ STRESS. Consider the input /pitaɾa/. To compute the designated state of stress for the constraint ✦ASPIRATE, we take it out of the system and find the stress in the optimal candidate. Because the designated state of stress for ✦FLAP has not been found yet, we have to treat it like a normal OT constraint. Given the assumed ranking, the optimal candidate is necessarily *pitáɾa*, with stress attracted by flapping. The computation of the designated state of stress for ✦ASPIRATE is illustrated below.

(31)

		IDENT[son]	✦ASPIRATE	DEP-*h*	✦FLAP	STRESS
/pitaɾa/	pítaɾa				*!	
	☞ pitáɾa					*
	píɾaɾa	*!				
	pʰítaɾa			*!	*	
	pitʰáɾa			*!		*
	pʰíɾaɾa	*!		*		

The optimal candidate is *pitáɾa*, with a non-default stress that has been attracted to a flap. This is exactly the type of pattern that the theory is designed to rule out! The problematic stress in *pitáɾa* now becomes the designated state of stress for ✦ASPIRATE, meaning that it would ✦-penalize all candidates where stress is on any other syllable. This is illustrated by the tableau below.

(32) DS of stress for ✦ASPIRATE: second syllable (*ta*)

		IDENT[son]	✦ASPIRATE	DEP-*h*	✦FLAP	STRESS
/pitaɾa/	pítaɾa		✦!		*	
	pitáɾa		*!			*
	píɾaɾa	*!	✦			
	pʰítaɾa		✦!	*	*	
	☠ pitʰáɾa			*		*
	pʰíɾaɾa	*!	✦	*		

This result is fatal for the typological predictions: the designated state of stress, as well as the stress in the winner, is now in a place where it should not be. The candidates with initial stress – the ones that should be optimal – end up with fatal ✦-violations of the ✦ASPIRATE constraint.

Moving on to the designated candidate for the constraint ✦FLAP, taking this constraint out of the ranking does not change the winner as illustrated by the tableau below (33).

(33)

		IDENT[son]	✦ASPIRATE	DEP-*h*	✦FLAP	STRESS
/pitaɾa/	pítaɾa		✦!		*	
	pitáɾa		*!			*
	píɾaɾa	*!	✦			
	p^{h}ítaɾa		✦!	*	*	
	☠ pit^{h}áɾa			*		*
	p^{h}íɾaɾa	*!	✦	*		

The designated location of stress for ✦FLAP is also on the second syllable *ta*. This means that any candidate with a stress on a different syllable will be ✦-penalized by ✦FLAP as well. In this particular case, as shown by the tableau below, the extra violations make no difference: the output is still the pathological candidate *pit^{h}áɾa*.

(34) DS of stress for ✦ASPIRATE: second syllable (*ta*)
DS of stress for ✦FLAP: second syllable (*ta*)

		IDENT[son]	✦ASPIRATE	DEP-*h*	✦FLAP	STRESS
/pitaɾa/	pítaɾa		✦!		✦	
	pitáɾa		*!			*
	píɾaɾa	*!	✦		✦	
	p^{h}ítaɾa		✦!	*!	✦	
	☠ pit^{h}áɾa			*!		*
	p^{h}íɾaɾa	*!	✦	*	✦	

This example shows that as soon as more than one constraint is present, applying the procedure to separate constraints overturns the result that antecedent properties cannot be modified.

This very serious problem arises because both of the constraints ✦ASPIRATE and ✦FLAP mention stress in their antecedent. Thus, if we try to determine the designated state of stress of one of these constraints, the other one can 'interfere' and produce an unwanted effect; this is what gave the bad result in the tableaux above.

Crucially, such interference only occurs if the two procedural constraints share their antecedent property. In the example under discussion, it would be sufficient to compute the designated state of stress ONCE for both the constraints ✦ASPIRATE and ✦FLAP, by taking them both out of the ranking simultaneously and finding the location of stress in the optimal candidate of the resulting grammar.

More generally, the set of procedural constraints can be partitioned into non-overlapping ANTECEDENT CLASSES based on what property is mentioned in the first part of the constraint: there is a STRESS class of constraints that have the form 'If a syllable is stressed, then *Q*', a TONE class, a WEIGHT class, a NASALITY class, etc. All constraints mentioning property *P* in their antecedent belong to the antecedent class *P*. The designated state of each property *P* can then be computed only once for each antecedent class, as given by the following definition, modified from (9).

(35) **THE DESIGNATED STATE** (Definition; modified)
Let there be a grammar *G*, an **antecedent class of constraints** *C*, an input /i/and its output [o] in *G*, and a phonological object *p* in [o]. Then the DESIGNATED STATE of *p* for /i/ and *C* in *G* is *p*'s correspondent in the optimal candidate in the grammar *G*' that is identical to *G* except that **all of the constraints** in *C* have been removed.

The intuition behind this modified definition of designated state is that we first determine the default behavior of some property (e.g. stress), and then apply all processes which are sensitive to that property.

Now the unwanted predictions do not arise. ✦ASPIRATE and ✦FLAP both have 'stressed syllable' as their antecedent property, and thus belong to the same antecedent class. The designated state of stress is thus computed by taking both constraints out of the ranking at the same time, as shown in the tableau below. The STRESS constraints ensure that the designated state of stress is on the initial syllable.

(36)

			IDENT[son]	✦ASPIRATE	DEP-*h*	✦FLAP	STRESS
/pitaɾa/	☞	pítaɾa					
		pitáɾa					*!
		píɾaɾa	*!				
		pʰítaɾa			*!		
		pitʰáɾa			*!		*
		pʰíɾaɾa	*!		*		

Now, any candidate with stress on a syllable other than the initial one would receive a fatal ✦ violation, and thus the typological predictions are preserved. The final evaluation is illustrated below.

(37) DS of stress: initial syllable

			IDENT[son]	✦ASPIRATE	DEP-*h*	✦FLAP	STRESS
/pitaɾa/		pítaɾa		*!		*	
		pitáɾa		✦!		✦	*
		píɾaɾa	*!	*			
	☞	pʰítaɾa			*	*	
		pitʰáɾa		✦!	*	✦	*
		pʰíɾaɾa	*!		*		

The remaining question is how to deal with more than one antecedent class in the same system. Suppose there are *n* antecedent classes of constraints; call them C_1, C_2, …, C_n. For each input, each of these antecedent classes is associated with a designated state of its antecedent property. The question is, how to determine that property? For each class, the designated state must be determined by taking all of the constraints in that class out of the system, and letting the remaining grammar decide the behavior of the relevant property. However, once this operation is performed for some class C_i, the violation pattern of all of the constraints in C_i changes. Therefore, the order in which the designated states for the different antecedent classes is determined can potentially affect the outcome. Furthermore, and more importantly, it is not immediately clear that the typological predictions of the theory remain with the addition of more than one antecedent class to the grammar.

I argue here that the typological predictions stand: no matter what order the designated state is computed for the several antecedent classes of constraints, no candidate where the antecedent property is modified can win. However, an unsettling result is that the order in which the operation is performed does in fact affect the outcome. Thus, while the system is typologically well-behaved, it is indeterminate: the same grammar can produce several outputs. It will then be necessary to fix an order for computing the designated states arbitrarily, by stipulation. This area of the theory remains for future research: it remains to be seen what empirical predictions are made by this arbitrary choice of the order of computation of designated states.

Let me now go through the first part of the argument, viz. that the typological predictions are not compromised by the presence of more than one antecedent class of constraints. Consider some antecedent class, C_i, whose antecedent property is P_i. The constraints express certain procedural generalizations about P_i, by putting this property in their antecedent and thus penalizing certain input-output mappings of it, by the ICP. By assumption, ALL procedural generalizations concerning P_i are expressed by the constraints in C_i; that is how the constraint system is designed. To take a specific example, if P_i is stress, then all generalizations about what stress cannot be attracted to are expressed by the constraints in its antecedent class, C_i. It contains constraints like '✦If stress, then aspiration', '✦If stress, then flapping', etc. – constraints that make it impossible to assign stress based on certain segmental properties. No constraint outside of C_i has the power to limit the set of properties to which stress can be sensitive, either because those constraints are not procedural, or because they belong to a different antecedent class.

Once again, by assumption, no grammar with all of the constraints in C_i taken out can produce a typologically undesirable stress pattern (e.g. aspiration-driven stress), because all constraints that can cause such effects are members of C_i. It immediately follows that the designated state of the property P_i is typologically well-behaved, and thus the ICP cannot cause an unwanted candidate to win. This completes the argument that the typological predictions of the theory are not compromised by the presence of more than one antecedent class of constraints.

However, there is an indeterminacy to the theory. If there is more than one antecedent class of constraints, each of which is associated with a designated state of various properties for each input, in what order should the designated state be computed? A simple example shows that the choice of the order can affect the outcome. Suppose there are two procedural constraints in the grammar, shown below. One is the familiar aspiration constraint, the other is the sonority-driven stress constraint that attracts stress to syllables with the nucleus *a*.[18]

(38) a. ✦ASPIRATE: If a syllable is stressed, then the onset is aspirated.
b. ✦STRESS-*a*: If the nucleus of a syllable is *a*, then the syllable is stressed.

Consider the input /pitaku/ in a grammar with default initial stress. Then, depending on the order in which the designated state-computing procedure is applied to this input, the winner can be either *pit*^{h}*áku* or *p*^{h}*ítaku* (I will spare the reader the tableaux that show this). Neither of these presents a typologically undesirable pattern (stress attraction to aspiration), but the grammar is indeterminate.

It is therefore necessary to fix the order in which the designated states are computed for several antecedent classes. I am not aware of any independent motivation for such an order. In the remainder of this dissertation, I will apply the procedure in the order in which the constraints are ranked, starting from the highest-ranked and proceeding downward. Once again, making this choice into an empirical one remains for future research.

---

[18] The exact form of this second constraint does not matter for now; I will come back to sonority-driven stress below. What is crucial here is that 'being *a*' is in the antecedent, while the location of stress is in the consequent.

3.4 Bidirectional interactions

The discussion so far was devoted to those cases where stress interacts asymmetrically with some segmental property. I have shown how my proposal to introduce a new type of constraint into the theory allows generalizations about the stress assignment process to be handled in OT. I suggested that by linking the asymmetry of the implicational statement of a constraint to the direction of interaction between the two properties mentioned in it, undesirable processes like aspiration-driven stress can be ruled out.

As explained in detail in Chapter 1, the typology of interaction between stress and segments contains not only asymmetrical interactions such as those analyzed in this chapter so far, but also the three broad classes of cases where stress interacts bidirectionally with some other property. The three properties that are able to influence stress are quantity, tone, and vowel sonority. In this section I explore the role of implicational constraints in accounting for these symmetrical interactions.

There are two general approaches one could take to these cases. One would be to argue that whenever stress interacts freely with some property, the constraints responsible for it are not implicational, but standard OT constraints expressing output-based generalizations that do not invoke the special interpretation mechanism proposed here. This approach is, of course, assumed in standard OT, where constraints like stress-to-weight (SWP) and weight-to-stress, stress-to-tone and tone-to-stress (de Lacy 2003) are interpreted as standard output-oriented markedness constraints.

The second approach to the bidirectional interaction of stress would be to employ, for each of the features, two implicational constraints with opposite directionality. If the procedural constraints 'If stressed, then heavy' (✦SWP), and 'If heavy, then stressed' (✦WSP) exist in the system, then both directions of interaction would be accounted for: each of the two constraints would be responsible for one of the directions of interaction. The former constraint would produce stress-driven weight, and the latter constraint would account for weight-driven stress. Here I argue that the second option, with two opposing implicational constraints is the correct view.

The empirical difference between the two approaches lies in the types of processes BESIDES stress-segmental interactions that the constraints are predicted to force. If the theory has only two constraints, ✦WSP and ✦SWP, then the only two kinds of repairs for stress-weight misalignment that are predicted to exist are stress attraction to heavy syllables, and lengthening/gemination in stressed syllables to make them heavy. The two constraints cannot cause any other process, under the theory proposed here. Because the effect of the ICP is to prevent constraints from modifying the property mentioned in its antecedent, ✦SWP cannot cause any repair that involves a stress shift, while ✦WSP cannot cause any repair that modifies the original heaviness of a syllable.

On the other hand, if stress-weight interactions are handled by the standard OT pair of constraint like SWP and WSP, no such predictions about possible repairs are made. My argument here is that the more restrictive predictions of the implicational version of ✦SWP and ✦WSP are empirically correct.

An important set of cases where SWP and ✦SWP make different predictions concerns vowel syncope. I will touch briefly on these cases here; they will be the topic of a more detailed investigation in Chapter 4. The standard OT constraint SWP can cause the deletion of a stressed vowel just in case the stressed syllable in the outcome is heavy, while the implicational version ✦SWP cannot do so. Let me illustrate this point with a hypothetical example. Suppose the input has three light syllables with stress on the second syllable, /$CV_1C\acute{V}_2CV_3$/. A high-ranking SWP constraint can potentially force the deletion of a the stressed vowel $\acute{V}_2$ just in case all other potential repairs creating a heavy stressed syllable in the output are ruled out by other constraints. The tableau below illustrates such a case: the fully faithful candidate ($CV_1C\acute{V}_2$)CV_3 is violates SWP; lengthening of the stressed vowel ($CV_1C\acute{V}_2$:)CV_3 violates DEP-μ; and the deletion of the final syllable to give ($CV_1C\acute{V}_2C$) is precluded by the high-ranking NONFIN constraint. This leaves the deletion of the stressed vowel $\acute{V}_2$ with resyllabification of its onset as the coda of the preceding syllable as the only viable option that would satisfy SWP. As a result, the constraint SWP causes the deletion of a stressed vowel and a concomitant stress shift.

(39) $/CV_1C\acute{V}_2CV_3/$

	SWP	DEP-μ	NONFIN	PARSE-σ	MAX-V
$(CV_1C\acute{V}_2)CV_3$	*!			*	
$(CV_1C\acute{V}_2{:})CV_3$		*!		*	
$(CV_1C\acute{V}_2C)$			*!		*
☞ $(C\acute{V}_1C)CV_3$				*	*

On the other hand, the implicational version ✦SWP cannot produce such an outcome, because this constraint has stress mentioned in its antecedent, and therefore any candidate with a stress shift is a perpetual loser. This is illustrated in the tableau below.

(40) $/CV_1C\acute{V}_2CV_3/$
Designated location of stress: V_2

	✦SWP	DEP-μ	NONFIN	PARSE-σ	MAX-V
$(CV_1C\acute{V}_2)CV_3$	*!			*	
$(CV_1C\acute{V}_2{:})CV_3$		*!		*	
☞ $(CV_1C\acute{V}_2C)$			*!		*
$(C\acute{V}_1C)CV_3$	✦!			*	*

Although several analyses of syncope processes that make use of the power of SWP to produce stressed vowel deletion have been proposed, I will argue in Chapter 4 that such analyses are incorrect, and better accounts are available for the relevant cases. As I will show, this inability of SWP to force stressed vowel deletion is part of a more general typological generalization that stressed vowels cannot be deleted through metrically-driven syncope. If that claim is correct, then it follows that the more restrictive constraint ✦SWP is to be preferred over the less restrictive version SWP, because the typology of repairs predicted by ✦SWP more tightly fits the observed typology.

3.5 Grounding the constraints

The mechanism of constraint interpretation introduced in this chapter allows constraints to refer directly to processes, and to penalize those candidates that involve undesirable processes. The proposal is intended to replace SOME of the current OT markedness constraints with constraints of the new type, but, crucially, output-oriented constraints are still present in the theory.

The rationale for allowing these two types of constraints in the system is that phonological generalizations are of two types. Some are surface-based, and require the standard OT constraints to handle them. Other generalizations are procedural, and require the new ✦-constraints to directly express them. This richness of the formal machinery in the proposed theory relative to standard OT is necessary because of the richer set of phonological generalizations than has been standardly assumed.

The next question then is, is it possible to predict generalizations are output-based and which generalizations are procedural? From the theoretical standpoint, is it possible to tell which constraints should be subject to the standard OT mechanism of constraint evaluation, and which constraints should be subject to the ICP? In other words, we have the problem of GROUNDING the constraints in some extratheoretical reality.

In this section I discuss some psycholinguistic evidence in favor of the observed proposed directionality of the stress constraints. However, I will argue that the best way of determining the nature of a given constraint is the typology of the interactions between the properties that it mentions.

3.5.1 Typological grounding

Rather than grounding the directionality of the asymmetrical procedural constraints in some extralinguistic factor, I rely on the more standard technique used by OT practitioners: grounding constraints in the observed typologies, and selecting, from the set of possibilities provided by the formalism, those constraints that account best for the crosslinguistic distribution of a particular feature. To use a familiar example, there is nothing a priori which leads to the two syllable structure constraints being ONSET and NOCODA: as far as the formal machinery is concerned, these constraints might have been the opposite, CODA and NOONSET. However, the typology of syllable structure points to the correct formulation of the constraints (Prince and Smolensky 1993).

The typological predictions on the interaction of phonological properties depend on the correct allocation of phonological categories among the two parts of the constraint, the causer (the antecedent) and the causee (the consequent). The form of each constraint, as stated above, must be decided on a case by case basis: whatever the direction of interaction the typology indicates must be reflected in the statement of the constraint. This lack of external, independent grounding is perhaps the weakest aspect of the current proposal.

As mentioned elsewhere in this chapter, there are yet more degrees of freedom than the choice of directionality for the procedural constraints. Whether or not a given constraint is procedural in the first place, and thus controlled by the violation assignment mechanism introduced here, is also subject to the analyst's choice. Once again, this choice must be made with reference to the observed typology.

The main objection of this dissertation to canonical OT is that its hypothesis that all phonological generalizations are in the output is too radical. I do not intend to replace this radical hypothesis with its opposite, however. Nothing I have written here should be interpreted as offering a counterargument to the observation, established since Kisseberth 1970, that phonologies often conspire to produce a certain output pattern. This fundamental claim stands. And yet, the systematic gaps in the typologies of interaction of some properties show that SOMETIMES phonologies respond to pressures other than purely surface-driven. These two sources of phonological generalizations, surface pressures and input-output mappings, lead, in my theory, to two types of constraints. Procedural constraints reflect the latter kind of generalization, and standard

OT constraints reflect the former. It is the analyst's job to assign each markedness constraint to one class or the other.

The argument in this chapter is that there are two kinds of constraints – the standard OT constraints that refer to output structures and prohibit marked configurations, and those that refer to processes and prohibit certain input-output mappings. The former type of constraints encode surface-based generalizations, while the latter type refer to procedural generalizations. For a given constraint, the observed typology of interactions serves as a clue to assigning the constraint to one type of the other. If the relevant phonological categories interact in a way that is most insightfully described in terms of a procedural generalization, then the constraint must be stated implicationally and is subject to the ICP. If, on the other hand, the generalization is best stated in terms of outputs, then it must be handled by a standard OT constraint. I will use this reasoning in Chapter 4 to argue that syllable structure constraints like ONSET and NOCODA, as well as sonority-sequencing and syllable contact constraints, must remain of the standard output-oriented type (section 4.1.4).

3.5.2 Psycholinguistic evidence

In this section I survey the psycholinguistic evidence that is relevant to the interaction of prosody and segments. If the directionality of the statement of constraints is to be grounded in such extragrammatical facts, the mode of explanation here would be analogous to Hayes' (1995) proposal to link the asymmetrical foot typology (the Iambic-Trochaic Law) to the effects of rhythm psychology: "I posit an EXTRASYSTEMATIC motivation, in a law of rhythm, for internal formal principles of the linguistic system" (Hayes 1995: 81). Cf. also the programmatic statement in Anderson 1981: "[W]hile linguistic constraints proper need not mirror the restrictions of other cognitive structures, nonetheless these other structures, insofar as they are involved in the development of linguistic knowledge, can be expected to have their own consequences for actual grammar… [T]he character of the linguistic system depends on the interaction of particular substantive considerations NOT specific to Language with an irreducible component which is" (1981: 536).

While the evidence I discuss below suggests that the asymmetrical statements of constraints are not arbitrary and have some basis in the human perception of stress and rhythm, I will remain cautious in directly grounding the claims about the formal structure of constraints in evidence about processing. There are two reasons to remain cautious in this respect: first, the evidence on processing and perception of stress is not complete, most importantly lacking in broad cross-linguistic data that would allow one to see similarities and differences between typologically distinct stress systems. Secondly, and more importantly, the theory I am proposing here is a theory of knowledge of language, which is in principle independent of processing and performance, and thus data about how humans process stress and segments cannot be directly used to argue for one or another formalism. With these caveats in mind, I turn to discussing the relevant experimental studies.

On the whole, the evidence indicates a psycholinguistic difference between those features with which stress is typologically observed to interact bidirectionally and those which stress can only influence in a unilateral fashion. Two sets of arguments from the psycholinguistic literature suggest that the prosodic component has a special status in phonology vis-à-vis the segmental component. First, patterns of speech errors and tip-of-the-tongue phenomena differ in the prosodic and segmental domains in a way that points to the logically prior status of stress with respect to segments. Second, the special status of quantity, intensity, and tone as correlates of stress is confirmed by the literature on the psychology of rhythm, which shows these same features to play a role in perception and production of non-linguistic prosody.

In this section I will review the available evidence that bears on the role of prosody in grammar. However, while the arguments from this external domain are suggestive, and at least do not contradict a theory of prosody-segmental interactions that gives special status to prosodic structure with respect to segmental features, I will not argue for a direct grounding of the asymmetrical prosody-segmental constraints in psycholinguistic data. There are two reasons to be cautious in this respect. First, the claims of this section must be taken with a grain of salt because most of the studies presently available have been done on English, so we know little about the crosslinguistic differences in processing and acquisition of prosodic structure, such as the difference between pitch accent and dynamic stress languages, let alone the psycholinguistic status of covert prosodic structure. The second reason why such external evidence does entail a particular view of constraint interaction in a direct way is that the theory I am pursuing is a theory of the knowledge of language, not of processing.

Let me begin by discussing the data on speech errors suggesting that prosody is processed prior to the segmental details. One source of information are tip-of-the-tongue (TOT) phenomena, which involve a temporary inability to produce a phonological output for a word whose semantics has already been accessed (Levelt 1989: 320). It TOT states, speakers have the intuitive feeling that they know the word they are trying to produce, but cannot remember how it sounds exactly. Because priming with phonologically similar but semantically unrelated words can cause a TOT state, it is generally accepted that such states are caused by an unsuccessful search for a phonological form of a given lemma, not for the lemma itself. Speakers in TOT states are typically aware of the prosodic structure of the form whose segments they cannot access, viz. the number of syllables and the stress pattern (60-80%). The initial segment of the target word is also available to TOT state speakers in 60-70% of the time. What is missing, however, are the details of the word's segmentation: the distribution of consonantal and vocalic features in the form. This clearly suggests that prosodic information is accessed earlier than segmental structure in word generation. TOT states occur when a speaker fails to proceed from the prosodic to the segmental stage.

Speech error data also suggest a separate behavior of prosody as opposed to segments. Particularly relevant is the study of malapropisms – word selection errors that are not due to substitutions, anticipations, and similar phonological problems, and are not semantic errors. In a survey of English malapropisms, Fay and Cutler (1977) showed that the mistakenly selected word in 98% of the cases has the identical stress pattern to the target word. This rate of agreement is higher than the 82% agreement in stress patterns found in errors classified as semantic (Fay and Cutler 1977: 508). The malapropisms typically have the same syllable count and the same grammatical category as the target word. This suggests that, at the point at which the error occurs, morphological and prosodic information, but not segmental structure, has been accessed. The locus of error in malapropisms is the same as in TOT states: it occurs between the prosodic and the segmental stages in processing. The difference between the two types of errors is that in the latter case the speaker fails to access any form, while in the former case he proceeds in the wrong direction. This confirms what TOT phenomena also suggest, that prosody is accessed relatively early in speech production.

Equally informative are error types and TOT states that are not reported to occur. While it is common for a speaker in a TOT state to have accessed the prosodic structure of a word while still searching for segmental information, the reverse type of TOT states has not been reported, i.e. having accessed and being able to produce the segmental content but searching for its correct prosodic parse, or its right stress pattern. Likewise, while malapropisms involve segmental mistakes with the correct prosodic structure, the reverse type of error appears to be unreported. Such a prosodic malapropism would contain the correct segmental structure with a prosodic aberration that cannot be attributed to substitution or anticipation errors.

To be sure, stress errors do occur in speech (Cutler 1980). However, they are almost entirely errors of a single type: an erroneous stress pattern is 'borrowed' from a morphologically related word. For example, Cutler reports the error *certífication*, due to the stress pattern of the related word *certíficate*. This type of error is clearly different from segmental malapropisms, where the incorrect segmentism is borrowed not from morphologically related words, but from PROSODICALLY similar unrelated words. Cutler's survey shows that in stress errors the prosodic pattern is not borrowed from segmentally similar but morphologically unrelated words. Once again, the error data indicate that prosody is accessed before segments.

These facts have lead researchers like Levelt to posit that prosodic information is accessed early, perhaps concurrently with morphological structure, and prior to most segmental information.

While the error studies have concentrated on English, there is some crosslinguistic data available on the comprehension side. The summary of previous research in Cutler and Van Donselaar (2002) indicates that stress is an important factor in word recognition in Dutch but not in English. Members of stress-based minimal pairs in Dutch, like *vornáam* 'first name' and *vórnaam* 'respectable' do not prime each other, while the members of analogous English pairs like *forbéar* and *fórbear* do. The relatively greater importance of stress in word recognition in Dutch vs. English was linked by these researchers to the relatively greater amount of vowel reduction in English. Because the functional load of stress per se is not great in a language with significant vowel reduction – in English, there are not many pairs like *forbéar* and *fórbear*, where the stress difference does not correlate with some vowel quality difference – stress is simply not a useful tool in word recognition in English. Instead, Cutler and Van Donselaar argue, English speakers, unlike Dutch speakers, act more efficiently by focusing on segmental cues in word recognition.

Note, however, that this result does not necessarily show that PROSODIC STRUCTURE itself is not relevant to word recognition. Another way to interpret the findings is simply that intensity is not a clue to prosodic structure used by English speakers; instead, they rely on the more robust segmental cues such as vowel reduction as a source of information about prosody, but use prosody to distinguish words. It would follow that pairs where the clues are abundant – those with vowel reduction – would be more easily distinguishable, while the prosody of pairs like *fórbear* and *forbéar* would be less accessible.

Another relevant line of research involves the relationship between the type of stress system found in a language, the presence of lexical exceptions in stress placement, and the ability of speakers to 'hear' stress (Peperkamp and Dupoux 2002, Peperkamp 2004). These researchers observe that speakers of some languages, such as French, are "stressdeaf" – are not able to identify explicitly where stress goes in a form, and are poor at distinguishing stress contrasts when presented with stimuli from a language that has contrastive stress. On the other hand, speakers of Spanish are relatively more aware of the location of stress and are better at distinguishing members of accentual minimal pairs. Peperkamp and Dupoux link this difference in behavior between the two languages to how predictable the location of stress is from utterance boundaries. A child learning a language like French can infer the location of stress solely with reference to utterance boundaries – no information on word or morpheme boundaries is necessary to arrive at the correct stress rule. Stress then can be acquired 'prelexically' – that is, before actual words and independently of them – and will not be stored in the mental lexicon for each lexical entry. Conversely, in a language like Spanish, where morphological and lexical information figures in the stress rule, that rule cannot be acquired solely from the position of stress relative to word boundaries, and thus cannot be available to the child prelexically. As a consequence of this relatively late acquisition of the stress rule, Dupoux and Peperkamp hypothesize, in Spanish-like languages accentual information is redundantly recorded in lexical entries. This accounts for the relatively greater awareness of stress by Spanish speakers than by French speakers. Peperkamp (2004) further links this difference to the existence of lexical stress exceptions in Spanish-type languages vs. French-type languages.

This work suggests that there may be a crosslinguistic difference in processing of stress. Because English, with its numerous stress exceptions, morphologically based accentual generalizations, and so forth, clearly belongs to the Spanish type rather than the French type, work done exclusively on English speech error data might present a skewed picture of the processing of prosodic structure. Fay and Cutler's (1977: 511) suggestion that stress patterns in English serve as an important organizational principle of the lexicon that aids both production and comprehension may not hold for languages of a prosodic type differing from English. For this reason the suggestive psycholinguistic data discussed in this section must be taken with a grain of salt.

So far I have discussed the psycholinguistic differences between prosodic and segmental structure in a given language with an assumption that it is established a priori which features count as segmental and which features count as prosodic. The next question, then, is: given that there are these two categories of phonological properties, why do intensity, quantity, and tone fall on the prosodic side of the divide, while the segmental features fall on the other side? To address this question I turn to the literature on the psychology of rhythm.

The function of rhythmic stress is to organize the speech string into manageable units for processing and storage, and that is why the nature of stress systems can be made clear by looking at the psychology of rhythm (cf. Hayes 1995). I suggest here that the special status of intensity, duration, and tone in stress systems, and ultimately the availability of bidirectional constraint sets linking these properties with abstract prominence, is due to the role of these features in the psychology of rhythm.

Duration, intensity, and pitch stand out among the linguistically relevant characteristics of a speech signal in that these properties are all and only that can be meaningfully applied to non-linguistic signals as well.[19] One can talk about, say, a musical note having duration and pitch and loudness, but not nasality or spread glottis or aspiration. What makes prosodic properties special in language is that these properties are coopted for linguistically relevant use from a non-linguistic domain. In language, prosodic features are also the only ones that express what can be called "paralinguistic meaning": the expression of emotional states such as surprise, disgust, excitement, and so forth, is done via exaggerating the prosodic features of the speech signal, or overlaying it with new prosodic content. One would not expect to find a language where, say, extreme surprise is expressed by nasality or voicing rather than by intensity or shifts in pitch register. And finally, the prosodic features are the only ones that can be meaningfully applied to non-linguistic vocalizations that are part of the speech signal, i.e. expressions like "arghh, hmm, phew", whistles, signs, and so forth.

[19] This insightful observation is due to Larry Hyman, p.c.

All such items must necessarily have some duration, some loudness, and, at least if voice is involved, some pitch – but not necessarily any segmental features.

In sum, prosodic properties are all and only properties that exist independently of language, but can be used by phonological systems for linguistically relevant ends.

This extralinguistic status of prosody is manifested also in its role in the psychology of rhythm. There is extensive evidence showing that humans have an innate tendency to organize sequences of like stimuli into small, regularly recurring, periodic constituents. Prosodic properties have a special role with regards to the perception and production of rhythm.

The basic human propensity for rhythmic organization is demonstrated by experiments showing that a succession of identical sounds is perceived as a rhythmic pattern, at a rate of between 0.5 and 5 Hz (Bell 1977, Fraisse and Oléron 1954, Allen 1975). The length of the rhythmic units is determined by the rate of the stimulus presentation: the higher the rate, the more units per group.

The two prosodic properties of intensity and duration have been shown to correlate with rhythmic strength in a way that has been subject to much discussion in the metrical stress literature. The iambic-trochaic law (Hayes 1995) expresses this asymmetry in perception: loud sound first, long sound last. A succession of sounds with alternating intensity is perceived as a trochaic sequence (Fraisse and Oléron 1954), though the effect diminishes at slow rates, below about 3.5 Hz (Bell 1977). If the successive sounds are distinguished by length rather than loudness, they are perceived as an iambic sequence, provided the difference in length is by a factor greater than about 1.5. Pitch seems to have a weak tendency to cause left prominence (Bell 1977).

A further relevant result of Bell 1977 is the relative language independence of the rhythmic effect. He found that Polish speakers are not influenced by the regular penultimate default stress of their language and do not perceive the penultimate sound in a series of identical sounds as more prominent.

This experimental work shows that the prosodic characteristics of at least intensity and duration have what we may call 'inherent salience': they are associated with rhythmically prominent positions. While we may remain agnostic as to the innateness of the human psychology of rhythm, it is clear from the experimental work cited above that, first, this ability and tendency to group stimuli into periodic units is extralinguistic, and second, that duration and intensity are inherently associated with rhythmic prominence. Tone may also have such a relationship with rhythm, although the details are less clear.

Once again, prosodic features in language are those that also have extralinguistic relevance – in this case, relevance to the psychology of rhythm – and have been coopted by languages to signal linguistically relevant rhythm, i.e. stress. This relationship between the extralinguistic and phonological manifestations of prosody becomes especially clear if we consider a typological asymmetry so obvious that hardly anyone ever bothers to note it. As I have mentioned many times above, stressed syllables tend to be both longer and louder than unstressed syllables. However, nothing in principle prevents the reverse use of prosodic cues: why are there no languages where stressed syllables are not louder, but LESS LOUD than unstressed syllables? Or languages where stressed syllables are not longer but SHORTER than unstressed ones?[20] Naturally, such a hypothetical language would appear absurd to most phonologists, but it is not as unreasonable as it seems. Stressed syllables in a language where they are cued by shortness rather than length would be just as easily distinguishable from unstressed stressed syllables as in a 'normal' language, so the absence of such languages from the typology has nothing to do with perceptibility. Instead, such counter-natural way of cuing prosodic prominence would go against the inherent association between rhythmic strength and prosodic features, and would arguably be unlearnable by a human being.

To sum up the discussion so far, the evidence from the psychology of rhythm suggests that prosodic features have a special status in language in two ways: first, they are the features that have non-linguistic meaning, and the only properties that can be

[20] There is no circularity here. Some readers may ask if in such a language we would simply call the shorter and the less intense syllables unstressed. The situation I claim here is unattested would be distinguished by having only ONE shorter or less intense syllable per word, on analogy to systems that have only one louder or longer syllable per word.

predicated of non-linguistic stimuli; and second, these prosodic features (with some complications for tone) have an inherent association with the human sense of rhythm. Evidently, these two facts are related: it is no surprise because rhythm psychology is an extralinguistic ability, and humans have a sense of rhythm in stimuli devoid of any linguistic content, that the features used to signal that rhythm are also not specifically linguistic.

Crucially, none of what I said so far with regards to the three prosodic features applies to any of the other potential effects of stress – i.e. any of the segmental effects having to do with consonantal features such as voicing, sonorancy, etc. These features are language-specific, not relevant to extralinguistic signals such as music, and have no inherent association of any kind with rhythm. This fact may suggest that the cause-effect relationship between prosodic features and linguistic prosodic constituent structure is bidirectional, but the effect of stress on consonants is not reciprocated by the like effect of consonants on prosody. As I stressed above, however, despite the suggestive nature of the psycholinguistic evidence, formal phonology must remain cautious about directly using that evidence in support of a particular formalism.

3.6 Sonority-driven stress and prominence

3.6.0 Introduction

Stress systems are often sensitive to vowel quality, more specifically, to the position of a vowel on the sonority scale (Kenstowicz 1994, de Lacy 2003). The standardly accepted sonority scale below corresponds to vowel height, with the low vowels being more sonorous that non-low vowels, and central vowels being the least sonorous of all. Sonority-driven stress systems typically have a preference for assigning stress to syllables more sonorous vowels.

(41) ə, ɨ 〈〈 i, u 〈〈 e, o 〈〈 a

De Lacy's (2003) theory of sonority-driven stress analyzes the interactions between stress placement and vowel quality with sets of markedness constraints in a stringent relationship, as shown below.

(42) a. $^*Hd_{Ft}/\{ə\}$
$^*Hd_{Ft}/\{ə, i/u\}$
$^*Hd_{Ft}/\{ə, i/u, e/o\}$
$^*Hd_{Ft}/\{ə, i/u, e/o, a\}$

b. $^*Non\text{-}Hd_{Ft}\{a\}$
$^*Non\text{-}Hd_{Ft}\{a, e/o\}$
$^*Non\text{-}Hd_{Ft}\{a, e/o, i/u\}$
$^*Non\text{-}Hd_{Ft}\{a, e/o, i/u, ə\}$

The constraints in (42)a penalize vowels in prominent positions starting at the low-sonority end of the hierarchy, while the constraints in (42)b penalize vowels in non-prominent positions starting at the high end of the hierarchy. In case these constraints are allowed to outrank stress placement constraints, the predicted effect is to force stress onto more sonorous vowels and repel it from less sonorous vowels. Conversely, if the constraints in (42) outrank vowel quality faithfulness constraints, the predicted effect is for vowels in prominent positions to move up on the sonority scale and for vowels in non-prominent positions to move down the scale.

As I mentioned in Chapter 2, the existence of sonority-driven stress is not easily compatible with de Lacy's the fixed ranking approach, because the constraints in (42) are the only constraints that mention segmental information that are allowed to outrank stress placement constraints. Sonority-driven stress can be analyzed using the proposal on procedural constraints in this chapter as long as the direction of the implicational statement of the constraint mentions stress in the consequent, not the antecedent. If the constraint says 'If a syllable is stressed, then its nucleus is an [a]', the stress cannot shift to satisfy it. If, on the other hand, the constraint is 'If a syllable's nucleus is [a], then that syllable is stressed', then stress can be attracted to low vowels.

While the theory proposed above can deal with the general case, there are some types of sonority-driven stress, uncovered by Vaysman, that are not compatible with any of the current theories. In section 3.5 below I will propose a further refinement of the analysis of stress-quality interactions that allows such cases to be accounted for, arguing for formally separating prominence from metrical structure.

First, however, I will discuss an alternative view of sonority-driven stress that relates it to quantity-driven stress.

3.6.1 Quantity or sonority?

Stress attraction by high-sonority vowels appears related to stress attraction by heavy syllables and long vowels. High sonority vowels are longer and, in some sense, more prominent than low-sonority vowels. The question is then, are sonority-driven stress and quantity-driven stress two manifestations of the same phenomenon, and can they be collapsed theoretically? In other words, can the sonority hierarchy be treated as a manifestation the weight hierarchy? Here I review arguments for and against such a move. It appears that the evidence at present is inconclusive.

Kenstowicz (1994) and de Lacy (2003) treat sonority-driven stress as a theoretically separate phenomenon from quantity-driven stress. The weight hierarchy CVV ⟩⟩ CVC ⟩⟩ CV is separate from the sonority hierarchy a ⟩⟩ e,o ⟩⟩ i,u ⟩⟩ ə, and the two can in principle act independently in the same language. An approach that treats weight- and sonority-driven stress as the same phenomenon would collapse the two hierarchies into one general quantity hierarchy. One way to do this is to subdivide each of the elements of the weight hierarchy based on sonority of the nucleus vowel, as shown below.

(43) Cā ⟩⟩ Cē,ō ⟩⟩ Cī,ū ⟩⟩
CaC ⟩⟩ CeC, CoC ⟩⟩ CiC, CuC ⟩⟩ CəC ⟩⟩
Ca ⟩⟩ Ce, Co ⟩⟩ Ci, Cu ⟩⟩ Cə

There are two arguments in favor of collapsing the weight and sonority hierarchies, and three arguments against doing so. I will go through these five arguments in this section, but withhold judgment on what the correct analysis is.

First, the majority of sonority-driven stress languages have either an 'unstressable *ə*' system, where stress is repelled from schwas, or a 'stressed *a*' system, where stress is attracted to low vowels, and no vowel length distinctions. All such systems can be analyzed as quantity-driven, with quantity predictable from vowel quality. At least in some cases such an analysis seems preferable. In Kara (Schlie & Schlie 1993, de Lacy 2003), for example, stress is assigned according to the following algorithm.

(44) a. Rightmost syllable containing [a]
b. Else rightmost closed syllable
c. Else stem-initial

The vowel system of Kara is shown below. Apart from [a] and [ə], all vowels exist in tense-lax pairs, whose distribution is at least in part predictable.

(45)	i ɪ				ʊ u	
	e ɛ		ə		ɔ o	
			a			

Omitting the application of laxness harmony and some other low-level rules, the distribution of tense-lax vowels is shown below in (46). The distribution of the mid vowel pairs is entirely predictable, while for high vowels, there appears to be a contrast in closed syllables. Schlie and Schlie, however, analyze tense high vowels in closed syllables as coming from underlying [iə]; if this analysis is correct, then the distribution of tenseness in high vowels is also entirely predictable from syllable structure. In contrast, [a] and [ə] are distinct in all syllables.

(46) a. [e, o]: in open syllables
[ɛ, ɔ]: in closed syllables
b. [i, u]: open syllables
[i ~ ɪ, ʊ ~ u]: contrast in closed syllables (S&S analyze [i] in closed σs as /iə/)
c. [a ~ ə]: contrast in all syllables

These facts suggest that the [a~ə] distinction is different from the tense-lax distinction in the other vowels. In addition to the distributional evidence given above, there are two other reasons not to treat the [a~ə] pair as tense-lax. First, when there is a partial tense-lax contrast in a language, it is usually the HIGHER, not the lower vowels that are contrastive (as in English): Kara would then be typologically unusual if [a~ə] are a tense-lax pair. Secondly, only the high and mid vowel pairs but not the [a~ə] pair participate in laxness harmony, as summarized below.

(47) **Kara laxness harmony**

In open σs: [ɛ] before [ɪ,ɛ]
[ɔ] before [ɔ] (no data on [ʊ])
[a, ə] do not undergo this process

I take these facts to suggest that [a~ə] is not a tense-lax pair, but should be analyzed in terms of quantity: [a] is a long (bimoraic) vowel, and [ə] is its short counterpart. If this is the case, the stress rule then takes the following form.

(48)
a. Stress the rightmost long vowel
b. Else the rightmost closed syllable
c. Else initial

This new formulation of the stress rule makes no reference to vowel sonority, but only to the weight hierarchy CVV 〉〉 CVC 〉〉 CV. There happens to be an indirect piece of evidence that this analysis is correct. The vowel system of Nalik, a closely related language (Volker 1998), shows a vowel length distinction only in the low vowels: there is a [a~ā] contrast, but no other quantity contrast. Unfortunately, Volker provides no description of stress in Nalik.

(49) Nalik vowel system (Volker 1998)

i		u
e		o
	a, ā	

In sum, at least some systems that superficially involve sonority-driven stress can, and perhaps should, be reanalyzed in terms of quantity. The proposed reanalysis of Kara is applicable to any system that has either a 'stressed *a*' or 'unstressable *ə*' generalization, and no other quantity distinctions.

The second argument in favor of collapsing the weight and vowel sonority hierarchies comes from the typology of interaction between weight and sonority. Arto Anttila (p.c.) observes that whenever a stress system has both quantity and sonority sensitivity, SONORITY NEVER TRUMPS QUANTITY. In other words, preferential assignment of stress to syllables with higher-sonority nuclei is only observed in case there is no other quantity distinction. While a syllable like [Ca] may attract stress over a syllable like [Ci], one never observes syllables like [Ca] attracting stress over BOTH [Ci] and [Cī]. This fact suggests that a hierarchy like the one in (43) might reflect the typological facts in a more restrictive way than two separate hierarchies for weight and vowel sonority.

To sum up, there are two arguments in favor of treating sonority-sensitivity as a kind of weight-sensitivity: the preponderance of 'unstressable *ə*' and 'stressed *a*' systems, and the typology of interaction between weight and sonority.

Now let me turn to the three arguments against collapsing weight and sonority (I rely on the discussion in Kenstowicz 1994). First, a very complex hierarchy like the one in (43) raises the question as to how weight is represented. Weight is thought of standardly as a structurally encoded property. The mora count, a unit of timing, is the device by which heavy syllables are distinguished from light syllables. Collapsing the weight and sonority hierarchies leads to more weight distinctions in a single language than can be reasonably handled by a moraic theory. Some languages which show sonority sensitivity in their stress systems do have stress-relevant weight distinctions. Asheninca (Payne 1981) and Finnish (Anttila ms.) are two such languages. In the most extreme cases, where both weight and sonority play a role in stress assignment, a distinction of as many as five levels of weight may be required. A theory that collapses the weight and sonority hierarchies would have to allow a range of syllables from mono- to pentamoraic.

(50)	CVV, CVC	μμμμμ
	Ca	μμμμ
	Ce,o	μμμ
	Ci,u	μμ
	Cə	μ

The second argument against collapsing weight is sonority is that, as pointed out by Kenstowicz 1994, vowel quality does not necessarily correlate with phonetic length even in those languages that have sonority-driven stress: it is not generally true that length decreases significantly with vowel height.

The third and final argument that the sonority hierarchy in its role in stress assignment is not simply a refinement of the stress hierarchy is that sonority fails to play a role in other weight-sensitive phenomena, most notably word minimality. I know of no languages where the distinction between an acceptable and a subminimal word is made according to sonority: e.g. a language where Ca would be an acceptable word, but Ce and Ci would not.

Based on this discussion, it appears that at present it remains inconclusive whether the vowel sonority hierarchy can be collapsed with the weight hierarchy for the purposes of stress assignment.

3.6.2 Separating prominence from metrical structure

Let me now turn to my proposed refinement of the analysis of sonority-driven stress.

There are three primary correlates of metrical prominence: pitch, duration, and intensity. On the standard view, pitch and duration behave differently than intensity in that they can be used contrastively independently of metrical structure. Both pitch contrasts and contrasts in vowel length can be misaligned with metrical prominence:

some languages tolerate high tones on unstressed syllables, and some languages tolerate unstressed heavy syllables. At the same time, there are universal preferences in languages that attract stress to tones and heavy syllables, or, conversely, that raise the pitch of stressed vowels or lengthen them. While for duration and pitch participate in a bidirectional interaction with metrical structure, the relationship between intensity and stress is commonly thought to be unidirectional: in dynamic stress systems, intensity realizes metrical prominence, but intensity itself, as is generally assumed, cannot be used contrastively independently of metrical structure.

I suggest here that intensity can be brought into line with the other two correlates of stress and treated as a contrastive feature. Constraints on the relationship between stress and intensity would then function analogously to stress-to-weight, weight-to stress, stress-to-tone, and tone-to-stress constraints. Such an analysis empirically depends on finding cases where there is a mismatch between intensity and metrical structure. I will suggest below that cases of this type do indeed exist.

Furthermore, I propose to treat stress-sonority interactions as mediated by intensity. This move has the advantage of limiting the bidirectional interactions of stress to the three main correlates of intensity, duration, and pitch, and for permitting an analysis of systems where metrical structure mismatches with sonority-driven prominence.

3.6.2.1 The dual system of Mari

In a recent paper, Olga Vaysman uncovered a sonority-sensitive stress system that may shed light on the nature of quality-sensitive stress in general (Vaysman 2005). Mari (Finno-Ugric, Russia) distinguishes two levels of sonority: [ə] and the rest of the vowel system. Main stress is assigned in a default-to-opposite fashion: the rightmost full vowel is stressed, but if every vowel in the word is a [ə], then stress falls on the initial syllable. The standard analysis (e.g. Halle and Vergnaud 1987) of such stress systems involves right-headed unbounded feet and an initial default. In a rule-based theory, every full vowel and every initial syllable project a line 1 grid mark, and a right-headed unbounded foot is constructed on line 1.

However, there is direct evidence that Mari also has foot structure which does not coincide with the sonority-sensitive stress. Underlying schwas can surface either as short vowels, or as full vowels whose exact nature depends on the harmonic context. Whether a vowel surfaces as [ə] or as a full vowel depends on syllable count: if the schwa is in an odd-numbered syllable, it will surface as a full vowel, otherwise it remains [ə].

The schwa vocalization process renders the stress assignment rule opaque: in words with no underlying full vowels, stress falls on the initial syllable even if some of the underlying schwas surface as full, because it is assigned at the stage in the derivation when all vowels in the word are still schwas.

Cases of metrical opacity, when different strata in the phonology have different stress rule, have been well-known in the literature. For example, in Huariapano, there is a segmental process of *h*-insertion sensitive to left-to-right syllabic trochees, while the surface stress pattern in most words involves right-to-left syllabic trochees (Parker 1998). In Jarawara (Dixon 2004), surface stress is assigned with right-to-left trochees, while there is a host of segmental processes all sensitive to left-to-right feet. What these cases of metrical opacity have in common is that it is the last round of stress assignment that survives on the surface qua stress; evidence for the earlier rounds comes from segmental processes conditioned by metrical structure that does not survive in the output. The Mari situation is different, however. If Mari were to be treated as a case of opacity, with one stress rule applying at one stratum and the other stress rule at a later stratum, the sonority-sensitive stress must apply first, because it is counterfed by schwa vocalization. The rhythmic stress rule that is responsible for the feet that condition schwa vocalization must then apply at a later stage in the derivation. And yet, it is the output of the earlier rather than the later stress rule that survives on the surface.

Vaysman's way of addressing the problem of the two incompatible stress systems in Mari is to suggest that only one of the two systems, the binary left-to-right feet responsible for schwa vocalization, has to do with metrical constituency. The sonority-sensitive system, contrary to the standard analysis, does not involve unbounded feet.

The Mari case analyzed by Vaysman demonstrates that not all stress systems involve metrical constituent structure, and it is the sonority-sensitive system that operates independently of foot placement. The question then becomes, can the generalization that foot-placement constraints have no access to the segmental details such as vowel quality be saved if all sonority-driven stress systems are analyzed in a Mari-like way that does away with metrical constituency? In this section I pursue an analysis that makes such a separation between constituent-based and prominence-based stress systems.

The separation between prominence and rhythm is not a new idea in the theory of stress. Functionally speaking, stress serves two separate functions: on the one hand, rhythmic organization of the segmental material, and on the other, providing cues to word boundaries. Stress placement constraints fall into two natural classes depending on which of the two functions of stress they cater to: on the one hand there are rhythmic constraints governing the placement of binary feet, such as FTBIN, *LAPSE, *CLASH, etc., and on the other hand, there are constraints such as CULMINATIVITY and edge alignment constraints that serve the second function of stress. The separation between these two sets of preferences and functions of stress has always been at least implicit in stress theories. For example, the analysis of prominence-driven systems in Walker (1997) assumes that not all stress systems involve metrical constituent structure: unbounded systems, on her analysis, are purely prominence-driven, with factors such as syllable weight and sonority, or lexical factors, determining the placement of stress, and with no iterative or unbounded feet. Languages that fall into this category are those with default-to-same and default-to-opposite stress patterns driven by quantitative factors (weight, sonority, length, such as in Mari and Mongolian), or by lexical factors (such as Indo-European accent). Walker's analysis relies on the following sets of constraints.

(51) a. ALIGN constraints: aligning the stressed syllable with a word edge;
b. Licensing constraints: licensing marked stressed syllables (i.e. $\acute{\sigma}_{\mu}$) at an edge;
c. PKPROM: constraint expressing quantity sensitivity
d. NONFINALITY

None of these constraints mention metrical constituent structure: the preferences have nothing to do with rhythmic organization. Instead, the constraints in (51) express two conflicting preferences: to have stress uniformly aligned with an edge, and to preferentially stress prominent syllables.

While this set of constraints works for the systems that Walker analyzed – purely prominence-based stress systems without any evidence for footing – nothing precludes the coexistence of prominence-based stress with prosodic constituency below the word level, as Vaysman's Mari case clearly demonstrates. The interesting cases are those that involve an interaction between footing and prominence.

3.6.2.2 Stress or prominence?

Along the lines of stress-to-tone and tone-to-stress attraction investigated by de Lacy (2003) and the standard OT assumptions about the interaction of stress and weight, I propose the following constraints to regulate the relationship between metrical structure and prominence. Formally, I treat prominence as a feature of vowels, much like nasality or height: a vowel that is [+prom] is realized with increased intensity. The usual set of faithfulness constraints regulates the mapping of this feature between the input and the output, and stress-to-prominence and prominence-to-stress constraints, here formulated in a parallel fashion to ✦SWP and ✦WSP constraints, regulate the relationship between this feature and metrical structure.

(52) Faithfulness constraints
MAX-PROM 'Input [+prom] is realized as output [+prom]'
DEP-PROM 'Output [+prom] realizes input [+prom]'
NOFLOP-PROM '[+prom] is associated with the same segment in the input and the output'

(53) Attraction of prominence to metrical structure

Stress-to-prominence:	S→P	'if stressed, then prominent'
Prominence-to-stress:	P→S	'if prominent, then stressed'

In addition, I posit markedness constraints that regulate the appearance of the [+prom] feature in words independently of stress. These are the constraints responsible for sonority-driven prominence.

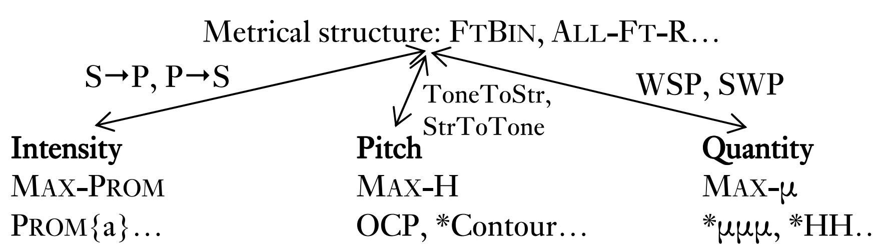
(54) a. Prominence-to-sonority
PROM{a}, PROM{a, o/e}, PROM{a, o/e, i/u}, PROM{a, o/e, i/u, ə}
*PROM{ə}, *PROM{ə, i/u}, *PROM{ə, i/u, o/e}, *PROM{ə, i/u, o/e, a}
'if a syllable contains *a*, it is prominent' …
'if a syllable contains *ə*, it is not prominent' …

b. Other constraints

CULM-PROM	'No more than one prominent syllable per phonological word'
*PROMCLASH	'No prominence clash'

On this analysis, the generalization that purely metrical, stress-related constraints do not make reference to the featural details of vowels. Stress can only be attracted to one of the three correlates of metrical structure: quantity, pitch, and loudness. All cases of stress-to-segments attraction are handled in a uniform way: constraints other than stress constraints may cause the attraction of quantity, pitch, or loudness to syllables with particular segmental features, and stress may in turn be attracted by constraints such as those in (53) to the resulting prominence, high pitch, etc. The general picture is given in (55) below.

(55)



At the core of my proposal is the idea that intensity can be treated on a par with the other two primary correlates of stress, pitch and duration. The set of constraints used to deal with intensity-stress interactions is entirely parallel to the constraints used for tone and weight: there is a pair of constraints aligning metrical heads with intensity, in two directions, and there is a set of constraints on the relationship between the feature [prom] and the content of a representation.

3.6.2.3 Factorial typology in standard OT

In this section I outline the predicted typology of stress given the enriched constraint set that separates stress from prominence. I go through the factorial typology generated by OTSoft using the following nine constraints (56). To simplify things, I consider under the heading RHYTHMIC the set of constraints that create a L→R syllabic trochee pattern, and of the sonority-prominence constraints I only include PROM{a}.

(56) a. RHYTHMIC (syllabic trochees L→R)
b. PROM{a}
c. S→P, P→S
d. MAX-PROM, DEP-PROM, NOFLOP-PROM
e. HAVE-STRESS, HAVE-PROM

I consider three inputs: an input with syllables without high-sonority vowels and without lexical stress, an input with a high-sonority vowel, and an input with a lexical stress. These inputs and the outputs included for each of the inputs are given below. Acute accents indicate prominence; the notation *sá* in the output indicates that a syllable has a stressed high-sonority vowel.

(57)	a.	/σ σ σ σ σ/	[(s s)(s s)s]
			[(ś s)(s s)s]
			[s s s s s]
			[ś s s s s]
	b.	/σ σ_a σ σ σ/	[(s s)(s s)s]
			[(ś s)(s s)s]
		2nd syl. has *a*	[s s s s s]
			[ś s s s s]
			[(s sá)(s s)s]
			[s (sá s)(s s)]
			[(ś sá)(s s)s]
			[s sá s s s]
			[ś sá s s s]
	c.	/σ σ́ σ σ σ/	[(s s)(s s)s]
			[(ś s)(s s)s]
		2nd syl. has	[s s s s s]
		lexical prom.	[ś s s s s]
			[(s ś)(s s)s]
			[s (ś s)(s s)]
			[(ś ś)(s s)s]
			[(s ś)(s ś)s]
			[ś ś s s s]

The factorial typology contains 45 output types, classified below based on the stress system type they generate. In each case, the three lines show the outputs of the three inputs from (57): the first line is for an input without any high-sonority vowels or lexical stress, the second line for the input with a high-sonority vowel in the second syllable, and the third line for the input with lexical stress on the second syllable. Thus, each column of three lines represents an output pattern.

The first pattern is the simplest: rhythmic stress with no sonority sensitivity. Here RHYTHMIC must outrank MAX-PROM, PROM{a}, and P→S. The ranking of S→P and DEP-PROM determines whether the binary rhythmic stress is realized overtly or covertly. There are many languages like this: any language with syllabic trochees built on the left edge without any sensitivity to factors such as vowel quality and without any exceptions.

I. Rhythmic stress, no sonority sensitivity, no faithfulness

RHYTHMIC ⟩⟩ MAX-PROM, PROM{a}, P→S

```
(ś s)(s s)s        (s s)(s s)s
(ś s)(s s)s     (s s)(s s)s
(ś s)(s s)s        (s s)(s s)s
```

The second pattern involves lexically unpredictable stress without sonority sensitivity, where MAX-PROM and P→S must outrank RHYTHMIC and DEP-PROM must outrank the sonority constraint PROM{a}. The system generates three possible outcomes. First, there is the system with a rhythmic default at the left edge but some lexical exceptions with stress on the second syllable. This involves the high-ranked S→P, forcing all stressed syllables to be realized as prominent. If this constraint is low-ranked, metrical structure will be covert just in those cases where there is no lexical stress (this pattern properly belongs with category VI below). In such a language, there could be a potential contrast between overt and covert metrical structure. The second sub-pattern in Class II involves metrical structure appearing only if there is unpredictable lexical stress, which also belongs in Class VI.

Languages instantiating the first subpattern include Polish and Macedonian, where a rhythmic default can have exceptions. The second and third subpatterns is harder to find for dynamic stress, but its analogue in pitch accent languages is common, where it is normal to have a contrast between accented and unaccented words.

II. Lexically unpredictable stress

MAX-PROM, P→S ⟩⟩ RHYTHMIC; DEP-PROM ⟩⟩ PROM{a}

```
(ś s)(s s)s        (s s)(s s)s
(ś s)(s s)s        (s s)(s s)s
s(ś s)(s s)        s(ś s)(s s)
```

The third class of outputs includes sonority-driven stress. Once again, there are two subpatterns here: one where metrical structure is overt in all words, and one where it is covert in words without high-sonority vowels. Examples of the former are any language with sonority-driven stress and footing dependent on it (e.g. Gujarati in de Lacy's description, Asheninca Campa). I am not aware of the examples of the second subpattern.

III. Sonority-driven stress

PROM{a}, P→S ⟩⟩ RHYTHMIC ⟩⟩ MAX-PROM, DEP-PROM

```
(ś s)(s s)s       (s s)(s s)s
s(sá s)(s s)      s(sá s)(s s)
(ś s)(s s)s       (s s)(s s)s
```

The system also predicts a combination of the patterns II and III, where both sonority sensitivity and lexical stress coexist.

II & III combined

PROM{a}, P→S, MAX-PROM ⟩⟩ RHYTHMIC, DEP-PROM

```
(s s)(s s)s       (ś s)(s s)s
s(sá s)(s s)      s(sá s)(s s)
s(ś s)(s s)       s(ś s)(s s)
```

Now we get to cases of mismatch between metrical structure and prominence. Class IV patterns have a default rhythmic structure throughout, but in case a word has a lexically marked prominent syllable, that syllable may not fall in the strong position of the foot.

IV. Mismatch between lexical prominence and metrical structure

RHYTHMIC, MAX-PROM ⟩⟩ P→S, PROM{a}

```
(ś s)(s s)s       (s s)(s s)s
(ś s)(s s)s       (s s)(s s)s
(s ś)(s s)s       (s ś)(s s)s
```

Next comes the type of stress system that was the original motivation for the reanalysis of stress proposed here: Mari-like mismatch between sonority-driven prominence and metrical structure. Once again, two variations on the system are predicted, depending on the ranking of S→P: cases where all feet are overt and cases where no prominence is realized just in case there is no high-sonority vowel in the word.

V. Mismatch between sonority-driven prominence and metrical structure

RHYTHMIC, PROM{a} ⟩⟩ MAX-PROM, P→S

```
(ś s)(s s)s        (s s)(s s)s
(s sá)(s s)s     (s sá)(s s)s
(ś s)(s s)s      (s s)(s s)s
```

The system also predicts a combined case of mismatch between rhythmic structure, sonority-driven stress, and lexical stress.

IV & V combined

RHYTHMIC, MAX-PROM, PROM{a} ⟩⟩ P→S

```
(ś s)(s s)s      (s s)(s s)s
(s sá)(s s)s       (s sá)(s s)s
(s ś)(s s)s        (s ś)(s s)s
```

Next come some potentially problematic cases. There is a large and heterogeneous class of outputs with some sort of contrast between lexically accented and lexically unaccented words. Not many of these outputs are represented in the typology; there appear to be few languages with dynamic stress where there is a contrast between accented and unaccented words, although such a pattern is common in the pitch accent analogues of the stress systems.

VI. Contrast between accented and unaccented words (possibly dependent on sonority)

```
s s s s s       (s s)(s s)s  s s s s s    s s s s s       s s s s s
s s s s s       (s s)(s s)s  s sá s s s   s(sá s)(s s)    s(sá s)(s s)
(ś s)(s s)s     (ś s)(s s)s  s(ś s)(s s)  (ś s)(s s)s     s s s s s

s s s s s       s s s s s    s s s s s     s s s s s      (s s)(s s)s
s(sá s)(s s)    s sá s s s   s sá s s s    s s s s s      (s sá)(s s)s
s(ś s)(s s)     (ś s)(s s)s (s ś)(s s)s   (s ś)(s s)s     (ś s)(s s)s

(s s)(s s)s     (ś s)(s s)s  (ś s)(s s)s   (ś s)(s s)s    s s s s s
s(sá s)(s s)    s sá s s s   s sá s s s    s sá s s s     s  s  s  s  s
(ś s)(s s)s     (ś s)(s s)s  (s ś)(s s)s   s(ś s)(s s)    s(ś s)(s s)
```

Another systematic set of patterns that are not attested involves "too much stress", i.e. cases where intensity realizes two independent things in a language.

VII. Too much stress

```
(ś s)(s s)s  (ś s)(s s)s    (ś s)(s s)s   (ś s)(s s)s   (ś s)(s s)s
(ś s)(s s)s  (ś sá)(s s)s   (ś sá)(s s)s  (ś sá)(s s)s  (ś sá)(s s)s
(ś ś)(s s)s  (s s)(s s)s    (ś s)(s s)s   s s s s s     (ś ś)(s s)s

(ś s)(s s)s  s s s s s      s s s s s
s sá s s s   s s s s s      s sá s s s
(ś ś)(s s)s  (ś ś)(s s)s    (ś ś)(s s)s
```

Below I give the rankings that generate the patterns in VII.

(58) a. S→P, RHYTHMIC, PROM{X} ⟩⟩ MAX-PROM
both sonority-driven prominence and rhythmic stress
b. S→P, MAX-PROM, PROM{X} ⟩⟩ RHYTHMIC
both sonority-driven prominence and unpredictable prominence
c. S→P, RHYTHMIC, MAX-PROM ⟩⟩ PROM{X}
both rhythmic stress and unpredictable prominence
d. all four constraint types high-ranked
all three stress types in one system: sonority, unpredictable prominence, rhythm

A system such as (58)a would have both sonority-driven prominence and, independently of it, a rhythmic stress pattern realized dynamically. All of the other systems in (58) would also have two independent stress patterns realized with the same cue.

This problem is not particular to my analysis, however: the constraints on tone-to-stress alignment also predict parallel 'too-much-tone' systems where a rhythmic pattern realized tonally coexists with lexically unpredictable tone.

(59) RHYTHMIC, STRESS-TO-TONE, MAX-H ⟩⟩ TONE-TO-STRESS
Unpredictable tones combined with a rhythmic tonal pattern

As (59) shows, the problem is not limited to my account: any theory that has both tone-to-stress and stress constraints and tone faithfulness constraints predicts that having both high-ranked would create a language with two separate functions of tone: it would realize both metrical constituent structure and appear on some vowels unpredictably, protected by the faithfulness constraints.

Although the constraint system systematically does not rule out such 'too-much-stress' and 'too-much-tone' systems, it is plausible to assume that they do not occur for other reasons, external to the OT grammar. Once again, this problem is not unique to my proposal, but is faced by any theory of stress where faithfulness plays a role, and is not unique to prosodic phonology either: too much faithfulness in segmental phonology can also lead to predicted but implausible languages. I follow the orthodox view that the non-existence of such systems does not pose a serious challenge for the theory.

Another small set of patterns predicted by my system involve no metrical structure at all.

VIII. No rhythmic structure

```
s s s s s       s s s s s
s s s s s       s sá s s s
s s s s s       s s s s s
```

And finally, there are a few oddball cases where the system predicts a rather bizarre outcome: stress is *not* realized just in case it is marked lexically. This arises because high-ranked NOFLOP, RHYTHMIC, and S→P make it impossible to realize stress in non-default positions: realizing it on the second syllable would violate S→P, shifting foot structure one syllable to the right would violate RHYTHMIC, and realizing it on the first syllable would violate NOFLOP. It remains to be seen how such cases can be ruled out.

IX. Stress not realized just in case it's marked lexically

```
(ś s)(s s)s     (ś s)(s s)s     (ś s)(s s)s       (ś s)(s s)s
(ś s)(s s)s     (ś s)(s s)s     (s sá)(s s)s      s sá s s s
s s s s s       (s s)(s s)s     (s s)(s s)        s s s s s

(ś s)(s s)s     (ś s)(s s)s
s(sá s)(s s)    s(sá s)(s s)
(s s)(s s)s     s s s s s
```

Finally, what the ill-behaved cases in VI-IX have in common is "extreme arrhythmia" (Paul Kiparsky, p.c.). They have either adjacent stresses (ś ś), or no stress at all (s s s s s). Once again, such extremely dysfunctional systems, predicted by any theory of stress that has faithfulness to prosodic structure, can be ruled out on general grounds of culminativity.

Although the typology of stress outlined here is richer than the typology predicted by the standard theory that does not treat intensity as potentially contrastive, in some areas my proposal makes more restrictive predictions. A hypothesis implicit in the preceding discussion has been that lexically unpredictable stress arises due to the action of faithfulness constraints that refer not to stress (i.e. metrical structure) directly but to one of its correlates, e.g. intensity. I propose here that this is the only way for lexical stress to arise: there are no faithfulness constraints that refer directly to stress, expressing the idea that the function of prosodic constituent structure is not to create contrasts for the language but to organize the material into manageable chunks.

In other words, the only way that unpredictable stress can arise is through the action of faithfulness constraints that operate on intensity, pitch, and quantity. MAX-PROM can lead to unpredictable stress such as in Russian, as outlined above; MAX-H creates unpredictable stresses in pitch accent languages such as Ancient Greek, and MAX-μ creates weight-sensitive stress.

Doing away with stress faithfulness constraints that refer to metrical structure, e.g. NO-FLOP-STRESS, predicts that *covert prosodic structure is always predictable*. If a language has only abstract feet – structure created by the action of constraints such as FTBIN or AL-FT-L, but not aligned to either pitch or intensity, such a stress system ought to have no exceptions. There should be no languages like Capanahua where the covert feet are left-aligned in some arbitrary set of words and right-aligned in others. Likewise, there should be no covert contrast in the placement of inaudible foot boundaries: because the are no faithfulness constraints referring to foot boundaries, there should be no contrast between, say, (σ σ) σ σ and (σ σ σ) σ.

3.7 Some applications

My proposal in this section is not specific to prosody-segmental interactions. Rather, I suggest that the too-many-solutions problem that is most clearly seen in the domain of stress-segmental interactions is in fact present elsewhere in OT and result from its general claim that all generalizations are in the outputs. I have argued that this position is too radical; prosody-segmental interactions are subject to a different type of generalization. In Chapter 4 below I will discuss in detail another domain where procedural generalizations are crucial – the typology of vowel syncope and epenthesis. In this section I briefly survey some other cases in phonology where my proposal can prove fruitful, leaving a detailed investigation for future research.

In (60) I give a sample of segmental constraints stated procedurally and the typological consequences of the ICP for each constraint. In each case, the effect that is ruled out involves modifying the property mentioned in the antecedent of the constraint.

(60) a. '✦If a velar is followed by a front vowel, then it is palatalized'
 ☞No vowel backing after velars
 b. '✦If a consonant is intervocalic, then it is voiced'
 ☞No vowel deletion to make voiceless consonants not intervocalic
 ☞No consonant epenthesis to make voiceless consonants not intervocalic
 c. '✦Short front vowel followed by *r* are realized as [e]' (cf. Latin)
 ☞*r* cannot be deleted, turned into *l*, etc. just after short front vowels
 d. '✦If a nasal is followed by a consonant, then it is homorganic with that consonant'
 ☞No epenthesis to relieve non-homorganic NC clusters
 ☞No deletion of C in non-homorganic NC clusters

Now consider harmony, which I will discuss in more detail. A notorious problem in any OT analysis of harmony is that the constraint penalizing disharmonic configurations does not say anything about how such configuration can be repaired. Consider a simple example of nasal harmony modeled after Pater 2003. Suppose that a language has progressive nasal harmony, and that glides are transparent to harmony while voiceless fricatives block it. Sample inputs and outputs and a tableau are given below.

(61) a. /ãwa/ → ãw̃ã
b. /ãsa/ → ãsa

(62)

			MAX[nas]	*s̃	AGREE	DEP[nas]
/ãwa/		ãwa			*!	
	☞	ãw̃ã				**
		awa	*!			
/ãsa/	☞	ãsa			*	
		ãs̃ã		*!		**
		asa	*!			

The problem with such an analysis is that the factorial typology is too rich. Given the input /ãsa/, the grammar in (62) chooses to sacrifice harmony for the sake of avoiding the marked segment *s̃*, and for the sake of avoiding a violation of MAX[nas]. However, nothing precludes a ranking where both harmony and **s̃* dominate the faithfulness constraint, resulting in a language that applies harmony where it can and deletes the harmonizing feature in blocking contexts. Such systems are unattested.

(63)

			*s̃	AGREE	MAX[nas]	DEP[nas]
/ãwa/		ãwa		*!		
	☞	ãw̃ã				**
		awa			*!	
/ãsa/		ãsa		*!		
		ãs̃ã	*!			**
	☞	asa			*	

The reason this problem arises is the same reason that was behind the incorrect predictions with respect to prosody-segmental interactions: there is a generalization about input-output mappings at work. This means that AGREE is a procedural constraint subject to the ICP. One possible statement of the constraint is given below, where the word 'domain' stands for the relevant harmony domain of the language (in our hypothetical example, the word).

(64) ✦AGREE
If *x* is a segment within a domain that contains a [+nas] segment, then *x* is [+nas].

The candidate *asa* in tableau (63) is exactly the type of candidate that the ICP is intended to rule out. In order to satisfy the implicational statement of the harmony constraint (64), this candidate modifies the property that is mentioned in the antecedent of the constraint rather than in the consequent. The ICP rules this out, as shown in the following tableau (65). Both the candidates *awa* and *asa* incur ✦ violations of the AGREE constraint by the ICP, because they have segments that are not within a domain that has a [+nas] segment while their t-correspondents in the designated candidates *ãwa* and *ãsa* are within such a domain. Incorporating the effects of ICP into the ranking will make it impossible for the candidates to win if the harmonizing feature is deleted just in those cases where harmony is impossible.

(65)

				*s̃	✦AGREE	MAX[nas]	DEP[nas]
/ãwa/		✦	ãwa		*!		
	☞		ãw̃ã				**
			awa		✦!	*!	
/ãsa/	☞	✦	ãsa		*		
			ãs̃ã	*!			**
			asa		✦	*!	

The full factorial typology for the four constraints contains only three patterns: (1) harmony applying to both /ãwa/ and /ãsa/, (2) harmony applying to neither, and (3) harmony applying to /ãwa/ but not to /ãsa/. In this particular case, the designated candidates of the constraint AGREE is always *ãwa* and *ãsa*, no matter what the ranking of the remaining three constraints, so the factorial typology can be computed simply by reranking the constraints as given in (65).[21]

(66) a. DEP ⟩⟩ AGREE (12 grammars)
Harmony applies in no forms

b. {AGREE ⟩⟩ DEP, *s̃ ⟩⟩ AGREE} (6 grammars)
Harmony applies in /ãwa/ but not /ãsa/

c. {AGREE ⟩⟩ DEP, AGREE ⟩⟩ *s̃} (6 grammars)
Harmony applies in all forms

Now consider a more complicated situation. In Guaraní and some related languages, the domain of nasal harmony is defined by the main stress foot: nasal harmony proceeds right-to-left up to the stressed syllable. One problematic prediction of standard OT is that it should be possible for stress to shift onto final nasal vowels just to prevent nasal harmony from applying. Such an outcome is given in the tableau below.

(67)

			*s̃	AGREE$_{\varphi}$	MAX[nas]	DEP[nas]	STRESSINITIAL
/tawã/		áwã		*!			
		ã́w̃ã				*!*	
	☞	awã́					*
		áwa			*!		
/tasã/		ásã		*!			
		ã́s̃ã	*!			**	
	☞	asã́					*
		ása			*!		

However, given the ICP, the unwanted outcome is no longer possible. The designated candidates of the constraint AGREE always have initial stress, because there is no constraint that can override default stress assignment. The candidates *awã́* and *asã́*

[21] Adding other constraints to the picture will, of course complicate matters somewhat. For example, high-ranking *NASALVOWEL will cause the designated candidates for AGREE to be *awa* and *asa*, resulting

violate the constraint ✦AGREE$_{\varphi}$ by : the first vowel *a* in *awã́* does not have the property of being in a stress foot that contains a [+nas] segment, while its correspondent in the DC does have this property. Likewise, the first vowels of *áwa* and *ása* also do not have the property in question, due to the absence of [+nas] feature in the domain, while their correspondents in the respective DCs do have the property. These candidates therefore incur ✦violations of ✦AGREE$_{\varphi}$.

(68)

				*s̃	✦AGREE$_{\varphi}$	MAX[nas]	DEP[nas]	STRESSINITIAL
/tawã/		✦	áwã		*!			
	☞		ã́w̃ã				**	
			awã́		✦!			*
			áwa		✦!	*		
/tasã/	☞	✦	ásã		*			
			ã́s̃ã	*!			**	
			asã́		✦			*!
			ása		✦	*!		

Once again, the factorial typology can be computed in a simple way, since the choice of the designated candidates of the constraint ✦AGREE$_{\varphi}$ does not depend on the ranking of the other constraints. The typology contains only three output patterns: (1) harmony applies in both /awã/ and /asã/; (2) harmony applies in neither; (3) harmony applies in /awã/ but not in /asã/ (as is the case in (68). The rankings are given in (69).

(69)
a. DEP ⟩⟩ AGREE (60 grammars)
Harmony applies in no forms

b. *s̃ ⟩⟩ AGREE ⟩⟩ DEP (30 grammars)
Harmony applies in /awã/ but not in /asã/

c. { AGREE ⟩⟩ s̃ ; AGREE ⟩⟩ DEP } (30 grammars)
Harmony applies in all forms

in outputs where [nas] is removed from vowels across the board.

CHAPTER 4

SYNCOPE AND EPENTHESIS WITH PROCEDURAL CONSTRAINTS

4.0 Introduction

4.0.0 Introduction

The last chapter I examined a case of asymmetrical interaction between two phonological domains, prosodic and segmental phonology. I have shown that the systematic failure of the featural details of the segmental makeup of a word to figure in stress assignment rules presents a general problem for output-oriented theories like OT. Because all significant generalizations are hypothesized in OT to lie in output structures, the theory is not designed to handle generalizations that are stated in terms of processes. Prosody-segmental interactions present just such a generalization: only some processes (stress-driven segmental changes), but not others (segmentally-driven stress) are observed typologically. I argued that such an asymmetrical generalization cannot be stated as a condition on output forms. An output condition requiring that metrical prominence and certain segmental features be localized within the same syllable or segment does not tell the whole story: what matters is not only where stress and segmental features are in the output, but how they got there. Derivationally, such generalizations about input-output mappings are quite straightforward to state. In OT, on the other hand, the input-output mapping emerges from the entire constraint ranking in the grammar. I have taken a direct approach by introducing procedural constraints that penalize candidates resulting from undesirable processes; violation patterns of such constraints must depend on the ranking of the remaining constraints in the grammar.

In this chapter I turn to a related type of generalization that cannot be stated in terms of output structures: ENVIRONMENT-based generalizations. In derivational theories, it is possible to state generalizations about processes being confined to apply only in some contexts but not others. As I will show, from the point of view of OT, such generalizations would again look like too-many-solutions problems, albeit from a different perspective. Environment-based generalizations appear as situations where a given process is used as a repair for some but not all constraints that could potentially force it. I will argue here that vowel epenthesis and metrically-driven vowel deletion (syncope) are subject to such generalizations, and that the proposed mechanism of procedural constraints is able to handle it. The procedural generalizations from the last chapter and the environment-based generalizations present two facets of the same problem, which has to do with the locus of the phonologically significant observation. In both cases, the most insightful statement of the typology should be made at a level other than the output. This property puts both kinds of problems within reach of the new procedural constraints introduced in the last chapter.

This chapter is organized around the two main empirical domains, epenthesis and syncope. I begin by showing that many markedness constraints systematically fail to force epenthesis. The real typological generalization about epenthesis will turn out to be an environment-based one: it only applies to resolve marked consonant clusters. I will then provide a solution in terms of the framework of constraint interpretation developed in the last chapter. Next, I move on to syncope, and show again that OT does not provide a solution to the straightforward environment-based generalization that syncope does not apply to stressed vowels, a generalization that can be easily accomodated by my theory. I will survey in detail analyses of several languages with syncope, showing that whenever stressed vowel syncope has been proposed, alternative and better analyses are available.

However, before turning to the empirical issues, I will devote the next section to clarifying some terms.

4.0.1 What is 'context' in OT?

Because there are only two levels of representation in canonical OT, the input and the output, the reader might rightly ask if there is any meaning to the term 'environment' in the context of parallel output-oriented theories like OT. The derivational notion of the context (structural description) of a rule has a straightforward analog in OT only in the special case where the entire structural description is present in the underlying representation. Only in that case is there a level of representation where the conditioning environment of the process is present but the process itself has not applied. For example, in rule-based terms, word-final obstruent devoicing applies in the context __#; the structural description of such a rule is [–son, +voi]#. In OT, it makes sense to talk about the environment of this process only if the [–son, +voi]# sequence to which final devoicing eventually applies is present in the underlying form. It is true of Russian, for instance: every consonant that devoices due to the final devoicing constraint *[–son, +voi]# is word-final in the underlying form, and thus the structural description [–son,+voi]# is present at that level of representation. The situation need not be so simple, however; it is easy to see what a case would look like where the structural description would not be present at ANY level of representation. Suppose final devoicing is fed by another process, e.g. final vowel deletion (apocope). A toy derivation is given below.

(1)	/taba/	
	tab	*apocope*
	tap	*final devoicing*
	[tap]	

In a rule-based theory, there is always a level of representation where the structural description [–son, +voi]# is present; in this case, it is the output of the apocope rule [tab]. Indeed, the existence of such a level of representation is necessary for the devoicing rule to apply at all. In OT, however, such intermediate representations are not available: the only two representations are the input /taba/ and the output [tap].

(2) /taba/

		*[–son, +voi]#	*V#	MAX	IDENT
	taba		*!		
	tab	*!		*	
☞	tap			*	*

Intuitively, the process that takes /b/ to [p] in (2) is final devoicing, but it is problematic to say so, because there is no level of representation where the voiced consonant [b] is final. The only form that qualifies is the losing candidate [tab], one among infinitely many other losing candidates. Thus, we must be more clear about what the words 'context of a process' mean in situations like (2), by specifying what is special about the candidate [tab], where the structural description is present, as opposed to all of the other losers.

Here we will work along the same lines as in the previous chapter, where I defined the notion of a constraint driving a process in terms of the optimal candidate in a grammar with the constraint in question removed. The intuition there was that a constraint causes a process to apply to a given phonological object if that object behaves differently depending on whether the constraint is present in the grammar or not. Here, the intuition is the same: we will look for the context of a process driven by a given constraint in the optimal candidate of the alternative grammar with that constraint removed. In our toy example, the relevant constraint is the one responsible for final devoicing, *[–son, +voi]#. The grammar without it picks [tab] as the winner, as shown below.

(3) /taba/

		*[–son, +voi]#	*V#	MAX	IDENT
	taba		*!		
☞	tab	*		*	
	tap			*	*!

In other words, we say that the environment of devoicing in this example is 'word-final' because the stop that devoices would have been word-final had devoicing not applied. A more strict definition is given below.

(4) ENVIRONMENT OF PROCESS (Definition)
Given a grammar *G*, a constraint *C*, and an input /i/, output [o], and some process that the constraint *C* forces in [o], the ENVIRONMENT of that process is minimal locus of violation of *C* in the optimal candidate [o'] in the grammar *G′* that is identical to *G* except that *C* has been removed.

With this machinery in place, it is now easy to state, in plain English, generalizations about environments of processes in OT terms. For example, the (false) claim that devoicing only applies word-finally would take the following shape: "Devoicing only applies to consonants that would have been word-final had devoicing not applied". The (true) claim that syncope does not apply to stressed vowels (see below in this chapter) would take the form: "Syncope applies only to vowels which would have been unstressed had syncope not applied". The form of these statements should already suggest to the reader that such generalizations can be handled by the constraint evaluation method introduced in the last chapter, which relies on a similar notion of the winner in an alternative grammar with some constraint removed. In this chapter I make the connection more precise.

4.0.2 Epenthesis and syncope: introduction

Let me now move on to the empirical domain of this chapter, vowel epenthesis and syncope. In a nutshell, the typological behavior of these two processes can be most economically and insightfully stated not in terms of the character of the output they produce but the environment in which they apply. Despite the superficial complementarity of these two processes, examining their environments shows that they respond to altogether different pressures. Epenthesis universally serves to break up marked consonant clusters, but, in general, does not cater to metrical constraints. Vowel syncope, on the other hand, applies to weak vowels – that is, vowels in light syllables that are unstressed, unfooted, posttonic, or word-final. Crucially, metrically-driven vowel deletion can never target stressed syllables. These are generalizations about ENVIRONMENTS in which the processes of syncope and epenthesis apply. There is no output condition that can capture such a generalization. Positing constraints that prohibit weak vowels on the surface – e.g., constraints against unstressed short vowels in open syllables, or constraints against vowels unparsed by feet – would ensure that no such vowels appear on the surface, but would do nothing to control context of syncope. If the generalization about syncope is best stated in terms of the hypothetical structure where syncope has not applied, then access to more than just the surface is needed: rather, the constraint responsible for syncope must have access to the entire ranking of the language, in a way that I will make precise in this chapter.

Let me now move on to the first case, vowel epenthesis.

4.1 Epenthesis

This section is organized as follows: I begin with discussing rule-based view of the typology of epenthesis, and then show in section 4.1.2 how OT makes a radically different claim about the potential environments in which vowel epenthesis may apply. Sections 4.1.3 and 4.1.4 are then devoted to my approach to constraining OT in order to achieve a better fit between the observed and predicted typology of epenthesis.

4.1.1 Rule-based views and the typology of epenthesis

In rule-based phonology, vowel epenthesis has been standardly assumed to be used in only two types of situations: as a response to syllable template requirements and as a repair for word minimality (McCarthy 1979, Selkirk 1981, Broselow 1982, Blevins 1985, Itô 1986, 1989). Pre-OT literature has been largely devoted to the debates surrounding the kind of information that epenthesis rules may make reference to. One derivational view of epenthesis is the skeletal rule theory (e.g. Blevins 1985). In this approach, syllabification rules created prosodic structure, which was linked to a CV or X-slot template. Epenthesis rules then inserted skeletal slots in certain positions, making reference to the "stray" or unsyllabified status of adjacent skeletal slots. Epenthesis rules in such a theory take the following form, where C' indicates the stray status of a skeletal slot.

(5) $\varnothing \rightarrow$ V / C'__ (Itô 1989: 217)
$\varnothing \rightarrow$ V / __C'

Contrasting with this view was the prosodic theory of epenthesis (Itô 1989), which argued against the diacritic use of the notion of 'strayness'. According to Itô's proposal, the site of epenthesis should follow directly from the independently needed principles of syllabification. This move, which eliminates the need for rules such as (5), allows to account for the facts in a principled way, making epenthesis a consequence of prosodic theory in general, rather than relying on arbitrary rules. For the purposes of the present discussion in the context of OT, this debate between prosodic and skeletal views of epenthesis is of largely historical interest. What is important, however, is the shared assumption by all pre-OT researchers that epenthesis is, universally, a response to prosodic requirements at the level of the syllable. It is this generalization, previously considered too obvious to merit extensive discussion, that is no longer available in canonical OT, precisely because it is a generalization about the ENVIRONMENT of a process, not just about output structures.

This generalization has been most explicitly laid out by Broselow (1982), who lists three types of epenthesis processes, in terms of their environments. I will follow Broselow's classification in the discussion here. First, there is SYLLABICALLY-conditioned epenthesis, which applies in order to relieve violations of the syllable template of a language. This type of epenthesis is by far the predominant one; the following table (6) presents a far from complete list of cases discussed in recent literature. The cases fall in three broad, and potentially overlapping, categories: languages that insert vowels into complex syllable margins (Icelandic, Kekchi, Mono, Lenakel, Arabic, Mohawk, Fijian, Welsh); languages that use epenthesis for restrictions on codas, whether due to a complete prohibition of codas (Maori, Selayarese, Tongan), or due to a prohibition of codas of a particular type (Sranan, Japanese); and, finally, languages where epenthesis is used to repair sonority sequencing violations.

(6)

LANGUAGE	ENVIRONMENT	SOURCE
Modern Icelandic	C__r#	Kenstowicz 1994: 79
Kekchi	ʔ__C#	Hall 2003
Mono	complex clusters	Hall 2003
Lenakel	complex clusters	Kenstowicz 1994: 126
Palestinian, Iraqi Arabic	complex clusters	Kenstowicz 1994
Mohawk	complex clusters; sonority	Rawlins 2006
Fijian	C-clusters in loans	Kenstowicz 2003
Welsh	Marked coda clusters	Hall 2003
Cook island Maori	after final C	Hall 2003
Selayarese	after final codas	Kenstowicz 2003
Tongan	after final codas	Kenstowicz 2003
Sranan Creole	after final obstruents	Alber & Plag 1999
Japanese	non-nasal codas in loans	Hall 2003
Ponapean	heterorganic clusters	Kenstowicz 1994
Alguerese Catalan	sonority sequencing	Lloret & Jiménez 2006
Berber	sonority	Hdouch 2004
Irish, Scots Gaelic	clusters of falling sonority	Green 1997
Mawu	CR in loans	Hall 2003

Second, Broselow terms METRICALLY-conditioned epenthesis the set of cases where epenthesis is used as a repair for word minimality. These cases might be more properly called MINIMALITY-conditioned, because, as I will show below, epenthesis is not used as a response to the broad set of metrical conditions it might be expected to respond to, but only to requirements on the minimal size of the phonological word. Let me give some examples of the alternations that result from minimality-driven epenthesis.

In Mono, for example, a two-syllable word minimum is satisfied by prefixing a word with a copy of its root vowel (7)a. The language also has syllable contact-conditioned epenthesis into CR clusters which interacts opaquely with respect to minimality-driven epenthesis (7)b. The data below are from Hall (2003); see also Olson 2001.

(7)	a.	/ʒī/	→	īʒī		'tooth'
		/bè/	→	èbè		'liver'
	b.	/jābrù/	→	jābùrù		'goat'
		/gré/ →	égré	→	égéré	'big'

In a similar case, Mohawk monosyllabic words are supplied with an epenthetic vowel to satisfy the disyllabic word minimum (8)a (cf. Rawlins 2006). Iraqi Arabic presents the same phenomenon from a slightly different perspective. The language has an initial epenthesis process in words that begin with a cluster. Normally, this epenthesis is optional, but becomes obligatory just in case the base word is a monosyllable that contains a short vowel not followed by a consonant cluster (8)b. Assuming final consonant extrametricality, this requirement amounts to a bimoraic word minimum enforced by epenthesis.

(8)	a.	Mohawk[22]				
		/keks/	→	í:keks	'I eat'	(Broselow 1982: 117)
		/weʔs/	→	í:weʔs	's/he/it is walking around'	(Rawlins 2006: 12)
	b.	Iraqi Arabic				
		ʔidrúus	~	drúus	'lessons'	(Broselow 1982: 125)
		ʔídrus	~	*drus	'study!'	

Another case with a disyllabic minimality-driven epenthesis is Lardil (Hale 1973), where monosyllabic roots receive a final epenthetic vowel. The following data are taken from Hale 1973: 427. In *r*-final stems, the vowel *a* is added to augment the form up to the disyllabic minimum (9)a; in other stems, the augment is the syllable *Ca*, where *C* is a stop homorganic to the stem final (9)b. The data in (9)c shows that consonant-final stems longer than one syllable are not augmented.

(9)	a. /ṭer/	→	ṭera	'thigh'
	/yur/	→	yura	'body'
	b. /ṛil/	→	ṛilta	'neck'
	/wun/	→	wunta	'rain'
	/ṭuṛ/	→	ṭuṛṭa	'excrement'
	c. /thuŋal/	→	thuŋal	'tree'
	/kentapal/	→	kentapal	'dugong'
	/kethar/	→	kethar	'river'
	/miyaṛ/	→	miyaṛ	'spear'

Note that these examples of minimality-conditioned epenthesis all involve word minimality, not foot minimality. That kind of epenthesis, which would bring degenerate feet up to the required minimum whenever they arise, is not attested.[23] This fact shows that, although word minimality is caused by a minimal foot size requirement, it is not FTBIN per se but the constraint that requires every grammatical word to to comprise a prosodic word (GRW=PRW) that is capable of causing epenthesis.

The third set of cases in Broselow's typology is what she calls SEGMENTALLY-conditioned epenthesis. Broselow's criterion for an epenthesis process to qualify as segmentally-conditioned appears to be its reference to information across syllable boundaries, rather than to information on the type and content of a syllable margin, which is entirely contained within one syllable. In this category are the cases such as Dorsey's Law in Winnebago, where a sequence of an obstruent followed by a sonorant is disallowed and broken up by epenthesis (Miner 1979, Alderete 2006[1995]).

---

[22] Mohawk also has a cluster-driven epenthesis process, discussed at length by Rawlins 2006, which interacts in an interestingly opaque way with stress.

[23] In such a language, for example, all odd-parity words could receive an epenthetic vowel.

(10) Epenthesis under Dorsey's Law (Alderete 2006[1995]: 32)

/hipres/	→	hiperés	'know'
/krepną/	→	kerepą́ną	'unit of ten'

Cases such as Winnebago have been reanalyzed as epenthesis that caters to syllable contact preferences (see Alderete 2006[1995]), which disprefer sonority rises across syllable boundaries. Alternative analyses of the Winnebago case are available (e.g. see Hall 2003 for a more phonetically-based analysis), but all analyses make reference to syllable structure. The boundary between syllabically-conditioned epenthesis and syllable-contact epenthesis might not be rigid and depend on analysis, but in any case that distinction is not relevant for the typological point I am making here; what is crucial is that epenthesis acts at the level of syllable structure.

Another case of syllable contact-conditioned epenthesis is Catalan, where rises in sonority across syllable boundaries are resolved in a variety of ways depending on the particular nature of the segments; if the cluster involves a nasal or *s* followed by a rhotic, epenthesis is used. Data below is from Pons Moll 2005: 9; the cluster in the underlying form and its output with epenthesis are underlined.

(11)

a.	/temɾiə/	→	təməɾíə	'(I) would be afraid'
	/temɾe/	→	təməɾé	'(I) will be afraid'
b.	/plaɲɾiə/	→	pləɲəɾíə	'(s/he) would complain'
	/plaɲɾa/	→	pləɲəɾá	'(s/he) will complain'
c.	/bɛnsɾiə/	→	bənsəɾíə	'(I) would win'
	/bɛnsɾe/	→	bənsəɾé	'(I) will win'

The typological claim made here is, thus, the traditional one: that epenthesis is universally used to resolve syllable structure markedness that has to with sonority sequencing, syllable contact, complexity of syllable margins, etc. Among metrical factors, only one may play a role in epenthesis, viz. word minimality. No other constraint can force vowel insertion. As I will show in this section, epenthesis cannot be used as a repair strategy for violations of purely metrical constraints such as *CLASH, *LAPSE, *NON-FINALITY, and so forth. Using the machinery discussed above in section 4.0.1, the generalization is equivalent to saying that in all cases where epenthesis applies, the winner of the alternative grammar without epenthesis contains either a marked consonant cluster, or is a subminimal word.

4.1.2 Non-observed epenthesis: metrical constraints

Apart from word minimality, other metrical factors are not observed to play a role in vowel epenthesis. There appear to be no examples in the typology of vowel epenthesis as a response to stress markedness constraints. In this section I show that the constraints NON-FIN, *CLASH, and *LAPSE can all potentially force the insertion of vowels, but these patterns are unattested. To make this clearer, I will construct some hypothetical examples of epenthesis driven by such stress markedness constraints.

Violations of NON-FIN, which penalizes candidates with a stress on the final syllable, can clearly be obviated by inserting material the intervenes between the stressed syllable and the word boundary. This would be simplest to observe in a language with lexical stress, where high-ranking MAX-PROM prevents any stress shift from its underlying position. If in such a language NON-FIN is ranked high enough – above DEP-V, to be exact – nothing would prevent the forms with underlying stressed final syllables from emerging with epenthetic final vowels serving as a buffer between the stressed syllable and the word edge. Consider the hypothetical inputs /pátak/ and /paták/: under the ranking MAXPROM, NON-FIN ⟩⟩ DEP-V, we predict final epenthesis in the latter case (/paták/ → *patákə*), but not in the former case (/pátak/ → *pátak*). This is illustrated by the tableau below.

(12)

			MAXPROM	NON-FIN	DEP-V
/pátak/	☞	pátak			
		pátakə			*!
/paták/		paták		*!	
	☞	patákə			*
		pátak	*!		

In a slightly more complex but equally reasonable hypothetical language, vowel epenthesis at word edge to satisfy NON-FINALITY could occur in a system with predictable stress, such as Mekkan Arabic, where stress falls on final superheavy syllables. In such a language, *ə* epenthesis would take place just in face the final syllable is superheavy, and therefore stressed, in order to satisfy NON-FIN. To flesh this out, consider a language with Mekkan Arabic's stress system: final if superheavy, else penult if heavy, else antepenult. If at least NON-FIN and SWP outrank DEP-V, such a language would have epenthesis after final superheavy syllables, in order to prevent them from being stressed, as illustrated in the following tableau.[24]

(13)

			NONFIN	SWP	DEP-V
/patak/		pa(ták)	*!		
	☞	(pátak)		*	
		(páta)kə		*	*!
		pa(tákə)		*	*!
/patāk/		pa(tā́k)	*!		
		(pátāk)		*!	
		(pátā)kə		*!	*
	☞	pa(tā́)kə			*

What is problematic with such an epenthesis process is that vowel insertion is a response to a stress constraint (NON-FIN), not syllable structure constraints. The differing behavior of inputs like /patak/ and /patāk/ shows this: in the language described in

[24] Constraints penalizing other alternative repairs, such as MAX-μ, which prevents vowel shortening to get rid of the final superheavy syllable, must be high-ranked.

(13), epenthesis cannot be a response to a syllable structure constraint like 'no final obstruent stop', because it fails to take place just in case stress is not in danger of falling on the final syllable.

Likewise, the constraint *CLASH is predicted to cause unattested epenthesis patterns. If a language has a potential stress clash – such as any language with left-to-right moraic trochees – nothing prevents the clash situation in such a language from being resolved by vowel epenthesis. This case is exactly analogous to the previous example, so no tableau should be necessary here: just in case *CLASH outranks the constraint DEP-V, a 'buffering' *ə* may be inserted to relieve the clash. However, typological evidence suggests that *CLASH can force only one of two processes: destressing or stress shift.

A slightly more surprising case of non-occurring vowel epenthesis involves *LAPSE. Any sequence of two unstressed syllables – that is, a *LAPSE violation – can be repaired in a number of ways: by shifting or redistributing stress more evenly, or by creating degenerate feet with clash (/σ́ σ σ σ́/ → [σ́ σ σ̲́ σ́]). However, another perfectly reasonable but unattested strategy would be to insert an extra syllable. In languages with obligatorily binary feet, such a response would yield a sequence of syllables that could be parsed with maximally unmarked feet, without violating *LAPSE or *CLASH. This is more concretely illustrated by the tableau below, with an English-like example. Suppose the constraint against unstressed heavy syllables (WSP) is ranked high enough that both the initial and penultimate syllables of a word like *abracadabra* must surface with a stress. The actual English form violates *LAPSE in the string *raca*; ranking DEP low-enough would cause this *LAPSE violation to be obviated by inserting a whole syllable [ʔə], allowing the form to be parsed by maximally unmarked feet. No such epenthesis pattern is attested in any language.

(14)

		WSP	*CLASH	*LAPSE	DEP
abracadabra	áb.ra.ca.dáb.ra			*!	
	áb.rá.ca.dáb.ra		*!		
	ab.rá.ca.dáb.ra	*!			
☞	áb.ra.cá.ʔə.dá.bra				*

In sum, OT predicts that the prosodic constraints NON-FIN, *CLASH, and *LAPSE can all potentially cause epenthesis. This is a typological claim about the environment of epenthesis, in effect denying the more traditional generalization that, apart from word minimality, metrical factors cannot cause the insertion of segments or whole syllables.

I will now turn to the application of my theory of constraint interpretation to these cases, showing that using the implicational version of the constraints NON-FIN, *CLASH, and *LAPSE will allow a better fit with the observed typology.

4.1.3 Too many solutions and procedural constraints: epenthesis

Let me now go through each of these hypothetical cases and show how my proposal handles the non-occurrence of epenthesis as a repair strategy for these constraints.

4.1.3.1. NON-FIN

The rationale of the non-finality constraint is to protect the final syllable of a word from some prosodic constituent being built over it. Because non-finality is routinely repaired across languages by moving the stress rather than by other conceivable means such as epenthesis, the statement of the constraint in implicational terms must express the asymmetry between the properties of 'being final' and 'being stressed': the former is non-negotiable, while the latter, as far as the constraint is concerned, can be modified. Therefore I propose the following version of NON-FIN.

(15) ✦NON-FIN
If *x* is a word-final syllable, *x* is not stressed.

As a consequence, because 'being final' is the antecedent property of this constraint, it would be able to force stress shifts but not any changes that leads to the syllable whose designated state is 'final' to becoming non-final. Indeed, taking this constraint out of the system, we would compute the default location of the final syllable. Any candidate where the correspondent of that designated final syllable is not final would get fatally ✦-penalized by the implicational version of the non-finality constraint. This is illustrated by the tableau below.

(16) DS of [tak]: final

			MAXPROM	✦NON-FIN	DEP-V
/pátak/	☞	pátak			
		pátakə		✦!	*
/paták/	☞	paták		*	
		patákə		✦	*!
		pátak	*!		

In both forms, the designated state of the syllable [tak] is final: taking the constraint out of the ranking, we get the winners [pátak] and [paták] for the two inputs. Thus any candidate with epenthesis after the final syllable – the candidates [pátakə] and [patákə] in the tableau above – would receive a ✦ violation, thus becoming a perpetual loser. The upshot is that, as desired, epenthesis is not a response to NON-FIN. On the other hand, stress shift is still possible as a response to this constraint, as illustrated below: reranking MAXPROM below ✦NON-FIN in (16) would produce exactly such a system, where the input /paták/ undergoes a stress shift to surface as /pátak/ as a response to ✦NON-FIN.

*4.1.3.2 *CLASH*

The bad predictions generated by the interaction of *CLASH with DEP can be seen in a variety of hypothetical situations. Here I will focus on only one of these, as the other ones have a similar solution. Crosslinguistically, repairs for *CLASH include destressing in clash and stress shifts. The former can be seen in e.g. Central Yupik, where a syllable standing between two stressed syllables loses its stress. In Chevak Yupik, any syllable following a stressed syllable loses its stress (Hayes 1995: 250). In fact, Hammond 1984 has suggested that destressing is the universal repair strategy for *CLASH. Therefore, the negotiable property of 'being stressed', just as in the case of the NON-FIN constraint, is in the consequent of the implicational ✦*CLASH, while the non-negotiable syllable adjacency is in the antecedent.

(17) ✦*CLASH
If σ_1 and σ_2 are adjacent syllables, they are not both stressed.

Consider how this version of the constraint prevents epenthesis from being used as a repair strategy for *CLASH violations. The tableau is shown below.

(18) Designated state of [pa] and [ta]: adjacent

			MAXPROM	✦*CLASH	DEP
/pátáka/		pátaka	*!		
		páʔətáka		✦	*!
	☞	pátáka		*	

The antecedent of the constraint mentions two syllables' adjacency status. This means that if the designated state of any pair of syllables with respect to this constraint is 'adjacent', then any candidates where those syllables are not adjacent would be penalized by the *CLASH constraint and thus not a possible winner. Indeed, the epenthesis candidate [páʔətáka] is just such a form: the designated status of the syllables [pa] and [ta] is 'adjacent', while in this candidate they are separated by the intervening epenthetic syllable [ʔə]. No such candidate can win, and thus epenthesis is not a possible repair strategy for *CLASH.

*4.1.3.3 *LAPSE*

The analysis of *LAPSE and its interaction with DEP works along the same lines as the *CLASH case. Here, too, the strategy is to express the asymmetry of repairs conditioned by this constraint in the asymmetrical statement of the constraint itself. The repairs for *LAPSE that are observed crosslinguistically involve stress shifts. Just as in the case of clash, I propose the statement of the constraint, given below, where the negotiable property of stressedness is in the consequent, while syllable adjacency is in the antecedent.

(19) ✦*LAPSE

If σ_1 and σ_2 are adjacent syllables, they are not both unstressed.

Consider the input /abracadabra/. This word has two light syllables separated by two heavy syllables, *ab.ra.ca.dab.ra*; this means that a weight-sensitive stress system runs the risk of violating *LAPSE, e.g. *áb.ra.ca.dáb.ra*. We are trying to rule out candidates that avoid such a *LAPSE violation by epenthesizing an extra syllable between the two heavies, e.g. *áb.ra.cá.ʔə.dáb.ra*. Adopting the ✦*LAPSE version given above allows us to rule out any repair that interferes with the adjacency status of syllables, including epenthesis.

(20)

			✦WSP	✦*LAPSE	✦*CLASH	DEP
abracadabra		áb.ra.ca.dáb.ra		*!		
	☞	áb.rá.ca.dáb.ra			*	
		ab.rá.ca.dáb.ra	*!			
		áb.ra.cá.ʔə.dáb.ra		✦!	✦	*

✦*LAPSE mentions syllable adjacency in the antecedent. Any pair of syllables which are not adjacent in some candidate but whose designated state with respect to ✦*LAPSE is 'adjacent' will render that candidate incapable of winning. It follows that epenthetic candidates must fail. For the input /abracadabra/ (see (14)), the candidate [ábracáʔədábra] violates ✦*LAPSE, because the syllables [ca] and [dab], whose designated state is 'adjacent', are separated in this candidate by another syllable.

Because ✦*CLASH mentions the same category in its antecedent as ✦*LAPSE, these two constraints belong to the same antecedent class, and the designated states for these constraints are computed simultaneously. This means that any candidate that ✦-violates ✦*LAPSE also ✦-violates ✦*CLASH.

4.1.4 Attested cases of epenthesis

Recall from section 4.1.1 above that observed cases of epenthesis fall into three categories: syllabically conditioned, minimality-conditioned, and sonority-conditioned. Sections 4.1.2 and 4.1.3 were devoted to plausible but typologically unattested metrically-conditioned epenthesis. Treating metrical constraints like NON-FIN, *CLASH and *LAPSE as procedural helps regulate the kinds of repairs these constraints can force, allowing stress shifts but ruling out segmental changes like epenthesis.

Thus the strategy in dealing with epenthesis is to regulate constraints not directly related to it. The attested cases of epenthesis, on the other hand, require no new machinery. If syllable structure, sonority-sequencing, and syllable contact constraints retain their standard OT output-oriented form, then epenthesis is automatically predicted as a repair for violations of those constraints, among other repairs. That is the strategy I will adopt.

In Chapter 3 above I discussed the creteria for deciding whether a constraint belongs to the procedural or the standard OT class. The two types of constraints say different things: the former regulate input-output mappings, while the latter express static surface pressures. The best way of separating the two classes is by observing the typology of repairs that the constraints force. On the one hand, there are constraints with a very limited set of effects: stress-segmental constraints like ✦ASPIRATE and ✦SWP are observed to condition only stress-sensitive modifications of aspiration and weight, respectively. On the other hand, there are constraints like NOCODA, ONSET, and *COMPLEX, with a rich attested set of effects both across languages and within a single language. Epenthesis, deletion and featural changes are all attested as repairs of syllable structure constraints; this wide range of repairs led de Lacy (2003) to call prosodic markedness constraints 'heavyweights'. Constraints on sonority profiles of syllables have a similarly broad set of effects, and include epenthesis, deletion of consonants, and featural changes.

Thus, as long as the metrical constraints like ✦NON-FINM, ✦*CLASH, and ✦*LAPSE are prevented from forcing unwanted epenthesis, nothing new must be added to the theory.

4.2 Syncope

The environment of syncope also presents a challenge for standard OT. This section is devoted to two goals: establishing the typological generalization on the environment of syncope, and accounting for this generalization using the formal proposals of this dissertation. I begin in section 4.2.1 with a survey of rule-based views of syncope, and show in section 4.2.2 how the predictions of OT differ with respect to syncope of stressed vowels. Unlike the older derivational theory, OT allows, in the general case, stressed vowels to syncopate as a response to the pressure of prosodic markedness constraints. I will argue that it is the more restrictive derivational view that is empirically correct, and show how my formal proposals can allow OT to account for the typology of the syncope environments. I will illustrate my argument with several case studies of syncope in section 4.2.3.

4.2.1 Rule-based views of syncope

Much pre-OT literature viewed metrical vowel syncope as vowel reduction taken to its limit. The environment of syncope, therefore, did not pose a problem: metrically-driven vowel deletion was thought to apply in the same environments as vowel weakening, i.e. in unstressed, unfooted, final, and posttonic positions. This view goes back to the long-held assumption that vowel syncope correlates with strong stress, expressed, among many others, by Bloomfield's statement that "[l]anguages with strong word-stress often weaken or lose their unstressed vowels" (1961[1933]: 382). The assumption that only unstressed vowels can syncopate had been considered obvious enough not to merit discussion, and was often used tacitly to argue for particular analyses of phonological systems. To take a well-known example, classical philology has devoted considerable attention to Latin syncope taken AS EVIDENCE that archaic Latin had a dynamic stress (cf. Vendryes 1902, Leumann 1977, Sihler 1995). In particular, syncope of vowels that would have been stressed by the classical Latin stress rule was seen to indicate that at the stage at which syncope applied the stress rule had been different[25] – an argument that clearly relies on the assumption that stressed vowel syncope would be nonsensical. This classical view was inherited by generative phonology, even though it once again seemed too self-evident to be questioned seriously.

Let me discuss here in more detail one theory of vowel syncope which did seek to explain, or at least express, the generalization that stressed vowel syncope is impossible. The theory of vowel deletion in Taylor 1994, couched in a constraint-and-repair framework, exemplifies the kind of mechanisms a derivational theory would need in order to prevent syncope from applying to stressed vowels. We will see that, although the central claim in Taylor's work is that conditions on environments of processes such as vowel deletion are best reformulated as conditions on the structure of outputs, the

[25] The often cited case is syncope in words like *balineum* > *balneum* 'bath', where the syncopating vowel would have been stressed by the Latin stress rule. Such cases are used as evidence to argue that the stress rule at the stage when syncope applied was different, and assigned stress to the initial syllable of *bálineum*.

type of output conditions Taylor envisions are very different from the surface markedness constraints of OT, and in fact amount to conditions on processes from the point of view of parallel OT.

Let me examine the part of Taylor's proposal that deals with the restriction on accented vowel syncope. In the constraint-and-repair framework, the general approach is to first apply a phonological rule, and then examine the output with respect to certain constraints. If the output produced by the rule violates a constraint, then a repair mechanism is invoked, which serves to obviate the violation. The application of this repair is the device that is intended to account for the absence of some pattern in the output. Vowel deletion, in Taylor's theory, consists of three separate stages: delinking of the vowel from the timing slot; deletion of that X slot; and deletion of the unassociated vocalic melody.

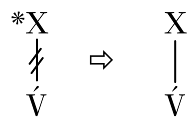
(21) Vowel deletion according to Taylor 1994
   a. Delinking of vowel from X slot
   b. Deletion of unlinked X slot
   c. Deletion of unlinked vocalic melody

At each of these stages, constraints on outputs can intervene to prevent vowel deletion from proceding all the way through, and repair strategies ensure that, in those cases, the vowel remains. Specifically, at the stage of vowel delinking, two types of constraints may have an effect on the derivation: constraints against delinked accented vowels, and constraints against delinked high-sonority vowels. The effect of these constraints is that unstressed vowels delete in preference to stressed vowels, and vowels lower on the sonority scale are more likely to be subject to deletion.

The following is the constraint used by Taylor to account for the failure of accented vowels to delete. It states that a delinked accented vowel is prohibited.

(22) *X (Taylor 1994: 14)
   ǂ
   V́

This constraint conditions a repair strategy, viz. relinking of the delinked accented vowel to the timing slot.

(23) 

This repair strategy is invoked as soon as the constraint (22) is violated, i.e. at the delinking stage of vowel deletion (21)a. Thus, once (23) has applied, stressed vowels never get a chance to delete.[26] In other words, Taylor's theory requires a Duke-of-York (A→B→A) derivation any time an accented vowel is involved: first the vowel is deleted by the general syncope rule, and then the relinking repair strategy puts the vowel back, in order to prevent a violation of the constraint against stressed delinked vowels.

Although the constraint (22), which ensures that stressed vowels fail to syncopate, is stated as an output condition, it is very different from an OT surface markedness constraint. Indeed, the representation to which the constraint refers is necessarily an intermediate one: it exists only after vowel delinking but before timing slot deletion and vocalic melody deletion. The constraint crucially refers to information that only exists at this intermediate stage of representation: the X slot and the vocalic melody to which the X slot is not linked. This intermediate stage preserves crucial information about the derivational history of the form, viz. the fact that a stressed vowel has been delinked from an X slot. Because Taylor's constraint necessarily causes a repair mechanism to apply that, in effect, undoes the action of the prior syncope rule, the condition is equivalent to a statement about DERIVATIONS: it simply prohibits syncope from applying in the environment of a stressed vowel.

---

[26] Also operative at this stage are constraints on vowel sonority, which invoke a similar relinking repair strategy when high-sonority vowels like [a] are deleted. An interesting question that falls outside of the present discussion is what accounts for the exceptionless application of the accentuation constraint (22) – stressed vowels NEVER delete – but language-particular application of the vowel quality constraints – [a]s sometimes do delete. Taylor appears to endorse these typological generalizations, but it is unclear how they follow from her account.

OT constraints, on the other hand, only have access to the output and the input, so Taylor's constraint is restatable in OT terms only in the special case where stress is present in the underlying form. In this special case, a positional faithfulness constraint requiring an output stressed vowel to have an input correspondent would effectively ensure that input stressed vowels are not deleted (as well as that stress is not shifted to a different syllable). As discussed above, a general property of OT is that environment-based generalizations are only statable when the environment is present in the underlying form; the situation discussed here is simply a special case of that fact. In the general case, however, Taylor's stressed vowel constraint cannot have a straightforward OT correspondent. Indeed, the derivational constraint (22) is not sensitive to whether the stress on the vowel comes from the underlying form or had been assigned by a prior rule. In the latter case, the crucial information necessary to prevent stressed vowels from syncopating is only present at the intermediate stage of the derivation, a stage not accessible to any OT constraint.

What is more, even in the special case where stress is present in the underlying form and a positional faithfulness constraint like MAX-σ́ is available, there is no general account of the typological IMPOSSIBILITY of stressed vowel deletion, because, like all constraints, MAX-σ́ would be violable. If the constraint forcing vowel deletion were to be ranked high enough, MAX-σ́ would fail to protect a stressed vowel from deletion, and stressed vowel syncope would be predicted to result.

The upshot of this discussion of one representative derivational theory of syncope is that if the empirical generalization is correct, then OT has no direct way of accounting for it. In pre-OT theories, the typological generalization that stressed vowels do not syncopate was not seriously questioned; syncope was standardly assumed to fit into the general pattern of vowel weakening, which ought by rights to apply only to weak vowels. Because OT has no way of dealing with the previously assumed typology of syncope, the strategy – just as the original assumption, left largely implicit – was to deny the generalization itself, a step that, I argue here, is incorrect. Instead, I will show that the generalization is in fact true and presents a serious and systematic problem for OT, and that the proposal on modifying the interpretation mechanism of OT constraints made in the previous chapter can be harnessed to account for it.

4.2.2 Syncope in OT

OT has quietly set aside the claim of previous theories that syncope does not apply to stressed vowels. I argue in this section that this step is a mistake. To show that stressed vowels indeed cannot syncopate, I will go through several OT analyses of languages where stressed vowel syncope was proposed, and show that alternative and better analyses are available that do not require the traditional typological generalization to be violated. I will also discuss cases where, apart from OT theorizing, the facts seem to indicate that stressed vowels have deleted, and will suggest ways of interpreting the data in a way that is consistent with the proposals made here.

Before buttressing the typological generalization with these arguments, I will propose an analysis using the proposals on constraint interpretation that are the subject of this thesis.

The typological claim here, to be supported in the remainder of the chapter, is that syncope applies in WEAK positions: posttonically, in unfooted syllables, in the weak branches of feet, and in the final position (apocope). Possibly, pretonic syncope and initial vowel deletion (aphaeresis) must also be added to the list (see Taylor 1994). All of these environments fall in the broader category of 'unstressed syllable'; some languages show syncope in an environment that cross-cuts these more narrow contexts, i.e. in all unstressed syllables. These environments are listed below.

(24)
a. Posttonic (Hopi, Southeastern Tepehuan)
b. Unfooted (archaic Latin, Tonkawa)
c. Weak branch of foot (Carib)
d. Word-final, unstressed (apocope) (Lardil)
e. Unstressed (Lebanese Arabic, Iraqi Arabic)

The goal of an OT analysis is to ensure that syncope-driving constraints force syncope only in these environments, but fail to cause stressed vowel deletion. Because deleting a stressed vowel would necessarily involve a stress shift, to account for the typology of syncope it would be sufficient to ensure that syncope constraints cannot force a stress shift; the proposal in the previous chapter is well-equipped to account for such a generalization. Recall that the strategy for ensuring that a given constraint does not force a stress shift is to put the category 'stressed syllable' in the antecedent of a procedural constraint. Then, given the proposed mechanism of interpreting procedural constraints, it would follow that the location of stress cannot be affected by such a constraint.

Thus, our syncope constraints must mention the environment, in terms of the location of stress, in the antecedent. What is in the consequent part of the constraint is slightly more problematic. The goal is to limit the effect of the syncope constraint to vowel deletion. Thus, whatever property is mentioned in the consequent of such a constraint must hold if and only if a vowel has deleted. The constraint must be able to "see" a trace of the deleted vowel in order to produce the desired typological prediction.

I will therefore adopt a representational assumption to ensure that such a trace is available to the constraint. The idea is that when a vowel deletes, it leaves an X-slot behind it, a featureless, empty syllable nucleus. This is not a new proposal; Kager's (1997) study of Carib and several other languages was aimed at establishing exactly this fact, viz. that in certain cases the syncopating vowel leaves an empty slot behind, which can be visible through its interaction with other processes such as assimilation.

Given this representational assumption about a trace left behind by the syncopating vowel, procedural syncope constraints can be formulated in a straightforward way.

(25) a. ✦SYNCOPE(POSTTONIC) 'If a nucleus is posttonic, it is empty'
b. ✦SYNCOPE(UNFOOTED) 'If a nucleus is in an unfooted syllable, it is empty'
c. ✦SYNCOPE(WEAK) 'If a nucleus is in a weak branch of a foot, it is empty'[27]

Given the procedural constraint interpretation mechanism we have been pursuing, these constraints can lead to syncope ONLY in the stated environment. Because the location of a vowel with respect to stress is mentioned in the antecedent of the constraint, no stress shift can result from its action. A fortiori, stressed vowels cannot be deleted – the removal of a stressed vowel would necessarily result in a stress shift. In the following hypothetical example, stress is assumed to be penultimate by default; the syncope constraint I use in this example is POSTTONIC. Empty nuclei (i.e. sites of deletion) are marked with the sign •.

(26) /$CV_1CV_2CV_3$/
Designated location of the posttonic nucleus: V_3

			✦SYNCOPE (POSTTONIC)	STRESS	MAX-V
	a.	$CV_1C\acute{V}_2CV_3$	*!		
☞	b.	$CV_1C\acute{V}_2C\,•_3$			*
	c.	$C\acute{V}_1C•_2CV_3$	✦!	*	*

The first two candidates, (26)a and (26)b, show default penultimate stress with and without postonic syncope, respectively. The candidate (26)c violates the constraint ✦SYNCOPE because the location of the posttonic nucleus is different with respect to its designated location. Indeed, in the grammar with ✦SYNCOPE removed, stress falls on V_2, and hence the nucleus V_3 is postonic. In the candidate (26)c stress falls on V_1 and V_2 is posttonic; this ensures that candidate's fatal ✦ violation. In effect, this violation is due to the fact that a stressed vowel has syncopated in this candidate, causing a stress shift

[27] The problem with formulating the constraints in negative terms, e.g. "If a nucleus is posttonic, it is not followed by an open syllable with a short vowel", is that this constraint could potentially cause repairs other than syncope: lengthening, coda insertion, epenthesis, etc., would all be possible. None of these is an attested process applying in the WEAK environments where syncope is typical.

and thus a difference in the location of the posttonic nucleus with respect to its designated location.

An important consequence of this proposal is that stressed vowel deletion is only impossible under METRICAL pressures, i.e. constraints such as ✦SWP. Crucially, other constraints may still force stressed vowel deletion, a fact which, I argue, is consistent with the typology. In particular, the constraint ONSET may force the deletion of vowels in hiatus. In many cases (e.g. Tonkawa; see section 4.2.3.2 below), it is always the FIRST vowel in hiatus that deletes, regardless of its stressed/unstressed status. On my proposal, this is a consequence of the fact that ONSET is not a procedural constraint, and may condition a range of repairs as a standard OT markedness constraint.

This discussion of the analysis sets the stage for the survey of the typology of syncope. The following case studies in a number of languages are divided into three groups: first, I deal with languages that conform with my generalization (Carib, Tonkawa, Lebanese Arabic, Old Russian, Latin); second, I consider several languages where recent analyses have made use of stressed vowel syncope, and show that in each case alternative analyses are available and possibly superior (Southeastern Tepehuan, Hopi, Central Alaskan Yupik). Finally, I briefly discuss a case where the vowel deletion processes applies to stressed vowels, and argue that in these cases the deletions do not fall under the scope of metrically-driven syncope (Bedouin Hijazi Arabic).

4.2.3 Case studies

4.2.3.1 Carib

My first illustration of a well-behaved vowel syncope process comes from Carib (Hawkins 1950; Kager 1997), an iambic language where weak vowels are deleted. In words consisting of only light syllables, iambic feet are built left-to-right, and unstressed vowels are either deleted or reduced to a consonantal release, under conditions I will not discuss here. Stressed syllables are lengthened. In odd-parity words, the vowel in the final syllable is lengthened and the syllable is stressed. Main stress falls on the rightmost foot. In the following examples, beneath each form is its schematic representation where the deleted syllables are crossed out (σ̸).

(27)	a. /pata/	(p^{ə}tá:)	'place'
	/σ σ/	(σ̸ σ́:)	
	b. /piripi/	(prì:)(pí:)	'spindle'
	/σ σ σ/	(σ̸ σ̀:)(σ́:)	
	c. /erepamɨ/	(^{ə}rè:)(pmɨ́:)	'I arrive'
	/σ σ σ σ/	(σ̸ σ̀:)(σ̸ σ́:)	
	d. /umanarirɨ/	(mà:)(nrì:)(rɨ́:)	'my cassava grater'
	/σ σ σ σ σ/	(σ̸ σ̀:)(σ̸ σ̀:)(σ́:)	

Heavy syllables disrupt the count relevant to stress and syncope: they bear stress, and the footing starts anew. Both CVV and CVC count as heavy.

(28)	a. /seepɔrɔ/	(sèe)(prɔ́:)	'along here'
	/σ: σ σ/	(σ̀:)(σ̸ σ́:)	
	b. /peʔmara/	(pèʔ)(mrá:)	'free'
	/σ: σ σ/	(σ̀:)(σ̸ σ́:)	
	c. /eerepamɨ/	(èe)(r^{ə}pà:)(mɨ́:)	'I arrive'
	/σ: σ σ σ/	(σ̀:)(σ̸ σ̀:)(σ́:)	
	d. /uyenkuʔtɨsaʔya/	(yè:)(kùʔ)(t^{ə}sàʔ)(yá:)	'if anyone deceives me'
	/σ σ: σ: σ σ: σ/	(σ̸ σ̀:)(σ̀:)(σ̸ σ̀:)(σ́:)	

The stress system in Carib is straightforwardly iambic: the only possible feet are (LH) and (H), assigned left-to-right. Degenerate feet (L) are prohibited, which accounts for the lengthening and final stress in odd-parity words like /piripi/→ (prì:)(pí:). In the weak position of the iambic foot heavy syllables are prohibited, resulting in the disruption of the left-to-right count whenever a heavy syllable occurs in an odd-numbered position, as in (28). Syncope or weakening applies to any vowel in the weak position of the (LH) foot. Thus, the data in (27)-(28) show that Carib syncope follows the predicted pattern and can be straightforwardly accounted for the constraint ✦SYNCOPE(WEAK).

There is one interesting complication to this pattern. Deletion does not apply when the vowel to be deleted is adjacent to a consonant cluster. This prohibition has to do with a well-motivated preference to avoid complex consonant sequences, expressed by a constraint like *CCC. However, not only does the vowel fail to delete in such circumstances, but the footing itself appears to be dependent on the outcome of reduction: the vowel that fails to undergo syncope surfaces as stressed, and the foot count restarts to its right.[28] If the vowel simply failed to reduce or delete in the form /kratupe/ (see below in (29)b), we would expect the footing to simply place it into the weak position of the iambic foot, resulting in the incorrect *(kratù:)(pé:). Instead, the actual form is (krà:)(t^{ə}pé:): the vowel in the first syllable, which otherwise would be subject to deletion, ends up with a stress. This stress attraction cannot easily be attributed to weight effects, because the consonant cluster that causes it is in the ONSET of the initial syllable. As Kager (1997: 473) puts it, "[a]n ordering paradox is evident: vowel reduction depends on metrical parsing, while metrical parsing must in turn depend on reduction".

(29)	a. /ptakaype/	(ptà:)(kày)(pé:)	'traira-fish now…'
		*(pt$^{(ə)}$kày)(pé:)	
		*(ptakày)(pé:)	
	b. /kratupe/	(krà:)(t^{ə}pé:)	'alligator now…'
		*(kr$^{(ə)}$tù:)(pé:)	
		*(kratù:)(pé:)	

These data present an apparent problem for my account. The implicational version of ✦SYNCOPE(WEAK) cannot derive the correct output (krà:)(t^{ə}pe) in this case, because this constraint in general can force only vowel deletion but not a stress shift. In other words, the only possible outputs for the input /kratupe/ are *(*kratù*:)(*pé:*) and *(*krtù:*)(*pé:*), depending on the ranking of *CCC. If *CCC is high-ranked, syncope simply fails and we get the form (*kratù*:)(*pé:*), where the vowel in the first syllable is

[28] A simple failure of vowel deletion following a cluster would not be unusual: it would amount to the blocking of syncope by a cluster condition. A similar situation exists in French: schwas delete unless a

protected by the consonant cluster (30). If *CCC is low-ranked, syncope applies and creates the marked consonant cluster in (*krtù:*)(*pé:*), as shown in (31).

(30) Designated footing: (σ σ̀)(σ́)

			*CCC	✦SYNCOPE	MAX-V	AL-R
/kratupe/	(☞)	(krà:)(t^{ə}pé:)		✦		*!
		(krtù:)(pé:)	*!		*	
	☹	(kratù:)(pé:)		*		

(31) Designated footing: (σ σ̀)(σ́)

			✦SYNCOPE	MAX-V	AL-R	*CCC
/kratupe/	(☞)	(krà:)(t^{ə}pé:)	✦!		*	
	☹	(krtù:)(pé:)		*		*
		(kratù:)(pé:)	*!			

Clearly, the actual candidate (*krà*:)(*t^{ə}pé*:) cannot win under these circumstances, because it incurs a fatal ✦ violation of the ✦SYNCOPE constraint.

There must be another constraint that would prefers the correct output (*krà:*)(*t^{ə}pé:*) over the incorrect *(*kratù:*)(*pé:*) in a way that affects the selection of the designated footing. One possibility is to invoke the set of constraints responsible for sonority-driven stress, which prohibit reduced vowels from being stressed. The series of constraints *NONHEAD/{a}, *NONHEAD/{a,e,o}, *NONHEAD/{a,e,o,i,u}, etc., as discussed in the preceding chapter, have stress not in their antecedent but in the consequent. This means that the sonority/stress constraint can affect the designated site of stress and as a result can force stress shifts.

(32) a. *NONHEAD/{a}
'If a syllable has the nucleus *a*, then it is stressed'
b. *NONHEAD/{a,e,o}
'If a syllable has the nucleus *a*, *e*, or *o*, then it is stressed'
c. *NONHEAD/{a,e,o,i,u}

---

consonant cluster would form. Cf. *événement* [evɛnmɑ̃] 'event' vs. *gredin* [grədɛ̃] 'rogue'.

'If a syllable has the nucleus *a*, *e*, *o*, *i*, or *u*, then it is stressed'

In other words, the solution to this puzzle is that Carib has, on top of the standard left-to-right iambic stress system, a sonority-sensitive restriction that prohibits full vowels from appearing in the weak branches of feet and causes stress attraction to full vowels. Although Kager's set of examples only contain cases where the exceptionally stressed vowel is *a*, it is safe to assume that all vowels except *ə* behave in this way, in absence of evidence to the contrary. This means the constraint at play here is *NONHEAD/{a,e,o,i,u}, which I will abbreviate as *NONHEAD(fullV). Because ✦SYNCOPE outranks ✦NONHEAD(fullV), syncope takes precedence, and the sonority-sensitive effect only emerges when syncope cannot apply due to other factors (*CCC).

(33) ✦SYNCOPE ≻ ✦*NONHEAD(fullV)
Designated state of footing: (σ σ́)
Designated state of initial vowel: [ᵊ]

			✦SYNCOPE	✦*NONHEAD(fullV)	*CCC	*CLASH
/pata/	☞	(pᵊtá:)				
		(patá:)	*!	*		
		(pà:)(tá:)	✦!	✦		*

(34) ✦SYNCOPE ≻ ✦*NONHEAD(fullV)
Designated state of footing: (σ̀)(σ σ́)
Designated state of initial vowel: [a]

			✦SYNCOPE	✦*NONHEAD(fullV)	*CCC	*CLASH
/kratupe/	☞	(krà:)(tᵊpé:)				*
		(krtù:)(pé:)	✦!	✦	*	*
		(kratù:)(pé:)	*!	*		*

Thus, Carib is a case of well-behaved syncope, once the additional complication is taken into account that, on top of the iambic system, the language has a minor sonority-sensitive stress 'subsystem'.

4.2.3.2 Tonkawa

Tonkawa syncope accords with the generalization that only unstressed vowels may undergo metrically-driven syncope. Interestingly, Tonkawa has not only metrical syncope, but also vowel deletion in hiatus, and the latter process indeed MAY apply to stressed vowels, as predicted by my proposal.

Stress in Tonkawa, as implied by Gouskova (2003: 124 and ff.) is based on the moraic trochee assigned left-to-right. Although the original description of Hoijer (1938; cf. also Kisseberth 1970) does not supply information about stress, it can be inferred from processes that depend on it, as Gouskova points out. Because of the circular reasoning that this uncertainty brings, Tonkawa has only a limited value for the point at hand, but the example is, I believe, still illustrative as a typical case of well-behaved syncope. It should be taken with a grain of salt.

Characteristically of trochaic systems, vowels shorten in iambic word-initial configurations: long vowels after initial L surface as short (35)a-b, but remain unchanged after an initial H (35)c. This suggests that a canonical moraic trochee, (H) or (LL), is enforced at the left edge, at the expense of unfaithfully mapping underlying LH sequences.

(35)	a.	/xa-kaana-oʔ/	(xaka)(noʔ)	'he throws it far away'
		/ke-yaaloona-oʔ/	(keya)(loo)(noʔ)	'he kills me'
		/we-naate-oʔ/	(wena)(toʔ)	'he steps on them'
	b.	/ke-soopka-oʔ/	(kesop)(koʔ)	'I swell up'
		/we-cˀaapxe-oʔ/	(wecˀap)(hoʔ)	'he puts up several beds'
	c.	/nes-kaana-oʔ/	(nes)(kaa)(noʔ)	'he causes him to throw it away'
		/yaaloona-oʔ/	(yaa)(loo)(noʔ)	'he kills him'
		/ʔatsoo-k-laknoʔo/	(ʔat)(sook)(lakno)(ʔo)	'came to life, it is said'

The data furthermore suggest that word-initial CVC syllables count as heavy, while word-internal CVC syllables are light: although CVC is acceptable when following an initial L, as shown in (35)b above, initial CVC syllables do not cause shortening of the following vowel, as can be seen in (35)c. This sensitivity of CVC weight to the position of the syllable in the word is attested elsewhere, and is a well-known feature of several Yupik languages.

Tonkawa has at least two vowel deletion processes, syncope and elision. Syncope applies to vowels in the weak branch of a foot, as seen in the examples in (36) below. An important restriction is that root-final vowels may not undergo syncope. Thus, the second vowel in /ke-yakapa-nes-ʔoʔ/ → (*key*)(*kapa*)(*nesʔoʔ*) deletes, but the fourth root-final vowel is protected from syncope. The second type of vowel deletion is elision, which automatically eliminates vowels before other vowels in hiatus. Both stressed and unstressed vowels can elide, nor are root-final vowels protected from elision.

In the data below, vowels that syncopate are crossed out in the middle column, and surface forms are shown in the rightmost column. Root-final vowels are underlined. (36)a shows weak vowel syncope, (36)b shows elision of unstressed vowels, (36)c elision of stressed vowels, and (36)d gives some additional examples of both elision and syncope, where the latter applies in a foot other than the initial foot in the word.

(36) a.	/yamaxa-noʔ/	(yam~~a~~)(xanoʔ)	(yam)(xanoʔ)	'he is painting his face'
	/ke-yakapa-nes-ʔoʔ/	(key~~a~~)(kapa)(nesʔoʔ)	(key)(kapa)(nesʔoʔ)	'they two strike me'
b.	/we-yakapa-oʔ/	(wey~~a~~)(kap~~a~~oʔ)	(wey)(kapoʔ)	'he hits them'
	/ke-yamaxa-oʔ/	(key~~a~~)(max~~a~~oʔ)	(key)(maxoʔ)	'he paints my face'
c.	/yakapa-oʔ/	(yak~~a~~)(p~~a~~oʔ)	(yak)(poʔ)	'he hits it'
	/ke-we-yamaxa-oo-ka/	(kew~~e~~)(yam~~a~~)(x~~a~~oo)ka	(kew)(yam)(xoo)ka	'you paint your faces'
d.	/nes-yamaxa-oʔ/	(nes)(yam~~a~~)(x~~a~~oʔ)	(nes)(yam)(xoʔ)	'he causes him to paint his face'
	/taa-notoso-oʔs/	(taa)(not~~o~~)(s~~o~~oʔs)	(taa)(not)(soʔs)	'I stand with him'

Crucial for our purposes here is the difference in sensitivity to prosodic structure between the two vowel deletion processes: syncope applies only to vowels in the weak branch of a foot, while elision mechanically deletes vowels in hiatus regardless of the prosodic context.

4.2.3.3 Lebanese Arabic

In this section I will go through an example of well-behaved metrically-driven syncope that applies to unstressed vowels, and illustrate the shortcomings of an analysis that takes syncope to be driven by constraints such as SWP and Parse-σ. The problem with this analysis, as I will show, is not so much empirical as explanatory: deletion of unstressed vowels is treated as a collection of seemingly unrelated processes, and it appears to be an accident that all of them coexist in the same language. However, a crucial generalization about the syncope's ENVIRONMENT is lost; nothing in the standard OT analysis expresses the simple fact that it is the unstressed vowels that delete. In other words, the Lebanese Arabic situation presents a non-surface-based conspiracy from the point of view of OT. Moreover, the unexplanatory nature of the analysis of the Arabic data is symptomatic of a broader typological problem, viz. the failure of the theory to express the crosslinguistic connection between stress and sycnope.

Lebanese Arabic has a stress generalization typical of many Arabic dialects: the final syllable is stressed if it is superheavy, otherwise stress falls on the penult if it is heavy, and, if neither of these two conditions is met, the antepenult is stressed by default, as illustrated by the following representative data (Gouskova 2003).

(37) a. Stress final superheavy

ʔakált	'I ate'
nazzált	'I brought down'
saʔalúuk	'they asked you'
maktabáat	'libraries'

b. Else penult if heavy, and in disyllables

názzal	'he brought down'
náazal	'he encountered'
maʕáarik	'battles'
maktábti	'my library'
sáħab	'he withdrew'
ʔákal	'he ate'

c. Else antepenult

ʔákalit	'she ate'
máktabe	'library'

Some high vowels /i, u/ in open syllables are deleted under certain conditions, and the job of any analysis of syncope is to provide an account for which of the vowels syncopate. The environment-based generalization is straightforward: a high vowel in an open syllable is deleted whenever it would otherwise be unstressed. In the following examples the syncopating vowel is underlined. The middle column shows the application of the stress rule to the form without syncope, illustrating that the vowel that is deleted is always the one that is unstressed.

(38)

UR	STRESS	SF	GLOSS
/nizi̲lit/	ní.zi̲.lit	nízlit	'she descended'
/saaħibi̲tuu/	saa.ħí.bi̲.tuu	saaħíbtu	'his friend'
/saaħi̲bitna/	saa.ħi̲.bít.na	saaħbítna	'our friend'
/ʔibi̲nii/	ʔí.bi̲.nii	ʔíbnii	'my son'
/bagi̲lii/	bá.gi̲.lii	bágli	'my mule'
/ni̲zilt/	ni̲.zílt	nzílt	'I descended'
/fi̲himna/	fi̲.hím.na	fhímna	'we understood'

These facts are quite straightforward in terms of the proposal made here about ✦SYNCOPE constraints: syncope targets unstressed high vowels in open syllables, and hence results from the action of the constraint ✦SYNCOPE(UNSTR).[29]

(39) Designated location of stress: *ni*

	✦SYNCOPE	MAX-V

[29] I abstract away from the selective syncope of high vowels *i, u*; this is not crucial for the purposes of the discussion.

		(UNSTR)	
/nizilit/	nízilit	*!	
☞	níz•lit		*
	n•zílit	✦!	*
	nízil•t	*!	*

The designated locus of stress is on the initial syllable *ni*. The first candidate, *nízilit*, has no syncope at all, so the ✦SYNCOPE constraint is violated. The next candidate does not violate the constraint, because it has no unstressed non-empty nucleus in an open syllable. The candidate that deletes the stressed vowel, *nzílit*, fatally ✦-violates the syncope constraint, because it involves a stress shift with respect to the designated locus of stress. The final candidate *nízilt* once again has an *i* nucleus in an unstressed open syllable, and thus violates the ✦SYNCOPE constraint.

Note that there is no complex interaction between syncope and any other constraints in the grammar: the ✦SYNCOPE(UNSTR) constraint alone can decide which vowels will syncopate. This might be taken as a symptom of a bad analysis or of incorrect constraints in canonical OT. However, I suggest here that this usurpation by one constraint of the entire process is in fact a welcome development, because the syncope process has ONLY ONE reason to apply: it simply targets unstressed vowels. The constraint ✦SYNCOPE(UNSTR) straightforwardly captures that generalization, both within the given language, and typologically.

It is instructive to compare what drives syncope on Gouskova's analysis and mine. The simple generalization stated above – that the vowel that syncopates is the one that would end up unstressed – does not enter directly into Gouskova's analysis. Syncope's nature as extreme weakening is obscured by the disparate foot structure constraints that end up driving syncope. On Gouskova's view, there is no constraint specific to syncope; instead, the general constraint against marked nuclei, *NUC/i,u, forces the deletion of high vowels in general, while other prosodic constraints determine which high vowels end up syncopated. Nothing in the analysis links lack of stress to syncope, because the theory is not designed to handle environment-based generalizations. Indeed, this connection between weak prosodic positions and syncopation has all the characteristics of a phonological conspiracy.

Consider disyllables. By the stress rule, the only situation when a disyllable can receive final stress is if the final syllable is superheavy; otherwise, stress will fall on the penult. Thus, *ʔakált* 'I ate' has final stress but *ʔákal* is stressed on the penult. It follows that, on the view that syncope applies to unstressed syllables, it only will apply to words like /nizilt/ 'I descended', which have a superheavy ultima, but not to /nizil/ 'he descended', with a heavy final. Indeed, we have *nzílt* but *nízil*, not **nzíl*.

On the *NUC/i,u analysis, what prevents syncope of the initial stressed syllable in *nízil* to **nzíl* (or, for that matter, the syncope of *i* in the closed second syllable to give **nízl*) is NON-FINALITY: in the syncopated forms the stress is on the final (and only) syllable, and this is taken as the reason why syncope fails to apply. This is captured by the ranking NONFIN(STRESS) ⟩⟩ *NUC/i,u. On the other hand, the form with an underlying superheavy syllable like /nizilt/ does syncopate because $WSP_{\mu\mu\mu}$, the constraint against unstressed trimoraic syllables, is ranked above NONFIN. The choice of the syncopating vowel in the trisyllable /nizilit/ is also made by NONFIN.

On the other hand, in forms like /saaħibituu/, the choice of the syncopating vowel is determined by another constraint, PARSE-σ: the optimal form (*saa*)(*ħíb*)*tu* parses more syllables into feet than the loser *(*sáaħ*)*bitu*. These examples are illustrated below.

(40) Lebanese Arabic syncope (Gouskova 2003: 238)

			NONFIN (STRESS)	NONFIN(FT)	*NUC/i,u	PARSE-σ
/nizilit/	☞	(níz)lit			**	*
		(nízi)lit			***!	*
		(nzílit)		*!	**	
		ni(zílt)	*!	*	**	*
		(nzílt)	*!	*	*	
/nizilt/	☞	(nzílt)	*	*	*	

		ni(zílt)	*	*	**!	*
/nizil/	☞	(nízil)		*	**	
		(nízl)	*!	*	*	
/saaħibituu/	☞	(saa)(ħíb)tu			*	*
		(sáaħ)bitu			*	**!

The form /fihimna/ presents still another case. Here, the choice between the correct (*fhím*)*na* and the incorrect candidate that syncopates the stressed vowel, *(*fíhm*)*na*, is actually not made by Gouskova's constraints, but a reasonable addition to the system, such as a constraint that bans trimoraic syllables (*μμμ), or a constraint against triconsonantal clusters (*CCC).

(41)

			*NUC/i,u	PARSE-σ	*CCC?
/fihimna/	☞	(fhím)na	*	*	
		fi(hím)na	**!	**	
		(fíhm)na	*	*	*!

Because the constraints responsible for the choice of the syncopating vowel in the various forms are freely ranked with respect to each other and the other constraints, the upshot of Gouskova's analysis is that the application of syncope in any one class of forms is in principle independent of its application in another class. There could be languages where PARSE-σ is ranked low enough that syncope applies to the stressed vowel in /saaħ ibituu/, but in /nizilt/ is still the unstressed vowel that syncopates. In other words, the generalization that syncope applies to whichever vowel would otherwise be unstressed has no place in the theory. This is a case of a phonological conspiracy: several constraints appear all to work toward the same goal of ensuring that only UNSTRESSED vowels syncopate.

Let me make this claim more precise. The ostensible virtue of analyses such as Gouskova's is that they link the location of syncope to the constraints that are responsible for the location of stress in the language: the same constraints that are necessary for stress placement, e.g. PARSE-σ, NONFIN, etc. are also responsible for the stress system and syncope.

In fact, a more detailed examination of the factorial typology shows that this claim is simply incorrect. Using the OT Soft program (Hayes et al. 2003), I have constructed a factorial typology from a subset of the Lebanese Arabic data. The inputs included all relevant forms for the stress system and two potentially syncopating forms: /nizilit/ and /fihimna/, where the correct outputs are *nízlit* and *fhímna*. For these two inputs I have included the following sets of candidates.

(42) a. /nizilit/
nízilit (no syncope)
nízlit (correct syncope)
nzílit (wrong syncope)
nízilt (wrong syncope)
nzílt (wrong syncope)

b. /fihimna/
fihímna (no syncope)
fhímna (correct syncope)

Overall, the system generates 155 distinct output patterns. Among these, I have selected only those that have a stress system that is surface-identical to the actual system of Lebanese Arabic. Within such systems, six different syncope patterns are possible, not counting the actually attested one. These are listed below, together with the rankings that produce them.

(43) a. *nízlit* (correct syncope)
fihímna (no syncope)
WSP(3), *CCC ⟩⟩ NONFIN(St) ⟩⟩ FTBIN, WSP, SWP ⟩⟩ Max-V ⟩⟩ *NUC/i,u, PARSE-σ ⟩⟩ NONFIN(Ft) ⟩⟩ ALFTR

b. *nízilit* (no syncope)

fhímna (correct syncope)

WSP(3), *CCC ⟩⟩ NONFIN(St) ⟩⟩ FTBIN, WSP ⟩⟩ PARSE-σ ⟩⟩ NONFIN(Ft), MAX-V ⟩⟩ *NUC/i,u, ALFTR, SWP

c. *nzílit* (wrong syncope)
fihímna (no syncope)

WSP(3), *CCC ⟩⟩ NONFIN(St), PARSE-σ ⟩⟩ MAX-V, WSP, SWP ⟩⟩ NONFIN(Ft), *NUC/i,u, FTBIN ⟩⟩ ALFTR

d. *nzílt* (wrong syncope)
fihímna (no syncope)

WSP(3), *CCC ⟩⟩ PARSE-σ ⟩⟩ WSP, SWP ⟩⟩ NONFIN(St), MAX-V ⟩⟩ NONFIN(Ft), *NUC/i,u, FTBIN ⟩⟩ ALFTR

e. *nzílt* (wrong syncope)
fhímna (correct syncope)

*NUC/i,u, WSP(3), *CCC ⟩⟩ NONFIN(St), MAX-V ⟩⟩ FTBIN, WSP, SWP ⟩⟩ PARSE-σ ⟩⟩ NONFIN(Ft) ⟩⟩ ALFTR

f. *nzílit* (wrong syncope)
fhímna (correct syncope)

WSP(3), *CCC ⟩⟩ NonFin(St), Parse-σ ⟩⟩ *Nuc/i,u, FtBin, WSP, SWP ⟩⟩ NonFin(Ft), Max-V ⟩⟩ AlFtR

The fact that the factorial typology contains languages that have a surface-identical stress system with that of Lebanese Arabic and a DIFFERENT syncope pattern shows that, in fact, the standard OT analysis fails to make a connection between stress and syncope. Two of the predicted systems show unstressed vowel syncope in some forms but not others (43)a-b. Four systems show stressed vowel syncope at least in some forms (43)c-f.

The diagnosis of the problem is clear. The syncope-driving constraint, *NUC/i,u, can be ranked at any point in the hierarchy; so can MAX-V and *CCC, which are not related to stress assignment. This freedom of ranking of at least the three constraints results in the loss of relationship between the location of stress, which is fixed by the stress constraints, and the location of syncope, which is fixed by these other constraints.

The actual system of Lebanese Arabic – with syncope of all and only unstressed vowels in open syllables – has no special status on this analysis; it is only one of several possible syncope patterns. This rather complicated relationship between the location of stress and the location of syncope is radically simplified by my proposal, which links the two in a direct way: given the setup of the theory, syncope simply cannot apply to a stressed vowel. The patterns in (43) that violate this generalization, predicted by the standard theory, become impossible on my proposal.[30]

4.2.3.4 Archaic Latin

Before moving on to the problematic cases in sections 4.2.3.5 and ff., let me briefly illustrate well-behaved unstressed vowel syncope in a language where two separate patterns appear to hold at the same time. Archaic Latin syncope is notoriously problematic, because the details of the sound change were different in different dialects. Syncope was never carried out to its full extent, and left a patchwork of exceptions and difficult forms in classical Latin (see general discussion of Latin syncope in Vendryes 1902 and in handbooks such as Leumann 1977 and Sihler 1995).

Despite the untidiness of the syncope's result in Latin, a certain set of patterns appear to emerge that make sense in terms of a particular theory of archaic Latin stress. Unlike the familiar classical Latin stress rule, prehistoric Latin had regular initial stress. A natural way of interpreting the archaic Latin stress rule is LEFT-TO-RIGHT MORAIC TROCHEES. I will not give the arguments for such an interpretation here; let us assume that the stage of Latin under discussion had this pattern. The first restriction is that syncope typically applied only to posttonic vowels, i.e. vowels in the second syllable. Furthermore, vowels syncopated in a different way depending on whether there were sonorants or obstruents in their context. As a rule, obstruent-adjacent vowels syncopated less often and in fewer environments than sonorant-adjacent vowels.

[30] Also supporting my analysis is Kiparsky's (2000) argument that, in a stratal theory of phonology, stress is lexical while syncope is postlexical and thus follows it in the derivation.

Secondly, the so-called TWO-MORA LAW restricted syncope to only those posttonic vowels that were followed by at least two moras in the word.

The following table, based on the data in based on Vendryes 1902, summarizes the tendencies of syncope in the second syllable depending on the weight of the preceding and following syllables. In obstruent contexts, vowels only syncopated when flanked by two heavy syllables, i.e. in words beginning with (H)<u>L</u>(H). In sonorant contexts, syncope also applied in case the initial syllable was light, in words beginning with (L<u>L</u>)X. When a light syllable was preceded by a heavy and followed by a light, syncope appears not to have taken place in regardless of segmental context.

(44) Syncope of posttonic vowels

	obstruent contexts	sonorant contexts
(L<u>L</u>)X	no	yes
(H)<u>L</u>(H)	yes	yes
(H)(<u>L</u>L)	unclear	no

Once again, due to the degree of irregularity in syncope's observed outcome, this table must be understood to indicate tendencies, not exceptionless facts. 'Yes' and 'no' in the table mean that syncope was predominant and rare, respectively.

Under the moraic trochee view of Latin stress, the difference between the contexts where the initial syllable is light and those where the initial syllable is heavy is reflected in the location of the right foot boundary: the initial foot parses the second syllable in the former case but not in the latter, as shown in (44). In words shaped like (H)(LL) the posttonic syllable is not only footed but is the head of its foot, i.e. bears secondary stress. The facts in (44) then invite a natural interpretation of syncope. In obstruent contexts syncope applies only to unfooted posttonic vowels, while syncope in sonorant contexts applies both to unfooted vowels and vowel in the weak branch of a foot.

This brief discussion, while not touching on the complexities of the Latin case, serves as another illustration of a well-behaved case of syncope, with the added twist that one language has two subpatterns in different contexts.

4.2.3.5 Yers in Old Russian[31]

A syncope process familiar to phonologists is the deletion of vowels known as yers, which applies in all Slavic languages. Originally a metrically-driven phenomenon, in the modern Slavic languages yer deletion has acquired the characteristics of a crazy rule, not motivated by metrical structure. More pertinent to the point at hand is the stage in the history of Slavic when yer syncope was still productive and conditioned by metrical factors. In this section I address the question of whether or not, at that stage, stressed yers could be deleted.

Late Common Slavic had two extra-short vowels *ĭ* and *ŭ*, known as the yers, weakened or strengthened depending on the environment. All Slavic languages point to the same original environments for strong and weak yers, suggesting that this division took place in the common period, but the languages disagree as to the eventual fate of the yers. In all Slavic languages the weak yers underwent syncope, but the strong yers changed into full vowels ("vocalized") in different ways depending on the language. All modern Slavic languages show synchronic vowel-zero alternations that reflect this original process of yer syncope and vocalization.

The environment for the distribution of weak and strong yers is known as Havlík's Law (see, a.o., V.Kiparsky 1979, Zalizniak 1985, Bethin 1998). Descriptively, a yer is strong only when it precedes a weak yer, counting from right to left. All other yers are weak. In other words, weak yers are word-final and those not adjacent to another yer. In yer sequences, even-numbered yers, counting from the right, are strong. In the data below, the Old Russian forms on the left have the strong yers underlined; the second column shows the modern Russian form with the strong yer vocalized and the weak yer deleted. In some cases, analogical leveling has replaced the expected forms with ones showing a different pattern of yer vocalization.

(45)	sŭmĭ̱rtĭnĕi	smertnoj	'death (adj.loc.sg.fem)'
	dĭnĭ̱sĭ	dnes^j	'today'
	dĭ̱nĭnica	dennica	'morning star'
	lŭ̱žĭno	ložno	'false (adv.)'
	lŭživyi	lživyi	'false (adj.)'
	lĭstĭ̱cĭ	l^{j}stec	'flatterer/antichrist'
	lĭ̱stĭca	lest^{j}ca > l^{j}steca	'flatterer/devil'
	otŭxodĭnikou	otxodnik-	'hermit'
	otŭšĭ̱lĭca	#otšel^{j}ca	'hermit'
	otŭ̱šĭlĭ̱cĭ	#otošlec	'hermit'
	rŭpŭ̱tŭ	rpot > ropot	'murmur'
	sŭlnĭ̱čĭnououmou	solnečnomu	'sunny'
	sŭnĭ̱mŭ	snem > sonm	'gathering'
	sŭ̱nĭmy	sonmy	'gatherings'
	šĭvĭ̱cĭ	švec	'shoemaker'
			(V.Kiparsky 1979: 97)

In addition to the purely rhythmic environment for yer strengthening, a yer can be semi-regularly protected from deletion in case a complex consonant cluster would result.

Note that in the statement of Havlík's Law as presented above, only right-to-left counting and a yer's immediate environment are relevant to its strong or weak status. Although the process has the characteristics of metrical syncope – most importantly, the fact that it applies in a rhythmic fashion to alternating yers – the location of stress is conspicuously absent from the descriptive statement of this rule. This fact appears problematic for the proposal made here, because if syncope is metrically driven, my account predicts that, at the very least, stressed yers should fail to syncopate. In this section I argue that although stress indeed did not matter in Havlík's Law, some complications having to do with the distinction between phonological and predictable accent in Slavic allow an analysis of the yer alternations that does not pose problems for my proposal.

At first blush, it appears that the statement of Havlík's Law given above is incomplete and should in fact include some reference to stress. It is frequently claimed

---

[31] I am grateful to A.K.Polivanova for discussion of the issues in this section.

in Russian historical grammars that stressed yers were strong, and thus stress could protect a yer from deleting. For example, Borkovsky & Kuznetsov (1965: 102) adduce the frequently cited example of the Russian word *dósku* 'board.ACC', which comes from Old Russian *dŭsku*. By the rule of yer vocalization given above, the yer in this form, not being adjacent to any other yer, should syncopate to give *dsku* and eventually *cku*. Although this form is occasionally attested, the surviving *dósku*, according to the Borkovsky & Kuznetsov, shows the influence of accent on the yer.

If Borkovsky & Kuznetsov's claim were not controversial, my discussion could stop here, since, after all, the failure of stressed yers to delete would be a welcome fact for my analysis. However, as Zalizniak (1985) has shown, the claim that stress could protect a yer from deletion, in the form that it is traditionally maintained, is false. Below I will go through Zalizniak's argument, and then suggest a way in which the facts can be reconciled with the theory of syncope that I am advocating.

Slavic accent is a descendant of Indo-European pitch accent system. Both roots and affixes can be specified in the lexicon for the presence or absence of an accent on a particular syllable. Historically, the yers come from the short vowels *ĭ* and *ŭ*, which could be accented just as any other vowel. In Slavic, if the underlying accent fell on a yer that is weak by Havlík's law, the accent shifted to the preceding syllable, or, in absence of such, to the following one. This is a change reconstructible for the entire Slavic domain. As a result, no weak yer could have a PHONOLOGICAL accent at the time of the fall of the yers, which means that the question of whether such a stress could protect a yer from deletion is not decidable: all phonologically accented yers were strong anyway.

Slavic inherited from Indo-European a rule known as The Basic Accentuation Principle, which assigns stress to the leftmost accent in case there is more than one underlying accent in a word, and assigns default stress to the leftmost syllable in forms without any underlying accent. Relevant for the discussion at hand is this latter class of forms, known traditionally as *enclinomena*: could their default initial accent protect a yer from deletion? Here Zalizniak's convincing argument is that stress plays no role in yer deletion: stressed initial weak yers delete just as unstressed ones.

A comprehensive list of phonologically unaccented forms with a potentially weak yer in the first syllable shows that, in the general case, this yer syncopates. These forms fall into several classes, shown below. First, there is a class of unaccented nouns with CV̆C-stems, where the yer is weak in those case forms that have a full-vowel marker. Relevant here are the inherently unaccented case forms, such as the masculine genitive *-a* and the feminine genitive *-i*. In (46)a I list several such noun stems, together with the modern Russian forms of the relevant case form, showing that the yer indeed deletes. (46)b shows some other forms that do not fall into the same declensional pattern, like the numerals *sŭto* and *dŭva* and the interrogative pronouns *kŭto* and *čĭto*. Here, too, the yer in the initial syllable is not protected from syncope by the default initial prominence of these phonologically unaccented forms.

(46)	a.	dĭn-	dn^{j}a	'day'	b.	kŭn^{j}azĭ	kn^{j}az^j	'prince'
		lĭn-	l^{j}na	'flax'		Tĭxvěrĭ	Tver^j	'Tver (city)'
		pĭn-	pn^{j}a	'stump'		sŭto	sto	'100'
		vŭš-	vši	'louse'		kŭto	kto	'who'
		rŭž-	rži	'rye'		čĭto	čto	'what'
						dŭva	dva	'2'

The environment of yer syncope is also met in a large class of verb stems of the shape CV̆C-, where in forms like the first singular the yer is weak and accented according to the initial default principle. Once again, all of the modern Russian forms show that this prominence does not shield the yer from syncope (47).

(47)	žĭgu	žgu	'burn'	mĭru	-mru	'die'
	žĭdu	ždu	'wait'	zĭr^{j}u	zr^{j}u	'see'
	čĭtu	čtu	'honor'	mŭču	mču	'hurry'
	rŭvu	rvu	'tear'	mĭn^{j}u	mnu	'crumple'
	pĭnu	pnu	'kick'	sŭpl^{j}u	spl^{j}u	'sleep'
	vĭru	vru	'lie'	lĭšču	l^{j}šču	'flatter'
	žĭru	žru	'devour'	mĭšču	mšču	'avenge'

Those forms where the yer fails to delete in the weak environment under default initial stress are shown below in (48). In most cases (48)a the yer in question is in an environment that would create a complex consonant cluster if the yer were to fall. Most commonly in such cases, the yer precedes a *st* group. In one single case, the word *sóty* 'honeycomb' (48)b, cluster phonotactics cannot explain the failure of yer deletion.

(48)	a.	lĭst-	l^{j}esti	'flattery'
		mĭst-	m^{j}esti	'revenge'
		čĭst-	česti	'honor'
		tĭst-	t^{j}est^{j}a	'wife's father'
		dŭsku	dosku	'board'
		dŭč-	doči	'daughter'
	b.	sŭt-	soty	'honeycomb (nom.pl)'

In nearly all such forms, cases with yer deletion are also attested.

(49)	OR	MR	also attested
	lĭsti	l^{j}esti	lsti
	mĭsti	m^{j}esti	msti
	čĭsti	česti	čti
	tĭst^{j}a	t^{j}est^{j}a	ct^{j}a
	dŭsku	dosku	dsku, cku
	sŭty	soty	sty

What is crucial is that the survival of the yer in the neighborhood of complex consonant clusters does not depend on the stress, as Zalizniak points out (1985: 170). The cases *dŭ žd^{j}á* > *dožd^{j}á* 'rain', *stĭkló* > *stekló* 'glass', *pĭstró* > *pestró* 'variegated' illustrate that the yer may fail to delete in weak positions in a consonant cluster environment even when it was not default-stressed, i.e. in forms with a stress on another syllable. In other words, whatever the explanation for yer survival in (48)a, it has nothing to do with stress at all. Finally, the isolated case *sŭty* > *soty*, with *sty* also attested, can be explained by analogy to the genitive plural form *sŭtŭ* > *sot*, or simply discounted as an exception – a not unnatural move in case of a process as complicated as the fall of the yers.

To summarize the discussion so far: (a) the role of phonological stress in yer deletion is unknown, because weak yers could not bear phonological stress due to an earlier stress shift, and (b) the default word-initial prominence appearing in phonologically unaccented words failed to protect a weak yer from deletion. Thus, yer deletion appears to present a problem for my proposal: stressed vowels seem to syncopate.

I believe that it is possible to reconcile the facts with the typological generalization that stressed vowels do not syncopate. Let me now take a broader look at the prosody of Old Russian. As mentioned above, given the Basic Accentuation principle inherited from Indo-European, Russian words fell into two broad classes: those with at least one underlying accent, and those without any underlying accent at all. In words of the former class, it is the leftmost stress which surfaced, while in words of the latter class, the *enclinomena*, default stress was assigned to the initial syllable. Jakobson (1963) suggested that the latter was cued by both pitch and intensity, while the former only by pitch; this proposal was in part designed to account for the different behavior of the two types of accent with respect to syncope, and thus cannot be used here without circularity to argue for any particular analysis of syncope.

What is crucial here is that the domain of the application of the Basic Accentuation principle, i.e. the phonological word, was much larger than the grammatical word, and included much cliticized material. This material included not only a host of function words such as prepositions, conjunctions, and particles, but also the *enclinomena*, even if they were lexical words. This is no longer true of Modern Russian, but Jakobson points out that in the archaic language of epic songs, one occasionally finds formulas that show the earlier system. In such formulas, phonologically unaccented words like *more* 'sea', *slovo* 'word', and *grudi* 'breasts' fail to surface with default initial stress but are cliticized to the preceding phonological word, as they would have been in Old Russian: *sinʲó* <u>*more*</u> 'blue sea', *takovó* <u>*slovo*</u> 'such a word', *belý* <u>*grudi*</u> 'white breasts' (Jakobson 1963: 161). Unstressed *enclinomena* in Old Russian occurred especially commonly in prepositional phrases, e.g *béz vorna* 'without a raven', of which there are many relics in the modern language: *ná zimu* 'for the winter', *zá gorodom* 'outside of the city', *ná golovu* 'onto the head'.

Facts such as these are standardly taken as evidence that, despite the existence of the default initial stress rule in Old Russian, in actual practice, phonologically unaccented forms frequently remained unaccented on the surface. Thus, each *enclinomenon* existed in two variants, one with the initial default stress, whatever its phonetic realization, and one without any stress at all. The obvious upshot of this fact is that the yers that deleted in the initial syllables of the *enclinomena* in the forms in (46)-(47) were not obligatorily stressed. On the assumption that yer deletion started in the unstressed variants of these forms, and then became lexicalized, the Russian yer deletion facts do not violate the typological generalization that stressed vowels do not syncopate.

In the modern Slavic languages, the synchronic reflexes of the old yer syncope rules undeniably cause stressed vowel deletion. These cases of vowel deletion are clearly not metrically motivated, falling into the category of "crazy rules" (Bach and Harms 1972), and thus are not subject to the generalization.

4.2.3.6 Southeastern Tepehuan

Rhythmic syncope in Southeastern (SE) Tepehuan has been assumed to target stressed syllables, in contradiction to my proposal. I show here that there is an alternative analysis of the stress system that makes it unnecessary to assume stressed vowel syncope. Because SE Tepehuan is standardly analyzed as an iambic language, syncope in examples such as /tɨroviɲ/ → [(tɨr)viɲ] 'rope' appears to target the second syllable, which, if syncope did not apply, would have been stressed. However, SE Tepehuan can also be analyzed as a trochaic, not an iambic language. Under such an analysis, syncope is regularly posttonic.

In SE Tepehuan stress falls on the initial syllable if it is heavy, and on the second syllable if the word begins with a LH sequence, as shown below.

(50) a. **H-initial words**: initial stress

vóohi	'bear'
vátvirak	'went to bathe'
táatpɨʃ	'fleas'

b. **LH-initial words**: second syllable stress

takáaruiʔ	'chicken'
kakárvaʃ	'goats'
sapók	'story'
tapɨ́ɨʃ	'flea'

This fact has led to the standard analysis of the stress system, endorsed by Kager 1997 and Gouskova 2003, in terms of left-to-right iambs. A single iambic (H) or (LH) foot is constructed at the left word edge in order to account for the stress facts in (50). However, the stress facts just mentioned, by themselves, do not inevitably diagnose an iambic system. There are many systems with identical stress generalizations but a clearly trochaic foot structure. For example, in Tümpisa Shoshone (Dayley 1989; Hayes 1995: 180), main stress falls on the initial syllable if it is heavy, and on the second syllable in words beginning with a LH sequence, just as in SE Tepehuan. These two patterns are illustrated below. The data below are from Dayley 1989: 436, cited in their phonemic transcriptions for simplicity, ignoring the effects of predictable processes like vowel devoicing. CVV syllables are heavy, CVC and CV are light.

(51) a. **H-initial words**: initial stress

té:wìngkɨppɨ̀hantɨ̀n	'told (someone)'
pá:kàntɨn	'having water'
tá:ttsɨ̀witɨ̀n	'seven'

b. **LH-initial words**: second syllable stress[32]

[32] Some LH-initial words in Tümpisa Shoshone may optionally have stress on the first syllable. There appear to be no such words that MUST be initial-stressed.

kukkwí:ppɨh	'smoke'
pihná:witɨn	'bee'
pomá:ppɨh	'grass, hay'
topó:mpi	'desert'
tuttsá:ppɨh	'dirt, dirty'
tsitó:hin	'push'
tɨtsɨ́:nna	'to count'
wɨkkɨ́:nappɨh	'fog'
pakɨ́:nappɨh	'cloud'

Despite the identity of the stress generalization with SE Tepehuan, Tümpisa Shoshone is standardly analyzed as trochaic. What would unambiguously diagnose a language with a pattern like the one in (50) or (51) as iambic or trochaic is how initial LL sequences behave: in iambic languages like Hixkaryana or Alaskan Yupik, words beginning with LL receive second syllable stress, while in trochaic languages like Tümpisa Shoshone, stress falls on the initial syllable of LL sequences, as illustrated below with data from Dayley 1989: 436 and ff.

(52) **LL-initial words**: initial stress

náttusùʔungkàntɨn	'doctor'
ké: námokkùpɨhkantɨn	'not having money'
nátɨpìnniyàha	'be named'
túkummàhannìngkɨnnà	'cook for'

The obvious analysis of the Tümpisa Shoshone stress facts is that a perfect moraic trochee, i.e. a (H) or a (LL) foot, is built as close to the left edge as possible. This results in H- and LL-initial words receiving initial stress, but LH-initial words receiving second syllable stress because no perfect moraic trochee can parse the initial syllable.

In SE Tepehuan, it is impossible to observe directly how LL-initial words behave, because, due to rhythmic syncope, no such words occur on the surface: in all underlying CVCV-initial words, the second vowel is deleted. Therefore, analyzing SE Tepehuan as an iambic language is a tempting choice in view of the stress of LH words, but an arbitrary one, because the crucial data which would decide between the iambic and the trochaic analyses are absent. If SE Tepehuan is treated as a trochaic system, then the generalization that it is the unstressed syllables that syncopate can be maintained.

The following examples illustrate the stress system analyzed in terms of moraic trochees.

(53) Stress

a.	(vóo)hi	'bear'
	(vát)virak	'went to bathe'
	(táat)pɨʃ	'fleas'
b.	ta(káa)ruiʔ	'chicken'
	ka(kár)vaʃ	'goats'
	sa(pók)	'story'
	ta(pɨ́ɨʃ)	'flea'

Black (1993) considers the trochaic analysis of SE Tepehuan briefly, but argues against it based on the following data, where the vowel that syncopates immediately follows a heavy syllable. In the following data (Blake 1993: 47), the input vowels that syncopate are underlined. In both cases, they follow the main stressed syllable.

(54)	/naan<u>a</u>kasɨr/	→	naankasɨr	'scorpions'
	/vapoot<u>V</u>podaʔ/	→	vapootpodaʔ	'worms'

The data in (54) are indeed not easily compatible with an analysis with standard L→R moraic trochees, because the syncopating vowel would bear secondary stress. Forms beginning with a HLL sequence would be footed as (H)(LL), and thus the vowel in the second syllable would be ineligible for syncope, according to the hypothesis pursued here. However, there is a trochaic analysis of SE Tepehuan that is consistent with the data in (54): one where a high-ranked *CLASH prevents adjacent syllables from being stressed. Two options are available: either the forms like those in (54) contain a "Germanic foot" (Dresher & Lahiri 1991), i.e. the uneven trochee (HL), or else forms beginning with HLLL contain an unparsed syllable separating the main stress from the first secondary stress, (H)L(LL). These two options are illustrated below; I will not argue for one or the other option here.

(55)	a.	(HL)(LL)L	(naana̸)(kasɨr)
	b.	(H)L(LL)L	(naa)na̸(kasɨr)

Whichever of the two analyses in (55) is chosen, vowel syncope is compatible with my hypothesis: it is the UNSTRESSED posttonic vowel that deletes.

SE Tepehuan has a vowel shortening process that applies to long vowels in unfooted final syllables. It can be seen in reduplication of words that begin with a CVV syllable: once the underlying long vowel is no longer in the main stress foot, it shortens (56)a. The data in (56)b show that this shortening process does not apply to long vowels that are parsed by feet.

(56) Vowel shortening in unstressed syllables

a.	(kóoʔ)	'snake'	/koo-ko̲o̲ʔ/	(kóo)ko̲ʔ	'snakes'
	(káam)	'cheek'	/káa-ka̲a̲m/	(káa)kam	'cheeks'
b.	ga(gáa)	'cornfield'	*(gága)		
	to(páa)	'pestle'	*(tópa)		
	ta(pɨ́ɨʃ)	'flea'	*(tápɨʃ)		

SE Tepehuan also has a final vowel apocope process that deletes short vowels word-finally (57)a, unless the output form would end in a consonant cluster or *h* (57)b. Final long vowels in LH words are not deleted (57)c.

(57) Apocope: short final vowels delete unless preceded by cluster or *h*

a.	/hiɲ#novi̲/	hiɲ#(óv)	'my hand'
	/hiɲ#noo-novi̲/	hiɲ#(ɲóo)nov	'my hands'
	/tu#huana̲/	tu#(huán)	'he is working'
	/nakasɨɽi̲/	(nák)sɨɽ	'scorpion'

b.	/hupna/	(húp)na	'pull out'	*(húpn)
	/voohi/	(vóo)hi	'bear'	*(vóoh)
c.	/gagaa/	ga(gáa)	'cornfield'	*(gág)

Syncope applies to even-numbered non-final open syllables, including syllables containing long vowels. Prefixes, here separated from the stem by the sign #, are ignored for the purposes of stress assignment and syncope. The data in (58)a show the deletion of short vowels, and (58)b the deletion of long vowels.

(58)	a.	/tɨrovin/	(tɨr)vin	'rope'
		/tɨɨtɨrovin/	(tɨ́ɨt)ropin	'ropes'
		/totopaa/	(tót)pa	'pestles'
		/taatapɨɨʃ/	(táat)pɨʃ	'fleas'
		/taatakaaruiʔ/	(táat)karuiʔ	'chickens'
		/tu#maamatuʃidʒaʔ/	tu#(máam)tuʃdʒaʔ	'will teach'
	b.	/gaagaagaʔ/	(gáaˀŋ)gaʔ	'he will look around for it'[33]
		/suisuimaɽ/	(súis)maɽ	'deer (pl.)'
		/hiɲ#ɲuuɲuutʃiʃ/	hiɲ#(ɲúuɲ)tʃiʃ	'my brothers-in-law'
		/hiʃ#maimaikak/	hiʃ#(máim)kak	'sweet (pl.)'

Syncope does not apply to long vowels following an initial light syllable (see (53)b and (56)b), showing that only unfooted vowels may syncopate. Apocope takes precedence over syncope, and feeds the stress rule, as shown below.

(59)	a.	/hiɲ#noo-novi/	hiɲ#(ɲóo)nov	'my hands'	*hiɲ#(ɲoon)vi
		/ʃi#ʔomiɲi/	ʃi#ʔo(míɲ)	'break it!'	*ʃi#(ʔóm)ɲi
		/naanakasɨɽi/	(náan)kasɨɽ	'scorpions'	*(náan)kasɽi

From the point of view of the theory developed here, SE Tepehuan is only problematic if the stress system is analyzed as iambic. If, on the other hand, the moraic trochee is the basis of the stress system, then syncope applies to UNSTRESSED vowels.

[33] This form shows the effects of an unrelated process, which turns coda voiced obstruents into preglottalized nasals.

The treatment of Southeastern Tepehuan as trochaic appears to be endorsed by Willet's (1991) description, although he does not formulate it in terms of foot structure. However, his statement that "vowels are deleted from every second nonfinal, open syllable following the accented syllable" (1991: 23) clearly suggests that on his view, derivationally, LL-initial words receive stress on their first syllable and then syncope applies to the posttonic – i.e. second – syllable.

In light of the reanalysis of the stress system of SE Tepehuan as iambic, the syncope process presents no problem whatsoever for the typological generalization under discussion.

4.2.3.7 Hopi

Of all the cases discussed here, Hopi presents the most serious challenge to my generalization: stressed vowels appear to syncopate. I argue that a closer look at the data shows that Hopi syncope is, in fact, well-behaved and applies in posttonic syllables.

Hopi has a typical iambic stress system: the initial syllable is stressed if heavy, otherwise the second syllable. Words beginning with the sequence LL are stressed on the second syllable, unlike Southeastern Tepehuan, which clearly diagnoses the Hopi system as iambic. Hopi also shows a preference, typical in iambic languages, to assign initial stress to disyllabic words, no matter what their weight profile. CVV and CVC syllables count as heavy, CV as light. The stress system is illustrated below.

(60) a. Initial if heavy
ʔác.ve.wa 'chair'
b. Otherwise the second syllable
ca.qáp.ta 'dish'

	qö.tö́.som.pi	'headband'
	ki.yá.pi	'dipper'
c.	Disyllables have initial stress	
	kó.ho	'wood'
	táa.vok	'yesterday'
	má.mant	'maidens'

The syncope pattern as presented by Gouskova (2003: 97) is as follows: in /LLL/ and /HLL/ words, the second vowel deletes, while in longer words of the shape LLL, the third vowel syncopates. In addition, because superheavy (CVVC) syllables are prohibited, whenever such syllables arise, their nucleus vowel shortens (61)b.

(61) a.	/LLL/ words					
	/soma-ya/	sóm.ya	'tie (pl.)'	cf.	sóma	'tie (sg.)'
	/soma-ŋ^{w}i/	sóm.ŋ^{w}i	'tie (nomic)'			
	/soʔa-ya/	sóʔ.ya	'die (pl.)'		sóʔa	'die (sg.)'
b.	/HLL…/-initial words					
	/tooka-ni/	tók.ni	'sleep (fut.)'		tóoka	'sleep (non-fut.)'
	/mooki-ni/	mók.ni	'die (fut.)'		móoki	'die (non-fut.)'
	/naala-ya-n-ta/	nál.yan.ta	'be alone'		náala	'alone'
c.	/LLL…/words					
	/navota-na/	na.vót.na	'inform'		navóta	'notice'
	/kawayo-sa-p/	ka.wáy.sap	'as high as a horse'		kawáyo	'horse'
	/aŋa-katsina/	a.ŋák.tsi.na	'Long Hair kachina'		áŋa	'long hair'
					katsína	'kachina'
	/tuhisa-tuwi/	tu.hís.tu.wi	'ingenuity'		tuhísa	'ingenious'
					túwi	'knowledge'
	/qövisa-tapna/	qö.vís.tap.na	'make pout'		qövísa	'bad sport'

While syncope in (61)b-c is clearly postonic, the cases in (61)a, i.e. the trisyllabic LLL words like /soma-ya/, are problematic. Here syncope appears to apply to a syllable that would otherwise bear stress. Gouskova's analysis of such a pattern is to attribute syncope to the joint action of NONFIN and SWP, as illustrated below in (39). The non-syncopating candidate *somáya* fails because its stressed syllable is not heavy, in violation of SWP, and lengthening the vowel to *somáaya* is not an option due to high-ranking

DEP-μ. Syncopating the final vowel to produce **somáy*, even though satisfying SWP, is also not possible because **somáy* would violate NONFIN. This leaves syncopating the would-be stressed vowel to give the output *sómya*.

(62) /soma-ya/

		DEP-μ	NONFIN	SWP	PARSE-σ	MAX-V
	(somá)ya			*!	*	
	(somáa)ya	*!			*	
☞	(sóm)ya				*	*
	(somáy)		*!			*

Such an analysis is impossible in my theory, where the procedural constraint ✦SWP cannot lead to a shift of stress, as illustrated below. The candidate *sómya* incurs a fatal ✦-violation of SWP, because it has initial stress, while the designated stress locus is on the second syllable.[34] Is Hopi then a counterexample to my proposal?

(63) /soma-ya/ DS of stress: second syllable

		DEP-μ	NONFIN	✦SWP	PARSE-σ	MAX-V
☞	(somá)ya			*	*	
	(somáa)ya	*!			*	
	(sóm)ya			✦	*	*!
	(somáy)		*!			*

In reality, Hopi syncope does not challenge the generalization that stressed vowels cannot syncopate; Gouskova's proposal as presented in (39) above results from an incorrect analysis of the syncope environment. The key to the solution is to be found in footnote 39 in Gouskova 2003: 97, which mentions that syncope applies only in derived environments, but leaves the account of the derived environment effect outside of the scope of the analysis. Indeed, forms like *ʔácvewa* 'chair' and *qötö́sompi* 'headband' from (60) and *navóta* 'notice' from (61) show that syncope does not apply to give outcomes like **navta*; in all such forms the environment of syncope is non-derived.

On the other hand, the problematic set of cases of the type /somaya/ → *sómya* all consist of a disyllabic stem followed by a monosyllabic affix. Recall that disyllables regularly have initial stress (60)c. Supposing that the stress of the affixed from /somaya/ inherits the stress of the base form /soma/, syncope is in fact posttonic. This hypothesis is presented in derivational terms below.

(64)	/soma-ya/	→	[sóma-ya]	→	*sómya*
	/soma-ŋʷi/	→	[sóma-ŋʷi]	→	*sómŋʷi*
	/soʔa-ya/	→	[sóʔa-ya]	→	*sóʔya*

It turns out that Hopi gives us direct confirmation of the claim in (64) that affixed forms retain the accetuation of the base. Jeanne's (1982) description lists a substantial number of forms which are exceptions to syncope, for arbitrary lexical reasons (1982: 248). There exist both LL and HL stems which fail to delete their second vowel when suffixed. The forms listed below should undergo syncope on the model of (61), but unpredictably fail to do so.

(65)		UR	FUTURE	NON-FUTURE	gloss
	a.	/maqa-ni/	máqa-ni	máqa	'give'
		/tɨwa-ni/	tɨ́wa-ni	tɨ́wa	'find'
		/sowa-ni/	sówa-ni	sówa	'eat'
	b.	/tɨɨva-ni/	tɨ́ɨva-ni	tɨ́ɨva	'throw away'
		/peena-ni/	péena-ni	péena	'write'
		/niina-ni/	níina-ni	níina	'kill'

What is crucial in these forms is their stress (Jeanne 1982: 256). Precisely these forms fail to surface with stress on the second syllable, as would be expected if the iambic stress rule had applied. Instead, we have *máqa-ni* rather than **maqáni*. This fact supports my proposal that the stress of the affixed form is retained from the base form, and thus syncope in forms from (61) like /soma-ya/ → *sómya* is in reality posttonic. In

[34] Using the procedural or standard version of NONFIN would have no effect on the outcome, so I keep the standard version for simplicity.

other words, the attachment of the suffix *-ni* and other suffixes counterfeeds the stress assignment rule. This opaque interaction is directly supported by the forms in (65)a, where there is no syncope to obscure the location of stress.

This reanalysis also solves Gouskova's derived environment problem: forms like *navóta* fail to syncopate on the model of /soma-ya/ because the environment of syncope is not met. The vowel in the second syllable of *navóta* is not posttonic, and hence not eligible for syncope in the first place.

To summarize the discussion of Hopi so far: Gouskova's claim that the forms in (61)a show stressed vowel syncope was based on an erroneous analysis of the stress system of the language. A broader look at the phonology of Hopi shows that the vowel that deletes in (61)a is in fact not stressed but posttonic. Now I move on to another process in Hopi, reduplication, which also appears to present a problem for my proposal, and also argue for a reanalysis that saves the typological generalization.

Hopi has a reduplication process whose outputs undergo syncope (data from Jeanne 1982: 249 and ff.). LL stems reduplicated by copying the initial syllable and syncopating the following vowel: $C_1V_1C_2V_2$ reduplicates to $C_1V_1C_1C_2V_2$ (see (66)a). HL stems copy the initial long vowel and shorten the second vowel of the resulting form: $C_1VV_1C_2V_2$ becomes $C_1VV_1C_1V_1C_2V_2$ (see (66)b). This shortening process, targeting long vowels after initial long vowels, is phonologically regular.

(66)		SG	PL	gloss
	a.	koho	kókho	'wood'
		como	cócmo	'hill'
		sihɨ	síshɨ	'flower'
		leŋi	lélŋi	'tongue'
		tamö	tátmö	'knee'
	b.	saaqa	sáasaqa	'ladder'
		tooci	tóotoci	'shoe'
		siivɨ	síisivɨ	'pot'
		sooya	sóosoya	'planting stick'
		noova	nóonova	'food'

There are two alternative analyses of the shape of the reduplicant. The first possibility is that in LL the reduplicant is CV, and syncope applies to the second syllable of the resulting LLL words, while the reduplicant in HL stems is CVV. The long vowel in the second syllable would shorten by the regular shortening process. The second possibility is to assume a CVV reduplicant in all forms, with syncope and shortening in closed syllables applying in LLL words (cf. shortening in closed syllables in (61)b). These two options are illustrated below.

(67) a. **Option 1**: CV reduplicant in LL stems, CVV reduplicant in HL stems

/CV-koho/	→	ko-koho	→	kókho		'wood'	
/CVV-saaqa/	→	saa-saaqa	→	sáasaqa		'ladder'	

b. **Option 2**: CVV reduplicant in all stems

/CVV-koho/	→	koo-koho	→	kookho	→	kókho
/CVV-saaqa/	→	saa-saaqa	→	sáasaqa		

Gouskova appears to endorse Option 1. However, the Option 1 analysis is incompatible with my proposal, because it would necessarily involve stressed vowel syncope in LLL forms like /kokoho/. If, on the other hand, Option 2 can be argued to be correct, then the generalization that syncope is posttonic can be preserved, because Hopi's iambic stress system would automatically force stress on the initial syllable of [koo-koho], and syncope would regularly apply to the second syllable. It is then crucial for the purposes of the discussion here to establish that it is Option 2, not Option 1, that is the correct analysis of the shape of the reduplicant.

Jeanne's (1982) description provides several pieces of evidence in favor of the Option 2 analysis. I present three arguments here. First, syncopation in reduplicated forms whose bases have more than two syllables clearly favors the CVV analysis. The forms in (68)a below have LL-initial bases, while (68)b begin with LH.

(68)		SG	PL	gloss	
	a.	kɨyapi	kɨ́kyapi	'dipper'	*kɨkɨ́ypi
		yɨŋʸapɨ	yɨ́yŋʸapɨ	'plague'	*yɨyɨ́ŋʸpɨ
		pitanakci	píptanakci	'hat'	*pipítnakci
		qötösompi	qö̋qtösompi	'headband'	*qöqö̋tsompi
	b.	caqapta	cácqapta	'dish'	
		panapca	pápnapca	'widow'	
		möcikvɨ	mö̋mcikvɨ	'trash'	
		melooni	mémlooni	'melon'	

For LL-initial bases, if the reduplicant were CV, then the resulting form would begin with LLL, where it is the third, not the second vowel that would be expected to syncopate (cf. (61)c). The two options are illustrated derivationally in (69). We would have the incorrect outputs **kɨkɨ́ypi* and **pipítnakci* along the lines of *navótna* from /navotana/. Instead, it is the second vowel that deletes, showing that the reduplicant must be CVV, which renders the second vowel post-tonic.

(69)						
	CV reduplicant:	pi-pitanakci	→	*pipítnakci		
	CVV reduplicant:	pii-pitanakci	→	piiptanakci	→	píptanakci

Second, reduplicants with initial CVC syllables contain a long vowel, showing that the reduplicant is indeed CVV. Note that the second vowel in these forms does not syncopate, because it is in a closed syllable.

(70)	SG	PL	gloss
	naqvɨ	náanaqvɨ	'ear'
	tɨsna	tɨ́ɨtɨsna	'body dirt'
	napna	náanapna	'shirt'
	ŋɨmni	ŋɨ́ɨŋɨmni	'flour'

The third and strongest piece of evidence in favor of the Option 2 (CVV) analysis of reduplication concerns exceptions. Just as in the case of LLL words with suffixes (65), there are reduplicated forms which undergo syncope optionally, creating variant forms. Just in those cases where the second vowel surfaces faithfully (and thus does not create an initial closed syllable), the long vowel of the reduplicant emerges, as illustrated by the forms in the third column below.

(71)	SG	PL *variant 1*	PL *variant 2*	gloss
	léŋi	lélŋi	léeleŋi	'tongue'
	ŋáhɨ	ŋáŋhɨ	ŋáaŋahɨ	'medicine'
	ʔówa	ʔóʔwa	ʔóoʔowa	'stone'
	kʷite	kʷíkʷte	kʷíikʷite	'braid'

I take these three arguments to suggest that the reduplicant is uniformly CVV, that Option 2 in (67) is the correct analysis of Hopi, and that, therefore, syncope is uniformly posttonic.

There are, in addition to the patterns discussed above, a small set of polysyllabic animate nouns that appear to reduplicated with CV- in the plural, and also take the suffix *-t*. These nouns undergo an unprecedented syncope of the FOURTH syllable, a fact for which Gouskova's analysis has no answer.

(72)	SG	PL underlying	PL surface	gloss
	koyoŋo	/ko-koyoŋo-t/	kokoyoŋt	'turkey'
	ʔaŋʷɨsi	/ʔa-ʔaŋʷɨsi-t/	ʔaʔaŋʷɨst	'crow'
	laqana	/la-laqana-t/	lalaqant	'squirrel'
	qapira	/qa-qapira-t/	qaqapirt	'goat'
	tokoci	/to-tokoci-t/	totokoct	'wild cat'

Although Jeanne does not mark stress in these forms, a reasonable hypothesis about stress assignment can explain the unexpected syncope pattern. If, just as suffixed forms, CV-reduplicants retain the stress of their bases, then syncope here would also be posttonic. The hypothesis is that the stress in the reduplicated form *ko-koyóŋo-t* corresponds to the base form stress *koyóŋo*. If this is the case, then the third syllable in the CV-reduplicants would indeed be stressed and the fourth syllable would be expected to undergo posttonic syncope. I repeat that this hypothesis cannot go beyond speculation, because Jeanne does not supply stress information for the forms in (72).

To summarize the discussion of Hopi, I have shown here that a closer look at the phonology of the language ensures that in all cases syncope is simply posttonic, and that Gouskova's proposal was based on an incorrect analysis of stress and reduplication. Let me conclude with presenting an overall picture of the relevant processes in Hopi phonology, from a stratal point of view (Kiparsky to appear).[35] Two strata are necessary: an earlier one with CVV reduplication and stress assignment, and a later one, with tense-aspect affixes and possibly CV reduplication, posttonic syncope, and shortening.

This proposal is illustrated derivationally below with the two representative examples of syncope: suffixed forms like *sómya* and reduplicated forms like *kókho*. Within each stratum, the interaction of the phonological processes is transparent; the only instance of opaque interaction between the two stata is the inheritance of the stress pattern of the base form *sóma* in the suffixed form *sómya*.

(73)

		/soma/	/koho/
Stratum 1:	CVV Reduplication		koo-koho
	Stress assignment	sóma	kóo-koho
Stratum 2:	Tense-aspect affixes	sóma-ya	
	Posttonic syncope	sómya	kóokho
	CV Reduplication (72) (?)		
	Shortening		kókho
		sómya	*kókho*

Once again, what is crucial for the purposes of the discussion here is that the picture in (73) does not require stressed vowel syncope, and thus Hopi is not a counterexample to my typological generalization.

4.2.3.8 Central Alaskan Yupik

Another serious challenge to the generalization that stressed vowels do not syncopate is presented by Central Alaskan Yupik (CAY). The discussion here is based on Jacobson 1995, Hayes 1995, and Gordon 2001. In a nutshell, stressed vowels in open syllables are lengthened, but long schwas are prohibited. Thus, if a stress were to fall on a schwa in an open syllable, the schwa deletes. In OT terms, this deletion process is standardly understood as a last resort in a situation when other options are not available: CAY has an absolute prohibition of a stressed short schwa *ə́ and of a long schwa *ə̄. The only way of satisfying both of these constraints when stress would otherwise fall on a schwa is to delete it.

The stress system of CAY is iambic. Initial CVC count as heavy, but CVCs elsewhere in the word count as light. The following data show the iambic stress generalization.

(74)	(əɣán)(ɣuq)	'she begins to cook'
	(kúi)(ɣú:q)	'it is a river'
	(áŋ)(já:)	'his boat'
	(íʁ)(nía)(ŋuk)	'she acquires a child'

Stressed non-final vowels in open syllables are lengthened, as is typical in iambic systems, as shown by the following data.

(75)	/akutamək/	(akú:)(tamə́k)
	/nunaka:/	(nuná:)(ká:)
	/aŋyaχpaka/	(aŋ)(yaχpá:)ka

[35] Nothing hinges here on a commitment to the stratal view of phonological opacity; I use it here because it allows an especially conspicuous picture of the phonology of the language.

Crucially for the present discussion, if a stress were to fall on a schwa, the schwa deletes and stress shifts one syllable to the left. This deletion applies not only in the cases where the schwa is preceded by a full vowel (76)a, but also when it is preceded by another schwa (76)b-c.

(76) a. /qanʁutəka:/ (qán)(ʁút)(ká:) *(qán)(ʁutə́)(ká:)
'he's talking about her' *(qán)(ʁutə́:)(ká:)
b. /aŋutə-ŋə-ciq-uq/ (aŋú)(tə́ŋ)(ciqúq) *(aŋú)(təŋə́)(ciqúq)
's/he will acquire a man' *(aŋú)(təŋə́:)(ciqúq)
c. /nuna-nətə-ɬini-luni/ (nuná:)(nə́t)(ɬiní:)(luni) *(nuná:)(nətə́
)(ɬiní:)(luni)
's/he apparently being in the village' *(nuná:)(nətə́
:)(ɬiní:)(luni)

Furthermore, as shown by the following datum, weak schwas in iambic feet do not delete.

(77) /qana:təkaqa/ (qan)(na:)(təka:)qa 'I speak about it'

These data indicate that schwa deletion applies IF AND ONLY IF the schwa would otherwise be stressed and in an open syllable. Deletion does not happen when the schwa is unstressed, or when it would be in a closed syllable. Thus, the standard analysis of this deletion process is that it results from the impossibility of simultaneously satisfying the two constraints against stressed monomoraic syllables (SWP), and against long schwas *$\bar{ə}$.

This interpretation is not compatible with the idea that syncope is what happens to weak vowels. In particular, SWP, on my theory, is not capable of causing syncope and the concomitant stress shift that results from deleting the stressed vowel. Indeed, the following tableaux illustrate that the actual winner, *(qán)(ʁút)(ká:)*, is a perpetual loser, because it incurs a fatal ✦ violation of the ✦SWP constraint. Depending on the ranking between ✦SWP and *ə:, the predicted winner is either the candidate with a long schwa, in violation of the *ə̄ constraint (78), or the candidate with a stressed short ə, thus violating ✦SWP (79).

(78) Designated location of stresses: (qán)(ʁutə́)(ka:)

			✦SWP	*ə̄	MAX-ə
/qanʁutəka:/		(qán)(ʁutə́)(ká:)	*!		
	(☞)	(qán)(ʁút)(ká:)	✦!		*
	☹	(qán)(ʁutə́:)(ká:)		*	

(79) Designated location of stresses: (qán)(ʁutə́)(ka:)

			*ə̄	✦SWP	MAX-ə
/qanʁutəka:/	☹	(qán)(ʁutə́)(ká:)		*	
	(☞)	(qán)(ʁút)(ká:)		✦	*!
		(qán)(ʁutə́:)(ká:)	*!		

The strategy I would like to pursue in dealing with this problem is along the lines of the solution presented for a similar problem in Carib above, although the situation is somewhat more difficult in CAY. Recall that Carib, I argued, has a sonority-sensitive stress subregularity on top of its iambic stress system, which accounts for an apparent deletion of a stressed vowel. In Yupik, the solution would run along the same lines: I would argue for making use of an additional sonority-sensitive constraint. Specifically, if the a constraint against stressed schwas is able to force stress shift onto a preceding syllable, and a SEPARATE constraint ensured the deletion of the now weak schwa, then we might solve the problem posed by (78)-(79) in a manner analogous to the Carib case. The constraint in this case would be the familiar constraint against ə in head syllables, ✦*HEAD{ə}. Ranked high, it would have precedence over ✦SWP, and would thus affect the designated place of stress for that constraint.

(80) Designated location of stresses: (qán)(ʁútə)(ká:)

		*ə̄	✦*HEAD{ə}	✦SWP	MAX-ə
/qanʁutəka:/	(qán)(ʁutə́)(ká:)		*!	✦	
	(qán)(ʁútə)(ká:)			*!	
	☞ (qán)(ʁút)(ká:)				*
	(qán)(ʁutə́:)(ká:)	*!		✦	

Of course, this reanalysis only takes care of the cases where stress shifts off a schwa to a full vowel, still leaving those like (76)b-c where it shifts to a preceding schwa. Consider the case (76)b, repeated below.

(81) /aŋutə-ŋə-ciq-uq/ (aŋú)(tə́ŋ)(ciqúq) *(aŋú)(təŋə́)(ciqúq)
's/he will acquire a man' *(aŋú)(təŋə́:)(ciqúq)

If the generalization that only unstressed schwas can syncopate is to be preserved, then something must cause the shift of the stress from *ŋə* to *tə*, i.e. some constraint must prefer the foot (tə́ŋə) over (təŋə́). The constraint ✦*HEAD{ə}, which was used to cause a stress shift from a schwa to a full vowel in (80), cannot do this job here. However, according to the iambic-trochaic law (Hayes 1995), iambic feet, unlike trochaic feet, have a tendency for quantitative unevenness. The quantitatively even trochaic foot (tə́ ŋə) better conforms to this preference than the quantitatively even iambic foot (təŋə́). The shift of stress off the second schwa onto the first schwa would then be a consequence of general principles of rhytmic organization. This shift can be accounted for by a constraint like UNEVENIAMB, which would penalize right-headed feet containing quantitatively identical syllables.

4.2.3.9 Bedouin Hijazi Arabic

The interaction of vowel deletion with stress in Bedouin Hijazi Arabic (BHA; Al Mozainy et al. 1985) appears to be a difficult counterexample to my proposal. In certain contexts, low vowels in this language delete, and this deletion process may apply to stressed vowels, in which case the stress migrates to the following syllable. Al Mozainy et al. used these facts to argue for deriving the directionality of the stress shift from the tree-based representation of the metrical structure of the language.

The stress rule in BHA is identical to the rule in Lebanese Arabic (see section 4.2.3.3 above): superheavy ultimas are stressed, otherwise heavy penults, otherwise antepenults. The low vowel *a* deletes in an open syllable if followed by another open syllable with the nucleus *a*. The following rule was used by Al Mozainy et al. (1985: 136) to deal with the *a*-deletion facts.

(82) $a \rightarrow \varnothing \,/\, C \_\_ [Ca]_{\sigma}$

The following examples illustrate that this process may apply to the stressed vowel (i.e. the vowel that would have been stressed had the deletion rule not applied). In the following data, non-deleting forms are in the left column, and the vowel that deletes in the right column forms is underlined.

(83)	a. sáḥab	'he pulled'	/sa̲ḥabat/	→	sḥábat	'she pulled'
	b. saḥábna	'we pulled'	/sa̲ḥabaw/	→	sḥábaw	'they (m.) pulled'
	c. náxal	'palm trees'	/na̲xalah/	→	nxálah	'a palm tree'
	d. gʸálaʕ	'castles'	/gʸa̲laʕah/	→	gʸláʕah	'a clastle'
	e. sálag	'hunting dogs'	/sa̲lagah/	→	slígah[36]	'a hunting dog'
	f. ɣánam	'sheep'	/ɣa̲nami/	→	ɣními	'my sheep'

(Al Mozainy et al. 1985: 136)

These forms show the deletion of the vowel that would be stressed by the normal stress rule of the language: e.g. stress would fall on the antepenult in *saḥabat* by the regular stress rule, but the *a* in that syllable deletes.

These data are compatible with an opaque serial interpretation: if syncope applies before stress assignment, the stress rule would give the correct outcome for forms like *sḥ*

[36] This and the following form, as well as the forms in (84), undergo a separate raising process, *a* → *i*, which is not important for the purposes of the discussion here.

ábat and other forms in the right column of (83). This solution, however, is not available for longer words. As the following forms show, stress must be assigned prior to *a*-deletion, because otherwise there is no way to account for its surface location.

(84)	a. ʔínkisaṛ	'he got broken'	ʔinksáṛat	'she got broken'
	b. ʔíntiðaṛ	'he waited'	ʔintðáṛan	'they (f.) waited'
	c. ʔíftikaṛ	'he remembered'	ʔiftkáṛaw	'they (m.) remembered'
	d. ʔíxtibaṛ	'he took an exam'	ʔixtbáṛaw	'they (m) took an exam'

(Al Mozainy et al. 1985: 137)

If *a*-deletion simply preceded stress assignment, we would expect incorrect antepenultimate stress, due to derivations like /ʔinkasaṛat/ → [ʔinksaṛat] → * *ʔínksaṛat*. Rather, stress must be assigned prior to vowel deletion, and separate principles (which were the focus of Al Mozainy et al.'s study) must ensure that it shifts rightward under stress deletion: /ʔinkasaṛat/ → [ʔinkásaṛat] → [ʔink∅́saṛat] → *ʔinksáṛat*. In other words, the BHA data appear to stand in direct contradiction to my proposal: stressed vowels can syncopate.

In the cases discussed above my strategy has been to reanalyze the environment of syncope itself and to show that a closer look at the process makes it unnecessary to treat syncope as applying to stressed vowels. Such was the approach in Hopi and SE Tepehuan. Here, however, I take a different course by suggesting that the process in question, although it does apply to stressed vowels, is not metrically-driven syncope, but rather an abstract process not unlike the deletion of the yers in modern Slavic languages. In fact, Al Mozainy et al. remark that the *a*-deletion process in BHA is unlike many other syncope process in Arabic dialects in at least two respects: it applies to stressed vowels, and it singles out the MOST sonorous vowel *a*, not the LEAST sonorous high vowels as is typical of syncope processes.

> "[M]ost syncope rules explicitly eschew deletion of stressed vowels or precede the assignment of stress and thus could not show stress shift … In other words, these syncope rules delete vowels that are less prominent by virtue of their lack of stress and, in some cases, their relatively lesser sonority… [The BHA deletion rule] appears to function as a fairly abstract kind of dissimilation, eliminating an underlying configuration of two successive identical nuclei by deleting the first of them". (Al Mozayni et al. 1985: 142)

I conclude that the *a*-deletion rule in BHA is not sensitive to metrical structure, but is a "crazy rule" in the sense of Bach and Harms (1972). While the existence of such rules weakens the proposed typological generalizations, they are an inevitable part of phonology, and the typology of syncope is no exception.

4.2.4 Syncope blocking conditions

One striking generalization is that syncope can be blocked in those and only those environments that allow vowel epenthesis. These are the consonant cluster conditions of two types (conditions on complex margins, and conditions on sonority), and word minimality conditions. In many languages, syncope fails to apply just in case it would create a marked consonant cluster, just as those consonant clusters may be resolved by epenthesis. In a smaller number of languages, syncope fails to apply if it would create a subminimal word. Syncope blocked by word minimality is attested both in systems with a disyllabic and a bimoraic minimum. An example of the former is Lardil (Hale 1973), where final vowels delete only in words longer than three syllables, which is part of the general prohibition against monosyllabic words (recall from (9) above that underlying monosyllables are augmented to respect the disyllabic minimum). The data below is from Hale 1973: 421, 424. The examples in (85)a show final vowel deletion in stems longer than two syllables, while (85)b shows that deletion does not apply to vowel-final disyllabic stems.

(85)	a. /yalulu/	→	yalul	'flame'
	/mayara/	→	mayar	'rainbow'
	/karikari/	→	karikar	'butter-fish'
	/kaŋkaṛi/	→	kaŋkaṛ	'father's father'
	b. /paṛŋa/	→	paṛŋa	'stone'
	/kela/	→	kela	'beach'
	/wanka/	→	wanka	'arm'
	/ṭarŋka/	→	ṭarŋka	'barracuda'

An example of syncope being blocked by a bimoraic word minimum is Ojibwe (Nishnaabemwin) (Piggott 1991, Valentine 2001), where final vowels fail to delete just in case the resulting word would have fewer than two moras.

This behavior is just what is predicted by the theory: epenthesis is driven by ordinary OT constraints on syllable structure and sonority sequencing, and these constraints can block the effect of the ✦SYNCOPE constraints in the ordinary fashion.

CONCLUSION

This thesis has addressed a recurring difficulty in OT phonology, the too-many-solutions problem. I have argued that the diagnosis of the problem has to do with the locus of phonologically significant generalizations. Contrary to standard OT thinking, at least some important generalizations are best stated not as surface conditions, but as conditions on input-output mappings and environments of processes.

The thesis is an exploration of what it would take for a pure-markedness approach to the problem. I have argued that, in order to handle non-surface generalizations, markedness constraints must be modified quite radically. Procedural constraints, whose job is to penalizes candidates involving certain PROCESSES, must have access to the rest of the ranking of the language. I have proposed a formal mechanism that allows such constraints to function, and showed how the addition of this new class of constraints can greatly restrict the typology of phonological processes, at least in two domains: prosody-segmental interactions, and vowel syncope. My theory of procedural constraints allows for a direct control over processes and environments in OT, thus bringing generalizations about input-output mappings within the scope of the theory.

REFERENCES

Alber, Birgit & Ingo Plag (1999) Epenthesis, Deletion and the Emergence of the Optimal Syllable in Creole. ROA-335.

Alderete, John (2006[1995]) Winnebago Accent and Dorsey's Law. ROA-860. Appeared originally in: Jill N. Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk (eds.) (1995). *Papers in Optimality Theory*, *University of Massachusetts Occasional Papers 18*. Amherst, MA: GLSA. Pp. 21-51.

Allen, George D. (1975). Speech rhythm: its relation to performance universals and articulatory timing. *Journal of phonetics* 3: 75-86.

Al-Mozainy, Hamza Qublan, Robert Bley-Vroman, & John J. McCarthy (1985). Stress shift and metrical structure. *LI* 16(1): 135-144.

Bach, Emmon and Robert T. Harms (1972). How do languages get crazy rules? *Linguistic change and generative theory*, ed. by Robert P. Stockwell and Ronald K.S. Macaulay, 1-21. Bloomington: Indiana University Press.

Bell, Alan (1977). Accent placement and perception of prominence in rhythmic structures. *Studies in stress and accent*, ed. by Larry M. Hyman, 1-13. (Southern California Occasional Papers in Linguistics, 4). Los Angeles: Department of Linguistics, USC.

Blevins, Juliette (1985). *A Metrical Theory of Syllabicity*. Ph.D. Dissertation, MIT.

Blevins, Juliette (2004). *Evolutionary phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.

Blevins, Juliette and Andrew Garrett (1998). The origins of consonant-vowel metathesis. *Language* 74: 508-556.

Bloomfield, Leonard (1961[1933]). *Language*. New York: Holt & Co.

Blumenfeld, Lev (2004). Tone-to-stress and stress-to-tone: Ancient Greek accent revisited. *BLS* 30.

Blumenfeld, Lev (to appear). Plautine trochees. Submitted ms.

Borkovsky, V.I. and P.S. Kuznetsov (1965). *Istoričeskaja grammatika russkogo jazyka*. Moskva: Nauka.

Broselow, Ellen (1982). On predicting the interaction of stress and epenthesis. *Glossa* 16(2): 115-132.

Bybee, Joan, et al. (1998). Prosody and segmental effect: Some paths of evolution for word stress. *Studies in Language* 22. 267-314.

Christmas, R. & J. Christmas (1975). Kupia phonemic summary, Kathmandu: SIL and Institute of Nepal and Asian Studies, Tribhuvan University.

Collins, B. & I. Mees (1990). The phonetics of Cardiff English. In: English in Wales: Diversity, conflict and change. Nikolas Coupland, ed. 87-103. Clevedon: Multilingual Matters.

Crosswhite, Katherine (2001). *Vowel reduction in Optimality Theory*. New York: Routledge.

Cutler, Anne (1980). Errors of stress and intonation. In: Victoria A. Fromkin, ed., *Errors in linguistic performance*. New York: Academic Press, pp. 67-80.

Cutler, Anne and Wilma van Donselaar (2001). *Voornaam* is not (really) a homophone: lexical prosody and lexical access in Dutch. *Language and speech* 44(2): 171-195.

Davis, Stuart. (1988). Syllable onsets as a factor in stress rules. *Phonology* 5:1, 1-19.

Dayley, Jon P (1989). *Tümpisa (Panamint) Shoshone grammar*. (*University of CA Publications in Linguistics* v. 115). Berekeley: U. of California Press.

De Lacy, Paul (2002). *The formal expression of markedness*. PhD Thesis, UMass.

De Lacy, Paul (2002). The interaction of tone and stress in Optimality Theory. *Phonology* 19: 1-32.

De Lacy, Paul (2003). Fixed ranking and the 'Too Many Solutions' problem. CASTL Kick-off conference handout.

Dixon, R.M.W. (1988). *A grammar of Boumaa Fijian*. Chicago: The University of Chicago Press.

Dixon, R.M.W. (2004). *The Jarawara language of Southern Amazonia*. With the assistance of Alan R. Vogel. Oxford: Oxford University Press.

Dresher, B.E. & A. Lahiri (1991). The Germanic Foot: metrical coherence in Germanic. *LI* **22**(2): 251-286.

Eekman, Thomas (1974). *The realm of rime: a study of rime in the poetry of the Slavs*. Amsterdam: Adolf M. Hakkert, 1974.

England, Nora (1983). *A grammar of Mam, a Mayan language*. Austin: University of Texas Press.

Everett, Daniel (1988). On metrical constituent structure in Pirahã phonology. *NLLT* 6: 207-46.

Everett, Daniel and Keren Everett (1984). Syllable onsets and stress placement in Pirahã. WCCFL 3: 105-116.

Fay, David and Anne Cutler (1977). Malapropisms and the structure of the mental lexicon. *LI* 8(3): 505-520.

Flemming, Edward (1994). The role of metrical structure in segmental rules. Ms., UCLA.

Fraisse, Paul and Geneviève Oléron (1954). La structuration intensive des rythmes. *L'année psychologique* 1954(1): 35-52.

Fraisse, Paul and Geneviève Oléron (1954). La structuration intensive des rythmes. *L'année psychologique* 1954(1): 35-52.

Gabas, Nilson (1998). *Estudo fonológico da língua Karo (Arara de Rondônia)*. München: Lincom Europa.

Goedemans, Rob (1996). An optimality account of onset sensitivity in QI languages. ROA.

González, Carolina (2001). The effect of prosody on glottal stop deletion in Capanahua. *NELS* 32.

González, Carolina (2003). *The effect of stress and foot structure on consonantal processes*. PhD Thesis, University of Southern California.

Gordon, Matthew (1999). *Syllable weight: phonetics, phonology, and typology*. PhD Thesis, UCLA.

Gordon, Matthew (2001). Syncope induced metrical opacity as a weight effect. In: K. Megerdoomian and L.A. Bar-el, *WICCFL 20 Proceedings*. Sommerville, MA: Cascadilla Press. 206-219.

Gouskova, Maria (2003). *Deriving Economy: Syncope in Optimality Theory*. PhD Thesis, UMass Amherst.

Green, Antony Dubach (1997) The Prosodic Structure of Irish, Scots Gaelic, and Manx. ROA-196.

Gregores, Emma and Jorge A. Suárez (1967). *A description of colloquiam Guaraní*. Janua linguarum, series practica 27. The Hague: Mouton.

Hall, Nancy (2003). *Gestures and segments: vowel intrusion as overlap*. PhD Thesis, UMass Amherst.

Halle, Morris and Jean-Roger Vergnaud (1987). *An essay on stress.* Cambridge, Mass.: MIT Press.

Halle, Morris and William Idsardi (1995). General properties of stress and metrical structure. In: John A. Goldsmith, ed., *The handbook of phonological theory*, 403-444. Oxford: Blackwell.

Hammond, Michael (1984). *Constraining metrical theory: a modular theory of rhythm and destressing*. PhD Thesis, UCLA.

Hayes, Bruce P. (1981). *A Metrical Theory of Stress Rules.* Ph.D. Dissertation. MIT.

Hayes, Bruce P. (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.

Hdouch, Youcef (2004) Phonological and Morphological Issues and Schwa Epenthesis in Berber. Rutgers Optimality Archive #696

Hoijer, Harry (1938). Tonkawa: an Indian language of Texas. In: Franz Boas, ed., *Handbook of American Indian languages*. Columbia University Press.

Itô, Junko (1986). *Syllable theory in prosodic phonology*. PhD Thesis, University of Massachussetts-Amherst.

Itô, Junko (1989). A prosodic theory of epenthesis. *NLLT* 7: 217-259.

Jakobson, Roman (1963). Opyt fonologičeskogo podxoda k istoričeskim voprosam slavjanskoj akcentologii: pozdnij period slavjanskoj jazykovoj praistorii. In: *American contributions to the Fifth International Congress of Slavists, Sofia, September 1963*. The Hague: Mouton. 153-178.

Jeanne, LaVerne Masayesva (1982). Some phonological rules of Hopi. *IJAL* 48(3): 245-270.

Kager, René (1997). Rhythmic vowel deletion in Optimality Theory. In: Iggy Roca, ed., *Derivations and constraints in phonology*. Oxford: Clarendon Press, 463-500.

Kariņš, A. Krišjānis (1996). *The prosodic structure of Latvian*. PhD Thesis, University of Pennsylvania.

Kenstowicz, Michael (1994). *Phonology in generative grammar*. Oxford: Blackwell.

Kenstowicz, Michael (1996). Sonority-driven stress. ROA-33.

Kenstowicz, Michael (2003). Salience and Similarity in Loanword Adaptation: A Case Study from Fijian. ROA-609.

Kiparsky, Paul (2000). Opacity and cyclicity. *Linguistic review* 17: 351-367.

Kiparsky, Paul (2004). *Universals constrain change, change results in typological generalizations*. Ms., Stanford University.

Kiparsky, Valentin (1979). *Russian historical grammar*. Ann Arbor: Ardis.

Kisseberth, Charles W. (1970). On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291-306.

Kisseberth, Charles W. (1970). Vowel elision in Tonkawa and derivational constraints. In: Jerrold M. Sadock and Anthony L. Vanek, eds., *Studies presented to Robert B. Lees by his students*. Edmonton: Linguistic Research, 109-137.

Lavoie, Lisa M. (2001). *Consonant strength: phonological patterns and phonetic manifestations*. Outstanding dissertations in linguistics. New York: Garland.

Leumann, Manu (1977). *Lateinische Laut- und Formenlehre*. München : C. H. Beck.

Levelt, Willem J.M. (1989). *Speaking: from intention to articulation.* Cambridge, MA: MIT Press.

Liberman, Mark (1975). *The intonational system of English*. PhD Thesis, MIT.

Liberman, Mark and Alan Prince (1977). On stress and linguistic rhythm. *LI* 8: 249-336.

Lloret, Maria-Rosa & Jesús Jiménez (2006). Prominence-driven epenthesis in Alguerese Catalan. ROA-806

Lombardi, Linda (2001[1995]). Why place and voice are different: constraint-specific alternations in Optimality Theory. In: Linda Lombardi, ed. *Segmental phonology in Optimality Theory*, 13-45. Cambridge: Cambridge UP.

Loos, Eugene E. (1967). *The phonology of Capanahua and its grammatical basis*. PhD Thesis, University of Texas (Austin).

Loos, Eugene E. (1986). Capanahua syllable restructuring in relation to comparative studies. In: Benjamin Elson, ed., 1986: 291-301.

McCarthy, John (2003). Comparative Markedness. *Theoretical linguistics* 29: 1-51.

McCarthy, John J. (1979). On stress and syllabification. *LI* 10(3): 443-465.

McCarthy, John J. (2002). On targeted constraints and cluster simplification. *Phonology*19: 273-292.

Mills, Elizabeth (1984). *Senuofo phonology, discourse to syllable*. UT Arlington: SIL.

Miner, Kenneth (1979). Dorsey's Law in Winnebago-Chiwere and Winnebago accent. *International journal of American linguistics* 45: 25-33.

Nivens, R. (1992). A lexical phonology of West Tarangan. In: D.A. Burquest & W.D. Laidig (eds.), Phonological studies in four languages of Maluku. Arlington, TX: SIL and U of Texas at Arlington publications in linguistics, 127-227.

Olson, Kenneth S. (2001). *The phonology and morphology of Mono*. PhD Thesis, University of Chicago.

Parker, Steve (1998). Disjoint metrical tiers in Huariapano. University of Massachusetts-Amherst. Ms.

Pater, Joe (2003). Balantak metathesis and theories of possible repair in Optimality Theory. Ms., UMass Amherst.

Patz, E (1991). Djabugay. In: Dixon and Blake, 1991, *Handbook of Australian Languages* v. 4. Oxford: OUP, 245-348.

Payne, Doris (1981). *The phonology and morphology of Axininca Campa*. Arlington, TX: Summer Institue of Linguistics and University of Texas.

Peperkamp, Sharon (2004). Lexical exceptions in stress systems: arguments from early language acquisition and adult speech perception. *Language* 80(1): 98-126.

Peperkamp, Sharon and Emmanuel Dupoux (2002). A typological study of stress 'deafness'. In: Carlos

Pons Mall, Claudia (2005). It is all downhill from here: the role of syllable contact in Romance languages. 13th Manchester Phonology Meeting.

Prince, Alan (1983). Relating to the grid. *LI* 14(1): 19-100.

Prince, Alan and Paul Smolensky (1993). Optimality Theory: constraint interaction in generative grammar. ROA.

Rawlins, Kyle (2006). Stress and epenthesis in Mohawk. Ms., UC Santa Cruz.

Safir, Ken (1979). Metrical structure in Capanahua. *MITWPL 1*: 95-114.

Samek-Lodovici, Vieri and Alan Prince (1999). Optima. ROA 363.

Schlie, Perry and Ginny Schlie (1993). A Kara phonology. In: Clifton 1993: 99-130.

Selkirk, Elizabeth (1981). Epenthesis and degenerate syllables in Cairene Arabic. *MITWPL* 3: 209-32.

Sihler, Andrew (1995). *New comparative grammar of Greek and Latin*. Oxford: Oxford University Press.

Smith, Jennifer (2005). *Phonological augmentation in prominent positions.* N.Y.: Garland.

Steriade, Donca (2001). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Ms., UCLA/MIT.

Taylor, Moira (1994). *Interaction of vowel deletion and syllable structure constraints*. PhD Thesis, Simon Fraser University.

Tesar, Bruce, and Paul Smolensky (2000). *Learnability in OT*. Cambridge, Mass.: MIT Press.

Valentine, J. Randolph (2001). *Nishnaabemwin reference grammar*. Toronto: University of Toronto Press.

Vaysman, Olga (2002). Against richness of the base: evidence from Nganasan. *BLS* 28.

Vaysman, Olga (2005). Foot structure and prominence: a case of mismatch. Handout, BLS 31.

Vendryes, Joseph (1902). *Recherches sur l'histoire et les effets de l'intensité initiale en Latin*. Paris: Librairie C. Klincksieck.

Vijayakrishnan, K.G. (1999). Weakening Processes in the Optimality Framework. ROA-329.

Volker, Craig Alan (1998). *The Nalik language of New Ireland, Papua New Guinea*. NY: Peter Lang.

Walker, Rachel (1997). Mongolian stress, licensing, and factorial typology. ROA-172.

Walker, Rachel (2004). Vowel feature licensing at a distance: evidence from Northern Spanish language varieties. *WCCFL* 23: 787-800.

Wilson, Colin (2001). Consonant cluster neutralisation and targeted constraints. *Phonology* 18: 147-197.

Yar-Shater, E. (1969). A grammar of Southern Tati Dialects, Columbia University Publications in Near and Middle Eastern Studies, series B3. The Hague: Mouton.

Zalizniak, A.A. (1985). *Ot praslavianskoi akcentuacii k russkoi*. Moskva: Nauka.