

Constraint interaction in language change: quantity in English and Germanic

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy in the Faculty of Arts

May 1999

Ricardo Bermúdez-Otero

Department of English and American Studies

Contents

<i>Abstract</i>	6
<i>Declaration</i>	7
<i>Notes on copyright and intellectual property rights</i>	8
<i>Dedication and acknowledgements</i>	9
<i>The author</i>	11
<i>Abbreviations, notation and nomenclature</i>	12
<i>I</i>	
Introduction	16
<i>Chapter 2</i>	
Morae in Optimality Theory	22
2.1 Introduction	22
2.2 The moraic representation of length and quantity	23
2.2.1 Weightless codas	23
2.2.2 Segmental length contrasts and Lexicon Optimization	29
2.2.2.1 Richness of the Base and Lexicon Optimization	29
2.2.2.2 Vowels	31
2.2.2.3 Long consonants	31
2.2.2.4 Short consonants	34
2.3 Issues in moraic faithfulness	36
2.3.1 Positional μ -licensing	36
2.3.2 First anomalous prediction	39
2.3.3 Second anomalous prediction	44
2.3.4 A solution	47

2.3.4.1	DEP [#] and IDENT [#] reformulated	47
2.3.4.2	Strict Surface Orientation and relativized faithfulness	50
2.4	Conclusion	55
<i>Chapter 3</i>		
	Opacity	56
3.1	Introduction	56
3.2	Definitions of opacity: intension and extension	58
3.2.1	Intensional definitions: opacity as misapplication	58
3.2.2	Two optimality-theoretic strategies against opacity	62
3.2.2.1	Upholding Strong Parallelism	62
3.2.2.2	Upholding Strict Surface Orientation	64
3.2.3	The extension of opacity in Optimality Theory	65
3.2.4	Conclusion	69
3.3	Optimality Theory with interleaving	71
3.3.1	Conceptual issues	71
3.3.1.1	Interleaving, level segregation, cyclicity	71
3.3.1.2	The arbitrariness of cophonologies	75
3.3.1.3	Derivational vs declarative implementations of interleaving	77
3.3.2	The place of articulation of coda nasals in Spanish	79
3.3.3	Restrictions on opacity in interleaved Optimality Theory	93
3.3.3.1	Internal restrictions: the architecture of the model	93
3.3.3.2	External restrictions: Lexicon Optimization in acquisition and change	98
3.3.3.3	Synchronic markedness reversal	104
3.4	Strongly parallel Optimality Theory	107
3.4.1	Paradigm effects	110
3.4.1.1	Alignment Theory	110
3.4.1.2	The priority of the base	113

3.4.1.3	Masking	125
3.4.1.4	Summary of paradigm effects	136
3.4.2	Non-paradigmatic opacity effects: sympathy	137
3.4.2.1	Sympathy Theory	137
3.4.2.2	★-Confinement and extended sympathy	150
3.5	Opaque mora preservation	158
3.5.1	Opacity in Hayes's canon of mora-preserving lengthening effects	158
3.5.2	West Germanic Gemination	164
3.5.2.1	Description	164
3.5.2.2	Common Germanic syllabification	166
3.5.2.3	The rôle of CONTACT	168
3.5.2.4	Opacity	169
3.5.2.5	Sievers' Law	175
3.5.2.6	Summary	181
3.5.3	Peripeteia: free reranking upheld, ★-Confinement refuted	182
3.6	Epilogue: where the moral of the tale is drawn	192
<i>Chapter 4</i>		
	Globality	197
4.1	Global effects in phonological change	197
4.1.1	Is sound change so blind after all?	197
4.1.2	Globality without top-down selection	200
4.1.3	Top-down selection of innovations in interleaved OT	203
4.1.4	Preview of the chapter	206
4.2	Middle English quantity: empirical issues	207
4.2.1	MEOSL: mora epenthesis or mora conservation?	207
4.2.1.1	The traditional account	207
4.2.1.2	Minkova's revision	209

4.2.2	The circularity of TRISH	211
4.2.3	Middle English lexical phonology: SHOCC in the <i>Ormulum</i>	212
4.2.4	Lengthening in unapocopated disyllables	215
4.2.5	Middle English in a West Germanic perspective	219
4.2.6	Summary	220
4.3	Globality in MEOSL	221
4.3.1	Teleology	221
4.3.1.1	Luick's <i>Normalmaße</i>	222
4.3.1.2	Lass: the Great Length Conspiracy	223
4.3.2	SHOCC	225
4.3.3	The global interaction of schwa loss and consonant extrasyllabicity	229
4.4	Globality in lexical change and stem-level defaults	233
4.5	Summary	237
5		
	Prospects	239
	<i>Appendix A</i>	
	Morphological conditioning of SHOCC in the <i>Ormulum</i>	241
	<i>Appendix B</i>	
	Unapocopated disyllables and MEOSL	247
	<i>References</i>	249

Abstract

This dissertation addresses the questions raised by two opposite types of interaction between phonological processes: opacity and globality. Special consideration is given to their dual nature as synchronic instantiations of UG capabilities and as products of language change. It is argued that, in order to achieve a synthesis, phonological theory must combine constraint ranking with an interleaved organization of the grammar.

Opacity is described, in Lightfoot's metaphor, as a 'Rube Goldberg feature of grammars': it is not catered for by specific UG resources, nor is it functionally grounded; rather, it develops contingently through the impact of misacquisition phenomena upon the interface between phonology and the rest of the grammar. Stochastic processes of hypocorrection give rise to surface-true phonological innovations represented at the postlexical level; these may interact opaquely with generalizations operating at higher grammatical levels. As innovations evolve, they tend to percolate upwards in the grammar; Lexicon Optimization drives this life cycle by prompting the restructuring of the input to a phonological level when relevant cues are attenuated in the child's trigger experience.

In strictly parallel versions of Optimality Theory, in contrast, opacity is regarded as the province of functionally grounded transderivational correspondence constraints. This approach to opacity is subjected to a comprehensive critique. In particular, the evidence of a process of gemination in West Germanic is shown to contradict the claim that, in interleaved models of grammar, the ranking of markedness constraints is always fixed across levels. West Germanic Gemination also refutes the assumption that sympathy-candidate selectors are confined to the set of IO-faithfulness constraints.

In turn, it is suggested that a subset of diachronic global effects involves the parallel interaction of old and innovative constraint rankings within the same phonological level. The implications of this proposal are illustrated with a reanalysis of the Middle English quantity adjustment which dispels the illusion of conspiracy conjured up by previous approaches.

Both case studies in quantity change are informed by a reformulation of constraints on moraic faithfulness. This overcomes a number of difficulties incurred by the current statements of DEP^u and IDENT^u in relation to Weight by Position.

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged in the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the Head of the Department of English and American Studies.

A mis padres

This dissertation would not have been possible without the dedicated support of Chris McCully at the University of Manchester and Teresa Fanego at the University of Santiago de Compostela. They have negotiated the intricacies of my somewhat quirky affiliation to both institutions with patience, skill and tact. To this they have both added extravagant measures of encouragement.

Whether in person, or by post or e-mail from UCLA, Donka Minkova was a constant benevolent presence during the gestation of this research. Needless to say, chapter 4 would not have existed but for her ideas. From her example I hope to have learnt a way of doing historical linguistics that combines boldness, philological responsibility and theoretical engagement.

I witnessed these virtues profusely displayed at the XII International Conference on Historical Linguistics (ICHL) and the X International Conference on English Historical Linguistics (ICEHL), both of which were held at Manchester. I am grateful to David Denison, Richard Hogg and Nigel Vincent for making Manchester an international centre of historical linguistic research.

It was at X ICEHL that I obtained copies of Kiparsky (1998a, b); I am deeply grateful to Paul Kiparsky for letting me have them. The experience of reading both manuscripts was a timely *coup de foudre*. It made me readdress issues from which I had withdrawn in perplexity and frustration, and it led me, at the eleventh hour, to rewrite no less than three quarters of this dissertation. I hope that, after the agony and the haste, the result does not reflect discredibly on his inspiration.

I have inflicted drafts of various portions of this thesis upon a number of patient conference audiences. Their apposite questions or comments have often steered me in the right direction. Among those who have contributed in this way I am pleased to thank Mike Davenport, Elan Dresher, Daniel Everett, S. J. Hannahs, John Harris, René Kager, Aditi Lahiri, Roger Lass, John McCarthy, Glyne Piggott, Niki Ritt and Bob Stockwell. Niki Ritt was kind enough to engage in a protracted e-mail exchange concerning Middle English quantity. I am also grateful for tips from Wiebke Brockhaus, Martin Durrell and Richard Hogg.

The Rutgers Optimality Archive has proved an invaluable research tool to a European postgraduate whose limited means have denied him access to the American conference circuit. Where once one would have spent weeks or months on the uncertain trail of unpublished manuscripts kept at US universities, the Archive has brought immediate access and *glasnost* world-wide. The founders and keepers of ROA are to be commended.

To John Hutton I am indebted both as a phonologist and as a friend. I also thank Linda van Bergen and Susan Rosser for making the English Department a jolly place to work at, and for becoming friends.

Those who have supported me with their love already know they have mine.

The research reported in this dissertation was supported until December 1998 with a scholarship from the Spanish Ministry of Education and Science (Subprograma de Formación de Profesorado Universitario en el Extranjero).

The author

Ricardo Bermúdez-Otero obtained the degree of *Licenciado en Filología Inglesa* from the University of Santiago de Compostela in July 1993. In 1996 he was awarded the degree of Master of Philosophy with distinction by the University of Manchester. He has published in the journal *English Language and Linguistics* (Bermúdez-Otero 1998a) and is co-editor of the forthcoming proceedings of X ICEHL. At the University of Manchester he has lectured in phonology for the Department of Linguistics.

Abbreviations, notation and nomenclature

1	first person
2	second person
3	third person
acc.	accusative
c.	century
C	consonant
dat.	dative
dim.	diminutive
gen.	genitive
(W)Gmc.	(West) Germanic
ind.	indicative
IO	input-output
IG	intonation group
LPM	Lexical Phonology and Morphology
(e)ME	(early) Middle English
MEOSL	Middle English Open Syllable Lengthening
(e)ModE	(early) Modern English
nom.	nominative
O	obstruent
OCP	Obligatory Contour Principle
(l)OE	(late) Old English
OHG	Old High German
OO	output-output
OS	Old Saxon
OT	Optimality Theory
OUpG	Oldest Upper German
PDE	Present-day English
pl.	plural
pres.	present
pret.	preterite

R	liquid consonant
S	sonorant consonant
SBPM	Sign-Based Phonology and Morphology
sg.	singular
SHOCC	Shortening before Consonant Clusters
SPE	Chomsky & Halle (1968)
SR	surface representation
str.	strong
subj.	subjunctive
TCT	Transderivational Correspondence Theory
TRISH	Trisyllabic Shortening
UG	Universal Grammar
UR	underlying representation
V	vowel
wk.	weak
•	root node
μ	mora
μ _p	positional μ-licenser
σ	syllable
σ̄	light syllable
σ̄	heavy syllable
Σ	foot
ω	prosodic word
*	ungrammatical
†	reconstructed
→	maps onto
>	historically becomes
»	is ranked higher than
«	is ranked lower than
⋄	is more harmonic than
⋄	is less harmonic than
∧	Smolensky-conjunction

■[...]	(opaque) winning candidate
□[...]	transparent losing candidate
⊗	sympathy / sympathetic
⊗O	sympathetic candidate - output
★	sympathy candidate selector
N _c	sympathy candidate selected by constraint C
● [✱]	counterfactually winning candidate
⊗	counterfactually losing candidate

Violable constraints are labelled in small capitals: e.g. FTBIN. Other grammatical statements, such as non-violable principles, are labelled in normal type: e.g. Proper Headedness.

Transcriptions follow IPA conventions unless otherwise stated.

The term 'morphosyntax' is used in a wide sense, denoting both morphology and syntax. Its narrow sense, in which it denotes those syntactic phenomena whose expression is morphological (e.g. agreement), will not be relevant.

History matters.

1

Introduction

Lightfoot (1999: 111) draws an apt analogy between grammars and the machines imagined by engineer-turned-cartoonist Rube Goldberg. Goldberg's contraptions attain their goals in comically circuitous ways: Lightfoot cites the case of an automatic stamp-licker assembled from a miniature robot, ants, and a hungry ant-eater. If this is madness, yet there is method in it; as Lightfoot points out, the machines comply with all known laws of physics. The serious point of Lightfoot's comparison is that grammars, whilst being exquisitely patterned designs attesting to the power of the language faculty, often have odd features which arise through the accidents of historical change:

He [sc. Rube Goldberg's character Professor Lucifer Gorgonzola Butts] never designed grammars, as far as I know, but grammars have their Rube Goldberg properties. It is true that grammars are formed in a child in accordance with the prescriptions of the linguistic genotype; consequently, they are economical and beautiful in ways imposed by UG. However, they are also shaped by the demands of haphazard experience, and this makes them historically contingent. There is no contradiction here; after all, Rube Goldberg machines conform to the usual laws of mechanics. But there is a tension.

Lightfoot (1999: 111)

In phonology, this tension is nowhere better illustrated than in the uneasy coexistence of opaque and global rule interactions. On the one hand, phonological alternations often apply in a concerted fashion, subserving some overarching constraint. In such phenomena we discern the order imposed by the hierarchical tiered structure of phonological representations. On the other hand, alternations just as often trip each other up: rules are found to fail in circumstances where some other process creates their conditioning environment or, conversely, they apply even where another generalization removes their trigger. Opaque interactions look odd, but, as our quotation from Lightfoot suggests, it is easy to understand them as the product of contingent historical events.¹ Children often misanalyse automatic phonetic effects as phonological alternations controlled by the target

¹ Lightfoot (1999: §5.4) gives ergative splits as an example of a Rube Goldberg property in syntax.

system; as these innovative surface-true generalizations enter the grammar, they often disfigure the output of previous alternations. In understanding why grammars are the way they are, one must therefore never neglect their historical dimension.

The tension between globality and opacity has made itself felt from the early days of generative phonology (see e.g. Kisseberth 1970 and Kiparsky 1971), but the relative emphasis accorded to each has since shifted dramatically. The framework of SPE, with its battery of extrinsically ordered linear rewrite rules, was eminently suited to the expression of opacity, and allowed grammars to recapitulate history to a potentially unlimited extent. As the theory of representations grew richer and more articulate during the 1980's, so the phonologist's ability to impose global conditions on derivations was extended. During the 1990's, with OT taking over the generative mainstream, the pendulum has swung to the opposite extreme. Harmonic evaluation directly captures the driving rôle of representational constraints; but in place of opacity the outline of the theory drawn by Prince & Smolensky (1993) leaves a descriptive gap. If one assumes that the mapping between underlying and surface representations is effected in a single step, and that output candidates are evaluated solely by markedness and IO-faithfulness constraints, then a large class of opaque phonological phenomena becomes rigorously intractable. Opacity is therefore back on the agenda, its very nature and its implementation in constraint-based grammars open to discussion. This dissertation addresses the current debate in OT; lying at the core of my contribution is a vision of the historical processes which fuel the rise and fall of opaque phenomena.

The problem of opacity has given rise to two competing research programmes within OT:

(i) Scholars working in the tradition of Lexical Phonology (broadly defined to include both its derivational and its representational interpretations) tackle opacity by means of so-called 'interleaving' mechanisms: stratification and cyclicity. In essence, the phonology is assumed to apply recursively over a hierarchy of progressively larger domains defined by morphological and syntactic structure. Domains of different types (usually stems, words and phrases) are assigned to different cophonologies, characterized by different hierarchizations of CON. This approach is represented by Lexical Phonology and Morphology in Optimality Theory (henceforth LPM-OT; Kiparsky 1998a) and by Sign-Based Phonology and Morphology (henceforth SBPM; Orgun 1996a, b).

(ii) Phonologists pursuing a more radical version of OT's parallelist agenda challenge the necessity and adequacy of interleaving. They propose that the input-output mapping is effected by means of a single pass through GEN and EVAL, with misapplication effects arising from multiple relationships of transderivational correspondence between the output and other surface forms (real or virtual). This strongly parallelist programme is embodied in the OO-correspondence theory of paradigm effects (Benua 1995, 1997; Itô & Mester 1997a; Kager 1995; Kenstowicz 1996) and in Sympathy Theory (McCarthy 1998).

These two approaches are evaluated in chapter 3. I provide evidence, both synchronic and diachronic, which undermines the strongly parallelist programme and supports a combination of harmonic evaluation with interleaving. The argument will be built around two case studies:

(i) In Spanish, resyllabification into the onset counterbleeds processes assigning default place specifications to unassimilated coda nasals. This counterbleeding effect is recursive, as stem-final nasals are resyllabified before word-level suffixes, and word-final nasals are resyllabified postlexically before vowel-initial words. Crucially, so-called 'velarizing' dialects select different defaults at the stem and the word levels. The upshot will be shown to be twofold. First, OO-correspondence is unable to account for the overapplication of default place assignment to the stem-final nasal of the plural noun [ðes.ˈðe.nes] 'disdain', which derives from the root /desdeɲ-/; the alveolar nasal of the plural is found nowhere else in the paradigm, as the singular [ðes.ðeŋ] shows the word-level default. Solutions based on alignment or sympathy prove descriptively inadequate, as they miss the links with default place assignment in non-resyllabification environments. Secondly, Spanish provides our first counterexample to Benua's (1997) claim that the ranking of markedness constraints remains fixed across levels; Benua's claim underpins a charge of overgeneration against interleaved OT.

(ii) My second case study is diachronic. In West Germanic, underlyingly short consonants geminated before [j] and before the liquids /r l/: e.g. /bid-I-an/ → [bid.djan] 'ask' (Gothic *bidjan*; OS *biddien*, OE *biddan*, OHG *bitten*). The historical record indicates that gemination was an opaque mora-conserving process involving the overapplication of Weight by Position. The clusters where gemination occurred had been heterosyllabic in Common Germanic: e.g. [bid^h.jan]. In West Germanic, their anomalous sonority profile was repaired by attracting the first consonant into the onset of the following syllable. Crucially,

the erstwhile coda consonant retained its moraic attachment despite no longer meeting the conditions for Weight by Position. Again, this phenomenon will illustrate how phonological change may give rise to synchronic markedness reversal between levels. More importantly, it refutes McCarthy's (1998) claim that selector constraints in Sympathy Theory are drawn only from the set of IO-faithfulness constraints: in a sympathy-theoretic analysis of West Germanic Gemination, O^* -candidates are selected by markedness constraints that penalize complex onsets. This would appear to favour Extended Sympathy Theory (Itô & Mester 1997b, de Lacy 1998) over McCarthy's (1998) original formulation.

I shall conclude that, at a deep conceptual level, strongly parallel OT cannot make sense of the existence of opacity. The theory offers no internal explanation, for the principles governing the selection of O^* -candidates are axiomatic and cannot be deduced from more fundamental postulates. Yet, at the same time, Extended Sympathy Theory defeats the hope that O^* -correspondence might be provided with functional grounding in lexical access: when a O^* -candidate is selected by a subordinate markedness constraint, it need not be more faithful to the input than the optimal output, in which case O^* -correspondence will do nothing to enhance the recoverability of the input. Furthermore, I shall argue that neither OO- nor O^* -correspondence can draw support from the success of BR-identity: BR-identity is a symmetrical syntagmatic relationship, whilst both OO- and O^* -correspondence are transderivational and require stipulative anti-symmetry devices.

Under interleaving, in contrast, opacity can be understood as a 'Rube Goldberg feature of grammars': it is not catered for by specific UG resources, nor is it functionally grounded; rather, it develops contingently through the impact of misacquisition phenomena upon the interface between phonology and the rest of the grammar. Stochastic processes of hypocorrection give rise to surface-true phonological innovations represented at the postlexical level; these may interact opaquely with generalizations holding at higher grammatical levels. The architecture of the system sets strict boundaries to the complexity of opaque interaction, for opacity arises between levels, but never within a single input-output mapping. Beyond these formal limitations, the accretion of opaque effects is also kept in check by contingent diachronic developments. In time, innovative constraint rankings tend to percolate upwards in the grammar; this occurs as Lexicon Optimization prompts the restructuring of the input to a phonological level when the evidence of alternation is critically attenuated in the child's trigger experience. In sum, interleaved OT

provides a non-stipulative approach to opacity that succeeds in integrating the synchronic and diachronic aspects of the problem.

In turn, chapter 4 will show that interleaved OT retains the ability to account for global effects. In a diachronic setting, globality often arises as the outcome of an innovative reranking is determined in parallel interaction with constraint rankings which already hold at the same level. I illustrate this proposal with an analysis of vowel length change in late Old and early Middle English (henceforth IOE and eME). I adduce new evidence (notably from the morphophonology of the *Ormulum*) to corroborate Minkova's (1982) reanalysis of so-called ME Open Syllable Lengthening as a compensatory process triggered by the loss of stem-final schwa. I will show that long established rankings in the lexical phonology demanded the conservation of the mora associated with the apocopated schwa; these rankings favoured final consonant extrasyllabicity, and were active in exempting -VVC monosyllables from closed syllable shortening.

Another form of globality is involved in lexical diffusion. I shall argue that lexically selective processes of vowel length change in ME propagated underlying specifications which minimized the infringement of high-ranking markedness constraint at the stem level. This proposal preserves Kiparsky's (1995) insight that diffusing changes are grammar-dependent, but replaces *ad hoc* lexical rules with independently motivated constraint hierarchies.

Two of the case studies from which my arguments draw empirical support involve syllable quantity and, more specifically, phenomena of mora conservation: namely West Germanic Gemination (chapter 3) and so-called ME Open Syllable Lengthening (chapter 4). Since the implementation of Mora Theory in an optimality-theoretic framework raises some non-trivial difficulties, notably in relation to Weight by Position, the dissertation opens with some ground-clearing discussion in chapter 2. In particular, I shall argue that faithfulness constraints on mora structure cannot penalize Weight-by-Position morae; otherwise, bizarre typological predictions emerge. I solve this problem by formally defining a relationship of positional μ -licensing, which holds between a mora μ and a segment α when α is non-moraic in the input and μ is the sole prosodic licenser of α in the output. DEP^μ and IDENT^μ are rendered blind to positional μ -licensing by means of an operation of local conjunction.

The conclusion briefly explores the new avenues which interleaved OT opens to research in historical phonology.

Chapter 2

Morae in Optimality Theory

2.1 INTRODUCTION

This dissertation contends that, in order to account satisfactorily for both opacity and globality effects, whether from a synchronic or from a diachronic point of view, phonological theory must combine interleaving (following LPM and SBPM) with constraint ranking (as developed in OT). Two of the case studies which I set forth in support of this argument involve segmental length change in the Germanic family: in chapter 3, evidence from West Germanic Gemination is used to refute some central claims of Sympathy Theory, which advocates a strongly parallel approach to opaque phonological phenomena; in chapter 4, I show that constraint ranking is necessary to capture the global connections between the several processes of change making up the ME vowel length adjustment. In both cases, my approach to the evidence is crucially based on Mora Theory. It turns out, however, that the implementation of Mora Theory within the constraint-based framework of OT poses some significant technical problems which have a direct bearing on my data. The present chapter will therefore be devoted to diagnosing and solving these problems in order to clear the way for the main argument of the dissertation, which is developed in chapters 3 and 4.

In section 2.2, I state my basic assumptions concerning the representation of segmental length and syllabic quantity. I adopt a Hayesian model of moraic structure, where onset consonants are directly adjoined to the syllable node, and weightless codas share the mora of a preceding rhymal segment (Hayes 1989). The view that weightless codas are linked to a mora headed by some other segment is supported with evidence from phonetics and from the behaviour of geminate consonants. I also discuss the input representation of short and long segments in the light of the principle of Lexicon Optimization, as revised by Inkelas (1994). Lexicon Optimization provides theory-internal support for Hayes's assumption that short vowels are underlyingly monomoraic, rather than non-moraic (cf. McCarthy & Prince 1986, 1988). Crucially, I also show how, in alternating environments, the representation of short consonants as non-moraic (i.e. prosodically underspecified) is compatible with Lexicon Optimization.

Interestingly, a hitherto unaddressed problem arises when a non-moraic input consonant is syllabified in the rhyme. In such cases, the output representation contains a segment-mora association which is absent from the input, and should accordingly be assessed as unfaithful by constraints such as DEP^μ or IDENT^μ. Yet there are conceptual and empirical grounds to suggest that DEP^μ and IDENT^μ cannot penalize such segment-mora links, which do not result in the neutralization of length contrasts. This difficulty is addressed in section 2.3. I define a relationship of positional μ -licensing, which obtains between a mora μ and a segment α when α is non-moraic in the input representation and μ is the sole prosodic licenser of α in the output. The concept of positional μ -licensing enables me to rectify the original formulation of DEP^μ and IDENT^μ, which are restated as conjunctive macro-constraints: e.g. in its conjunctive reformulation, DEP^μ is violated only by a mora which lacks an input correspondent *and* does not act as positional licenser of some segment.

2.2 THE MORaic REPRESENTATION OF LENGTH AND QUANTITY

2.2.1 Weightless codas

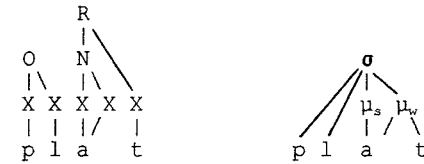
I assume that segmental length and syllabic quantity are represented in terms of morae, rather than X-positions.¹ Mora Theory exists in several variants; here I adopt a Hayesian model in which the contrast between moraic and non-moraic segments mirrors the distinction between rhyme and onset segments in X-position Theory (Hayes 1989). Implicit in this model is an inviolable principle of Rhyme Exhaustivity,² which rules that, within syllable boundaries, post-nuclear segments are exhaustively parsed into morae (cf. Al-

¹ For X-position Theory, see e.g. Levin (1985) and Lowenstamm & Kaye (1986). For Mora Theory in its various flavours, see e.g. Hayes (1989), Hyman (1985) and McCarthy & Prince (1986, 1988). Following Hayes (1989), it has generally been accepted that Mora Theory provides a more restrictive model of compensatory lengthening than X-position Theory (see e.g. Bickmore 1995, Broselow 1995), although Rialland (1991) and Noske (1992) challenge this view. Mora theorists, however, have experienced difficulties dealing with so-called 'failures of moraic consistency' (Broselow 1995: 199). Moraic consistency is said to fail when the same segment sequence appears to have different weights for different grammatical purposes (e.g. stress assignment and compensatory lengthening). See Hayes (1995: §7.3) for responses to this problem.

² Following a widespread and useful notational convention, I give the names of violable optimality constraints in SMALLCAPITALS. Other grammatical statements, such as non-violable principles, are labelled in Normal Type: e.g. Rhyme Exhaustivity, Weak Layering, Proper Headedness.

Mohanna 1997). In this light, consider the representation of a hypothetical form *plaxt* in a language with a binary weight distinction:

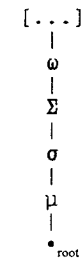
(2,1) (a) *X-position Theory* (b) *Mora Theory*



In (2,1b), the onset sequence *pl* is characterized by the fact that both segments are directly linked to the syllable node without the mediation of the mora tier. The rhyme *a:t*, in contrast, comprises all segments linked to a mora. If we assume, with Prince (1990: 377), that rhyme structure is 'micro-trochaic', then the nuclear segment, i.e. *a*, may be identified simply by reference to its leftmost position within the rhyme. Following this assumption of micro-trochaicity, I shall henceforth refer to the leftmost mora of any syllable as its 'head mora' or 'strong mora' (see Zec 1995b: 91 and references therein); thus, in (2,1b), μ_s is the head mora. Similarly, a mora is said to be headed by the leftmost segment it dominates: e.g. in (2,1b), μ_w is headed by *a*.

The representation of onset consonants as non-moraic relies on the assumption that, within the limits of Proper Headedness, the prosodic hierarchy is weakly, rather than strictly, layered (Itô & Mester 1992; cf. Selkirk 1984b):

(2,2) **The prosodic hierarchy**³



³ I treat root-nodes, rather than morae, as the terminal units of the prosodic hierarchy; cf. Itô & Mester (1992: note 7). Itô & Mester's arguments refer to ambisyllabicity, on which see footnote 10 below.

(2,3) **Weak Layering**

A prosodic unit of level n may immediately dominate prosodic units of levels lower than $n-1$.

(2,4) **Proper Headedness**

Every prosodic unit of level n must be headed by a prosodic unit of level $n-1$.

Under Rhyme Exhaustivity, then, syllable-internal weak layering is enlisted to express the distinction between onset and rhyme: rhyme segments are moraic; onsets are non-moraic. In line with Proper Headedness, the minimal syllable consists of a single moraic segment. Outside the syllable, weak layering arises under certain constraint rankings as a comparatively marked configuration. The constraint WEAKC, for example, requires word-final consonants to be dominated by the fewest possible prosodic nodes; therefore, if WEAKC is superordinate, a word-final consonant becomes extrasyllabic and is licensed directly by the ω node (Spaelti 1994).⁴ OT postulates a family of violable PARSE constraints which favour strict layering of all prosodic units other than onset segments (Spaelti 1994: 578-9; cf. Lin 1997: 416 for different labels):

(2,5) **The PARSE family⁵**(a) **PARSE^{seg}**

Every segment must be dominated either by a mora node or by a syllable node.

⁴ In recent OT work, edge extraprosodicity phenomena are generally viewed not as 'invisibility', but as 'weak edge' effects (Spaelti 1994). According to this conception, the right edge of the prosodic word is inherently weak, in the sense that designated rightmost constituents are preferably dominated by the fewest possible nodes.

⁵ In the original version of OT developed by Prince & Smolensky (1993), PARSE constraints functioned simultaneously as faithfulness constraints and as guardians of strict layering. In McCarthy & Prince's (1995) more recent version, faithfulness enforcement has been taken over by the correspondence family (MAX, DEP, IDENT, etc.). Accordingly, PARSE can now specialize in penalizing weak layering.

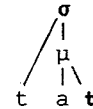
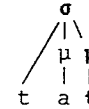
(b) **PARSE^h**

Every mora must be dominated by a syllable node.⁶

(c) **PARSE^σ**

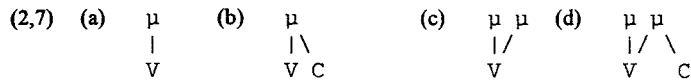
Every syllable must be dominated by a foot node.

As is well known, there are languages in which coda consonants do not count for weight: i.e. whilst CVV syllables are heavy, CVC syllables are light (e.g. Lardil, Malayalam, Selkup). Under the assumption of Rhyme Exhaustivity, the contrast between weight-contributing codas and weightless codas must be expressed in terms of mora sharing: a weight-contributing consonant heads its own mora, whilst a weightless consonant is adjoined to a mora headed by some other segment.

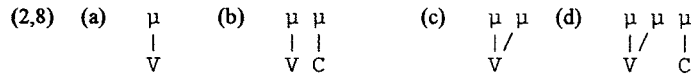
(2,6) (a) *weight-contributing coda* (b) *weightless coda*

In other versions of Mora Theory, coda segments may be directly adjoined to the syllable node, behaving as the mirror image of onset consonants (e.g. McCarthy & Prince 1986; Sherer 1994; Zec 1995a, b); in such models, weight-contributing codas head their own morae, as in (2,6a), but weightless codas are immediately dominated by σ . As shown by Broselow, Chen & Huffman (1997), there are two arguments which speak in favour of Rhyme Exhaustivity, with its corollary that weightless codas are represented as moraic non-heads. First, this principle throws light on the phonetic facts of segment duration, for mora sharing has been found to have a consistent subphonemic (i.e. non-contrastive) shortening effect. Broselow, Chen & Huffman's experimental results indicate that, in any given language, the duration of a short vowel will be phonetically larger in structure (2,7a) than in structure (2,7b), where the vowel shares its mora with a following consonant; similarly, the long vowel in (2,7d) will be phonetically shorter than that in (2,7c).

⁶ If PARSE^{seg} is ranked relatively high and PARSE^h is ranked relatively low, a situation can arise where segments are licensed by extrasyllabic morae (contrary to the hypothesis of Mora Confinement; Itô & Mester 1992: 11). Evidence for precisely such a state of affairs has been found in Gokana (Hyman 1985), Bulgarian (Zec 1988), Bela Coola (Bagemihl 1991), Piro (Lin 1997), and Arabic (Kiparsky 1998a). Extrasyllabic morae are also involved in the representation of word-initial geminate consonants; see §2.2.2.3.



Conversely, a short vowel will have roughly the same phonetic duration, *ceteris paribus*, in structure (2,8a) as in structure (2,8b), where the following consonant heads its own mora. The same applies to the long vowel in (2,8c) and (2,8d).



Broselow, Chen & Huffman's findings about the phonetic effects of mora sharing confirm the analysis of full lengthening and half lengthening in Bantu propounded by Hubbard (1995a,b), Maddieson (1993) and Maddieson & Ladefoged (1993); see (3,69). A second argument in favour of Rhyme Exhaustivity is provided by the behaviour of geminate consonants. Mora sharing is required to account for the fact that geminate consonants, though moraic, can be non-weight-contributing (Broselow 1995: 193; Broselow, Chen & Huffman 1997: §5.2; Sprouse 1996); see §2.2.2.3. In this light, a unified representation of coda weightlessness in terms of mora sharing appears more parsimonious.

Whether or not coda consonants are weight-contributing depends on the ranking of a family of sonority-sensitive constraints on mora headship (for the format of the constraints, see Broselow, Chen & Huffman 1997:65; for their sensitivity to sonority, see Zec 1988, 1995b):

(2,9) **Constraints on mora headship**⁷

(a) *** μ /OBS**

A mora must not be headed by an obstruent.

(b) *** μ /SON**

A mora must not be headed by a sonorant consonant.

⁷ It may well be the case that finer grained sonority distinctions will be required in some languages; see Zec (1995b: note 5).

These constraints interact with *BRANCH^μ, which requires that a mora should not dominate more than one root-node (Broselow, Chen & Huffman 1997: 65; Sprouse 1996: 398, 406; Walker 1994: 103), yielding the following results:

(2,10) If *BRANCH^μ » * μ /OBS, then (2,6a) > (2,6b)

If *BRANCH^μ « * μ /OBS, then (2,6a) < (2,6b)

Constraints (2,9a) and (2,9b) are arranged in a fixed harmonic ordering * μ /OBS » * μ /SON (see Sherer 1994: §2.6.2.2). Thus, if obstruents are allowed to head morae (*BRANCH^μ » * μ /OBS), so are sonorant consonants (*BRANCH^μ » * μ /SON, by transitivity), but not necessarily vice versa.⁸

Under the principle of Rhyme Exhaustivity, the interaction of the constraints PARSE^{Seg}, *BRANCH^μ and * μ /OBS yields a full typology of the behaviour of short obstruents in rhymes (cf. Sherer 1994: ch. 2). For the sake of simplicity, I assume that, in all three cases, superordinate PARSE^μ blocks licensing by an extrasyllabic mora.

(2,11)

Ranking	Effect	Example
{PARSE ^{Seg} , *BRANCH ^μ } » * μ /OBS	Rhymal obstruents contribute to weight	English
{PARSE ^{Seg} , * μ /OBS} » *BRANCH ^μ	Rhymal obstruents are weightless	Lithuanian
{*BRANCH ^μ , * μ /OBS} » PARSE ^{Seg}	Obstruents not licensed in rhyme	Beijing

For Lithuanian, where CVO syllables are light, whilst CVS syllables are heavy, see Zec (1995b: §2.3.1); for Beijing Chinese, where the only coda consonants are /n ŋ ʅ/, see Broselow (1995: 227-228).

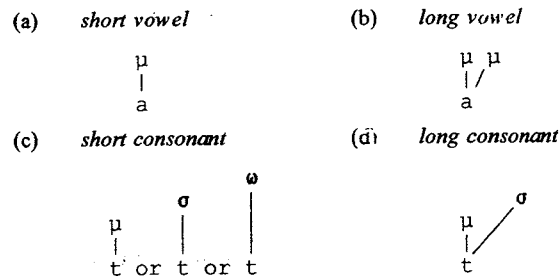
⁸ In those versions of Mora Theory where weightless codas are immediately dominated by σ (e.g. McCarthy & Prince 1986; Sherer 1994; Zec 1995a, b), the function of *BRANCH^μ is performed by a constraint which prohibits non-moraic codas (labelled WEIGHTBYPOSITION in Morén 1998 and Struijke 1998, *APPENDIX in Sherer 1994). This constraint interacts with a family of other constraints which require that segments attached to morae should be highly sonorous.

2.2.2 Segmental length contrasts and Lexicon Optimization

2.2.2.1 Richness of the Base and Lexicon Optimization

Let us now turn to the representation of segmental length. At the output level, matters are relatively straightforward (subject to a minor caveat discussed in footnote 10): any segment linked to a single prosodic node is interpreted as categorically short; any segment linked to two prosodic nodes is interpreted as categorically long. Typical representations are given in (2,12):

(2,12) Segmental length contrasts at the output level



The real difficulty lies in establishing how length contrasts are specified in input representations. In OT, this problem must be considered in the light of Lexicon Optimization (Prince & Smolensky 1993: §9.3; Itô, Mester & Padgett 1995).

Under Lexicon Optimization, the content of input representations is determined by reference to the constraint hierarchy. Simplifying somewhat, the principle can be stated as follows:

(2,13) Lexicon Optimization (initial version)

Let there be a set of potential input representations $I = \{I_1, I_2, \dots, I_n\}$, all of which converge onto the output form O under the constraint hierarchy \mathcal{M} . Within the set I , there will be a representation I_i such that the mapping $I_i \rightarrow O$ best satisfies \mathcal{M} . The learner chooses I_i as the input representation for O .

Accordingly, OT establishes no a priori limits on the content of underlying forms, this property of the theory being known as Richness of the Base. Notably, no assumption of underspecification, whether radical or contrastive, is made. In practice, since faithfulness

constraints penalize input-output disparity, Lexicon Optimization and Richness of the Base conspire to render input representations as close as possible to surface forms.

In an important development, Inkelas (1994) has fine-tuned the principle so that it can select the optimal input representation for a morpheme with more than one surface alternant.

(2,14) Lexicon Optimization (revised by Inkelas 1994: 6-7)

Given a grammar \mathcal{G} and a set $S = \{S_1, S_2, \dots, S_i\}$ of surface phonetic forms for a morpheme M , suppose that there is a set of inputs $I = \{I_1, I_2, \dots, I_j\}$, each of whose members has a set of surface realizations equivalent to S . There is some $I_i \in I$ such that the mapping between I_i and the members of S is the most harmonic with respect to \mathcal{G} , i.e. incurs the fewest marks for the highest ranked constraints. The learner should choose I_i as the underlying representation for M .

Under the revised version of Lexicon Optimization, a major restriction upon underspecification becomes apparent: only predictable alternating structure may be left lexically unspecified (Inkelas 1994: 7). Trivially, all unpredictable or idiosyncratic information must be fully specified in the input. More interestingly, Lexicon Optimization also favours the full specification of predictable non-alternating structure, since this reduces the disparity between the input and output representations, leading to the best satisfaction of faithfulness constraints. Thus, underspecification becomes possible only when a morpheme has different (predictable) surface realizations, for, in such cases, it is inevitable that at least one of the alternants will depart from the input representation.

(2,15) Specification of phonological structure in input representations

	<i>predictable</i>	<i>unpredictable</i>
<i>alternating</i>	unspecified <i>or</i> fully specified	fully specified
<i>non-alternating</i>	fully specified	fully specified

In a number of respects, addressing the input specification of segmental length in terms of (2,15) leads to proposals similar to those of Hayes (1989), particularly as regards vowel

(2,19) *Leti* /r:a:/ → [r:a:] 'again' in absolute initial position⁹



Note, incidentally, that the output representation diagrammed in (2,12d) does not entail that geminate consonants will always be weight-contributing. Research has shown that geminates may be either weight-contributing or weightless; *pace* Tranel (1991), the choice is open even to languages where short coda consonants are consistently weightless (cf. the Hindi dialect described by Gupta 1987; see Davis 1994). In our framework, a geminate will be weight-contributing if, and only if, it is allowed to head its own mora in the output representation (Broselow 1995: 193; Broselow, Chen & Huffman 1997: §5.2; Sprouse 1996). Abstract examples of both possibilities are given in the following tableau:

(2,20)

	μ μ μ a t a	IDENT ^u	MAX ^u	*BRANCH ^u	* μ /OBS
(a)	σ σ \ / μ μ μ a t a				*
(b)	σ σ μ μ μ \ / a t a		*	*	

(2,21) IDENT^u (initial version)

Let α be a segment in the input.

Let β be a correspondent of α in the output.

If α is linked to n morae, then β is linked to n morae.

⁹ I assume that the syllabicity of /a:/ is not subject to alternation, hence the presence of the syllable node in the input representation.

In both candidates, output *t* is phonetically long, as it is linked to two prosodic nodes.¹⁰ In candidate (2,20a), the geminate is weight-contributing; it makes the preceding syllable heavy. In candidate (2,20b), in contrast, the geminate is weightless; since it shares the mora headed by *a*, it does not contribute to mora count. Both candidates fulfil IDENT^u,¹¹ the selection of the winner depends on the relative ranking of * μ /OBS relative to MAX^u and *BRANCH^u.

2.2.2.4 Short consonants

Let us finally turn to the input representation of short consonants. We saw under (2,12c) that, from a prosodic point of view, short consonants are the most versatile of segments; they may be immediately dominated by potentially any prosodic node. In non-alternating environments, this versatility poses no particular difficulty: under Lexicon Optimization, the consonant will already be adjoined to the appropriate node in the input representation. In alternating environments, however, a choice has to be made as to its underlying prosodic specification. In (2,22), for example, I have given an abstract illustration of the range of prosodic alternations in which short consonants routinely participate:

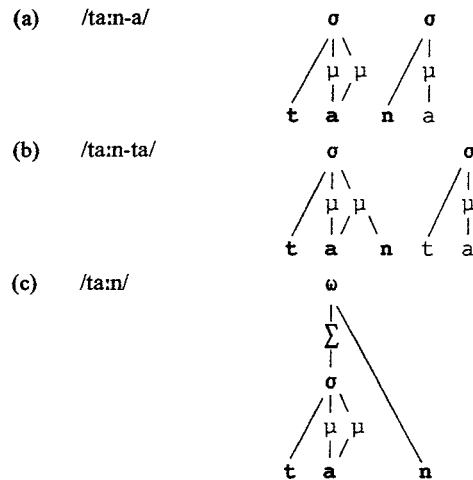
¹⁰ If one countenances ambisyllabicity (for arguments against, see J. Harris 1992, 1994), one faces the problem that ambisyllabic consonants are doubly linked, like geminates, yet phonetically short. There are two possible solutions to this problem. One involves assuming that ambisyllabics and geminates are phonologically identical; their realization as long or short would depend on a postulated parameter of phonetic interpretation for doubly-linked consonants (Borowsky, Itô & Mester 1983; Hammond 1997). This would entail that ambisyllabics and geminates cannot contrast within any one language, which appears to be correct (to the extent that confidence in any statement about ambisyllabicity is possible). Alternatively, ambisyllabics and geminates may be assigned different representations. Again, there are two possibilities. Itô & Mester (1992: 7) suggest that geminate consonants head their own mora (as in 2,20a), whilst ambisyllabics share the mora headed by a preceding segment (as in 2,20b). This proposal is untenable, because, as we have seen, the doubly-linked mora-sharing structure in (2,20b) is needed to represent non-weight-contributing geminates. A second possibility would be to represent ambisyllabics as simultaneously and immediately dominated by two syllable nodes, thus:



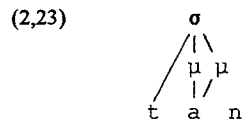
Given our assumptions about mora structure, this representation would imply that ambisyllabics are somehow extrasyllabic. Additionally, it would be necessary to claim that a doubly-linked segment is phonetically long if, and only if, at least one of its attachments is moraic.

¹¹ The statement of IDENT^u in (2,21) follows the general schema for the IDENT family established in McCarthy & Prince (1995: §2.2). The constraint is equivalent to Broselow, Chen & Huffman's (1997: 65) MORAFaITH. See §2.3 for further discussion and a revised version.

(2,22) Hypothetical surface alternants of a morpheme /ta:n/



How should the morpheme /ta:n/ be lexically specified in the light of such alternations? To begin with, the principle of Lexicon Optimization, with its corollary of minimum input-output disparity, decrees that non-alternating structure should be fully specified. In the case of /n/, however, some input-output disparity is inevitable as a result of the segment's alternating prosodification. Assume, however, that alternations such as (2,22) are preferably handled in structure-filling fashion;¹² if so, then /n/ will be prosodically unattached in the input. The preferred underlying representation of /ta:n/ will accordingly be the following:



¹² Inkelas supplements her reformulation of Lexicon Optimization with a further principle of Grammar Optimization:

Grammar Optimization (adapted from Inkelas 1994: 11)

The optimal grammar is the most transparent, i.e. the one in which alternations are maximally structure-filling (Kiparsky 1993). In terms of Optimality Theory, this means that MAX constraints are ranked as high as possible.

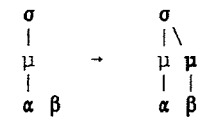
In (2,23), all non-alternating structure is fully specified; only the prosodic status of /n/ is underspecified; see table (2,15). In sum, short consonants in alternating environments will be lexically specified as non-moraic, as in McCarthy & Prince (1986, 1988) and Hayes (1989); cf. Hyman (1985). For a similar result, see Sprouse (1997: §5.6). In the next section we shall see, however, that the representation of short consonants as underlyingly non-moraic creates non-trivial problems for the assessment of input-output faithfulness.

2.3 ISSUES IN MORAIIC FAITHFULNESS

2.3.1 Positional μ -licensing

The previous section established that, in line with Rhyme Exhaustivity, coda consonants are licensed by the μ node. In Hayes's (1989) rule-based version of Mora Theory, in which all short consonants are underlyingly non-moraic, this requires the application of two rules ordered at the end of the syllabification algorithm (Hayes 1989: 258-9). First, weight-contributing codas are assigned a new mora by a parameterized rule of Weight by Position (2,24a); if Weight by Position is set to 'off', then CVC syllables will be light.¹³ Subsequently, an adjunction operation (2,24b) links any remaining weightless consonants to an already existing μ node:

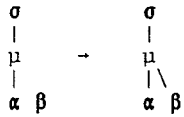
(2,24) (a) Weight by Position



Active?: yes/no

If active, iterative?: yes/no

¹³ If, in contrast, Weight by Position is not only active but also iterative, then the language may tolerate trimoraic syllables (Hayes 1989: §7).

(b) **Adjunction**

As the previous section indicated, the situation in OT is slightly different. In non-alternating environments, short coda consonants will already be linked to the appropriate μ node in the input representation. It is only in alternating environments, where short consonants can be prosodically underspecified, that licensing in the coda will require the creation of segment-mora associations absent from the lexical representation. This section will be devoted to the treatment of such non-underlying mora-segment associations in OT, focusing on the rôle of faithfulness constraints. Adapting Hayes's terminology for my present purposes, I shall henceforth use the expression 'positional μ -licensing' to denote the relationship between a short, underlyingly non-moraic coda consonant and its mother μ node. Thus, in both (2,24a) and (2,24b), we will say that μ is the 'positional licenser' of β .

Our discussion so far has indicated that coda consonants contribute to weight only if the ranking {PARSE^{Seg}, *BRANCH^M} » * μ /C obtains; conversely, they will be weightless under the hierarchy {PARSE^{Seg}, * μ /C} » *BRANCH^M; see (2,11). In alternating environments, however, the phenomenon of positional μ -licensing adds one further layer of complexity to the issue. Since the segment-mora associations involved in positional μ -licensing are absent from the input representation, the null assumption under Correspondence Theory (McCarthy & Prince 1995) is that they will be penalized by constraints on moraic faithfulness, specifically by IDENT^M (see (2,21) above) and by DEP^M:

(2,25) **DEP^M (initial version)**

Every mora in the output has a correspondent in the input.

Both constraints play a vital rôle in controlling the neutralization of segmental length contrasts. By inhibiting changes in the moraic specifications of segments, IDENT^M discourages shortening and lengthening. DEP^M is slightly more specialized; it prohibits mora

insertion, and therefore prevents all forms of lengthening other than compensatory lengthening.¹⁴

Nonetheless, if, in accordance with the null assumption, we maintain that both DEP^M and IDENT^M penalize positional μ -licensing, we are led to the intuitively repugnant conclusion that syllabifying an underlyingly non-moraic consonant in the rhyme amounts to an instance of lengthening. Using rule-based terminology, the implication would be that the application of Weight by Position or Adjunction deserves the same penalty as the application of a vowel lengthening rule. I suggest that this implication is wrong, and that DEP^M and IDENT^M must be reformulated so as not to inhibit positional μ -licensing.¹⁵

The source of this problem lies in the fact that morae do not behave like segmental features for the purposes of neutralization. Consider a language with a symmetrical inventory of oral and nasalized vowels. In such a language, if a nasal autosegment spreads to an underlyingly oral vowel, neutralization automatically ensues: the vowel becomes a member of the nasalized set. The prosodic status of the vowel is, in this respect, irrelevant. With morae, however, the situation is different: association of a mora μ with a segment α results in lengthening of α if, and only if, α is also associated with some prosodic node other than μ . Hence, not all cases where GEN creates a new mora-segment association have the same import in relation to length.

It is easy to see that, in the rule-based version of Mora Theory, both the formulation and the ordering of Weight by Position and Adjunction are contrived to be non-neutralizing.¹⁶ First, both rules apply only to unsyllabified segments, i.e. to segments with no previous prosodic attachment; see (2,24). In addition, Hayes stipulates that, within syllabification algorithms, Weight by Position and Adjunction are universally ordered *after*

¹⁴ Compensatory lengthening transfers an already existing mora between segments. This is reflected in Hayes's (1989: 285) principle of Moraic Conservation:

Moraic Conservation

Compensatory lengthening processes conserve mora count.

See section 3.5.1 for a discussion of Mora Conservation in OT.

¹⁵ So counterintuitive is the equation of Weight by Position with lengthening, that some scholars have ignored it, tacitly reformulating DEP^M. Thus, Green (1997: §1.2) states that, "If a candidate lengthens an input short vowel or consonant, it violates DEP(μ). This is incorrect in two respects. First, it ignores the possibility of compensatory lengthening (see previous footnote). Secondly, it overlooks the fact that DEP^M, as defined in (2,25), penalizes positional μ -licensing, despite the fact that the relevant consonants remain short.

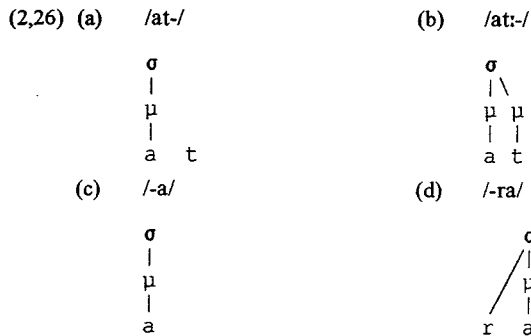
¹⁶ As Kiparsky (1993: 278) observes, "Notoriously, the class of neutralization rules cannot be formally specified, for whether a rule is neutralizing depends on the derivations of the grammar."

the Onset Formation rules (see Hayes 1989: 259). As a result, an underlyingly non-moraic coda will never emerge with a double linkage from the initial round of syllabification. Gemination can only be effected by some later process of resyllabification: e.g. (3,56).

Of course, conceptual arguments grounded on intuition are not sufficient to reject the initial versions of DEP^μ and IDENT^μ. The next two sections provide fully worked out empirical arguments. I will show that, if one allows these two faithfulness constraints to penalize positional μ-licensing, then one can construct grammars which describe unattested languages. In other words, the initial versions of DEP^μ and IDENT^μ lead to overgeneration by UG. In each section, I will design a toy grammar where geminates syllabify in an entirely different way from short consonants, leading to unnatural alternations. Both examples will exploit the fact that geminate consonants are underlyingly moraic, and can therefore surface in a moraic position without violating the initial versions of DEP^μ and IDENT^μ. Underlyingly non-moraic consonants, in contrast, will be barred from the rhyme owing to the inhibiting effect of superordinate DEP^μ and IDENT^μ upon positional μ-licensing.

2.3.2 First anomalous prediction

Let L_a be a language whose lexicon comprises two lexemes, /at-/ and /at:-, which are inflected by means of two suffixes, /-a/ and /-ra/. Following the conclusions of section 2.2.2, assume the following input representations:¹⁷



Furthermore, let the following rankings obtain in the grammar of L_a :

¹⁷ As the reader will be able to ascertain in due course, the lexical representations given in (2,26) comply with the requirements of Lexicon Optimization relative to the constraint hierarchy of L_a .

- (2,27) (a) { DEP^μ, *BRANCH^μ, PARSE^{Seg}, MAX^{Seg} } » *_σCC » CONTACT
 (b) { MAX^μ, *BRANCH^μ } » *μ/OBS

(2,28) *_σCC (see Prince & Smolensky 1993: 87)

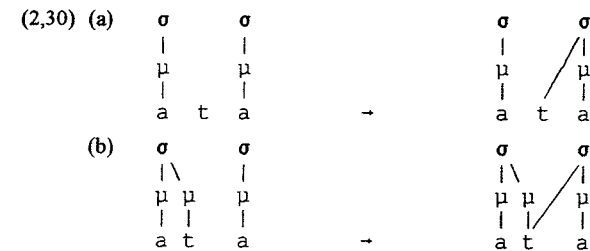
The onset comprises no more than one segment.

(2,29) CONTACT (see Clements 1990, 1992; Vennemann 1988: 40)

$\alpha_{\sigma}][_{\sigma}\beta \rightarrow (\text{Son}(\alpha) > \text{Son}(\beta))$

Given a syllable contact $\alpha_{\sigma}][_{\sigma}\beta$, the sonority value of α is greater than that of β .

Under this constraint hierarchy, the behaviour of the vocalic suffix /-a/ is relatively uninteresting: see (2,30). To confirm that the output representation in (2,30b) is optimal under the constraint ranking established in (2,27b), refer to tableau (2,20) above.



Forms containing the suffix /-ra/, however, behave in a rather peculiar way.

Let us first consider the results of adding /-ra/ to the lexeme /at:-/ in tableau (2,31). Under the ranking *_σCC » CONTACT established in (2,27a), heterosyllabic consonant clusters $\alpha_{\sigma}][_{\sigma}\beta$ are more harmonic than branching onsets $[_{\sigma}\alpha\beta$, irrespective of whether sonority rises or falls from α to β .¹⁸ Accordingly, candidates (a) to (c) fail. In the remaining two candidates, the input cluster /t:r/ becomes heterosyllabic. Concomitantly, the final consonant of the base, which is underlyingly moraic, degeminates, since it is not allowed

¹⁸ Such is the case in, for example, Turkish. In [se.vap.tan] 'good deed' abl., [p] has undergone coda devoicing and triggered closed syllable shortening of the preceding vowel; cf. the possessive form [se.va: bi]. Crucially, the syllabification of the nom. pl. form [se.vap.lar] is parallel to that of [se.vap.ten], even though the cluster [p.l] has a rising sonority contour, unlike [p.t] (data from Kaye 1990: 302). The ranking *_σCC » CONTACT also obtained in the Germanic protolanguage (see §3.5.2.2); this created rising-sonority coda-onset clusters which underwent developments of great theoretical interest in West Germanic (see §3.5).

to link up to the following syllable node (see §2.2.2.3). Finally, MAX^μ and *BRANCH^μ prefer candidate (e) to candidate (d). Observe that candidate (e) does not violate either DEP^μ or IDENT^μ, as the input correspondent of [t] is linked to its own mora.

(2,31)

	μ μ μ a t r a	MAX ^μ	*BRANCH ^μ	*[_σ CC	*μ/OBS	CONTACT
(a)		*!	*!	*!		
(b)		*!		*!		
(c)				*!	*	
(d)		*!	*!			*
(e)					*	*

Let us now turn to tableau (2,32) to find out the fate of input /at-ra/ under the assumption that positional μ-licensing violates DEP^μ.

(2,32)

	μ μ a t r a	DEP (μ)	*BRANCH (μ)	PARSE (Seg)	MAX (Seg)	*[_σ CC	CONTACT
(a)		*!					*
(b)			*!				*
(c)				*!			
(d)		*!					
(e)					*!		
(f)						*	

In candidate (a), the input cluster /tr/ is heterosyllabified, in compliance with the ranking *_σCC » CONTACT. In the coda, [t] projects a mora, which acts as its positional licensor, thereby fulfilling superordinate *BRANCH^μ—in contrast with candidate (b), which therefore fails. However, the positional licensor μ_p has no input correspondent, and so DEP^μ rules out

candidate (a).¹⁹ It follows that, under the constraint hierarchy established in (2,27a), the parse [_o[_oat][_ora]] is not viable, despite the high ranking of *_oCC.

Candidates (c) to (e) represent further attempts to avoid syllabifying /tr/ as a complex onset. In (c) and (d), /t/ is extrasyllabic. Candidate (c) fails because superordinate PARSE^{Seg} requires every segment to be dominated by either μ or σ , see (2,5a). Candidate (d), where /t/ is licensed by an extrasyllabic mora, is again ruled out by DEP^u, as the extrasyllabic mora has no input correspondent; it also violates PARSE^u (2,5b). Finally, MAX^{Seg} blocks segment deletion, and so disqualifies candidate (e). Despite the ranking *_oCC » CONTACT, therefore, the optimal parse is [_o[_oa][_otra]], with a branching onset. Interestingly, candidate (a) would have emerged victorious had we assumed that positional licensers do not violate DEP^u or IDENT^u.

In conclusion, L_2 exhibits a highly unnatural set of alternations:

(2,33)

	/at-/	/at:-/
/-a/	[a.ta]	[at.ta]
/-ra/	[a.tra]	[at.ra]

In particular, I know of no language where clusters of short obstruent plus short sonorant are tautosyllabic if the obstruent is single in intervocalic position, but heterosyllabic if the obstruent is geminate in intervocalic position. However, if we assume that morae involved in positional μ -licensing do not violate DEP^u, a much more natural set of alternations results:

(2,34)

	/at-/	/at:-/
/-a/	[a.ta]	[at.ta]
/-ra/	[at.ra]	[at.ra]

¹⁹ Note, additionally, that both candidate (a) and candidate (b) fail IDENT^u, as input /t/ has no moraic links.

The forms in (2,34) correspond to a system with no branching onsets and with consonant degemination in preconsonantal position.

2.3.3 Second anomalous prediction

Let L_3 be a language with two lexemes, /an-/ and /an:-/, inflected by means of two suffixes, /-a/ and /-ta/. Assume input representations analogous to those in (2,26). Moreover, let the following rankings obtain in the grammar of L_3 :

- (2,35) (a) { ONSETFORM, IDENT^u, PARSE^{Seg} } » MAX^{Seg}
 (b) { MAX^u, *BRANCH^u } » * μ /SON

I use ONSETFORM as a cover term for two constraints. First, onset clusters must have rising sonority contours, in accordance with the Sonority Sequencing Generalization (Selkirk 1984a: 116). Secondly, onset-initial segments are minimally sonorous (Clements 1990, 1992; Vennemann 1988: 13-14).

Given (2,35), the inflected forms /an-a/ and /an:-a/ behave in exactly the same way as the L_2 forms shown in (2,30):

- (2,36) (a) σ σ σ σ
 | | | / |
 μ μ μ μ
 | | | | |
 a n a a n a
- (b) σ σ σ σ
 | \ | | \ |
 μ μ μ μ μ μ
 | | | | | |
 a n a a n a

But, yet again, the -CV suffix induces bizarre prosodic behaviour in the base.

Let us first look at tableau (2,37) to find out what happens when we add the suffix /-ta/ to the lexeme /an:-/. The results are straightforward. First, input /n/ undergoes degemination in preconsonantal position: ONSETFORM rules out candidate (a), which contains an onset cluster with a falling sonority contour, whilst MAX^{Seg} weeds out candidate (c), where onset well-formedness is achieved through segment deletion. Of the remaining two candidates, MAX^u and *BRANCH^u choose (d), where [n] is syllabified as a weight-

contributing coda. Observe that the optimal parse trivially satisfies IDENT^μ, as the input correspondent of coda [n] is monomoraic.

(2,37)

$\begin{array}{c} \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad n \quad t \quad a \end{array}$	ONSETFORM	IDENT ^μ	MAX ^μ	*BRANCH ^μ	MAX ^{Seg}	*μ/SON
(a) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad n \quad t \quad a \end{array}$	*!					*
(b) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad n \quad t \quad a \end{array}$			*!	*!		
(c) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad n \quad a \end{array}$					*	*
(d) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad n \quad t \quad a \end{array}$						*

In contrast, tableau (2,38) shows that, if we continue to adhere to the assumption that positional μ -licensing infringes moraic faithfulness, input /an-ta/ behaves unexpectedly. Since /n/ is underlyingly non-moraic, superordinate IDENT^μ debars it from being linked up to a mora, as a result, /n/ is excluded from the rhyme by Rhyme Exhaustivity (see section 2.2.1). However, an extrasyllabic position is not an option either, for PARSE^{Seg} is also superordinate. Eventually, since [nt] is an ill-formed onset, and ONSETFORM dominates MAX^{Seg}, one of the two consonants must be deleted. Because ONSETFORM encodes a preference for onsets with low sonority, it is /n/ that has to go. Note, however, that, if we assumed IDENT^μ to be blind to positional μ -licensing, then the parse [_a[_σan][_σta]], represented by candidate (a), would be adjudged optimal.

(2,38)

$\begin{array}{c} \mu \quad \mu \\ \quad \\ a \quad n \quad t \quad a \end{array}$	ONSETFORM	IDENT ^μ	PARSE ^{Seg}	*BRANCH ^μ	MAX ^{Seg}	*μ/SON
(a) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu_p \quad \mu \\ \quad \quad \\ a \quad n \quad t \quad a \end{array}$		*!				*
(b) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad n \quad t \quad a \end{array}$		*!		*!		
(c) $\begin{array}{c} \sigma \\ \quad \backslash \quad / \quad \\ \sigma \quad \mu \quad \sigma \\ \quad \quad \\ a \quad n \quad t \quad a \end{array}$			*!			
(d) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad n \quad t \quad a \end{array}$	*!					
(e) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad n \quad a \end{array}$	*!					*
			[_σ n < [_σ t			
(f) $\begin{array}{c} \sigma \quad \sigma \\ \quad \backslash \quad / \quad \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ a \quad t \quad a \end{array}$						*

In sum, L_0 has the following set of alternations:

Here, μ_j is delinked from its underlying anchor a_j and transferred to a_i ; at the same time, a_j projects a new mora μ_k . Observe that the relationship between a_j and μ_k fulfils condition (2,40ii): a_j is uniquely dominated by μ_k . However, if DEP^μ is to be blind to positional μ -licensors, then μ_k must somehow be excluded from the definition; otherwise, the lengthening of a_j would be interpreted as compensatory (i.e. as not involving mora epenthesis), despite the fact that the segment from which μ_j is transferred still surfaces as moraic. In other words, the contrast between compensatory lengthening by vowel loss, e.g. /tata/ → [ta:t] (Hayes 1989: §4.2, §5.1.6), and simple open syllable lengthening, e.g. /tata/ → [ta:ta], would dissolve. To avoid this clearly undesirable result, the definition of positional μ -licensing must include clause (2,40i).

Equipped with (2,40), we can now revise the definitions of DEP^μ and IDENT^μ provided in (2,25) and (2,21), respectively. In the following statements, the Boolean operator \wedge indicates local constraint conjunction (see below).

(2,43) **DEP^μ (final version)**

Let μ be a mora in the output.

DEP^μ = (a) \wedge (b)

(a) μ has a correspondent in the input.

(b) μ is a positional μ -licenser.

(2,44) **IDENT^μ (final version)**

Let α be a segment in the input.

Let β be a correspondent of α in the output.

Let α be linked to n morae.

IDENT^μ = (a) \wedge (b)

(a) β is linked to n morae.

(b) β is positionally μ -licensed.

In (2,43) and (2,44), DEP^μ and IDENT^μ are defined as macro-constraints (Crowhurst & Hewitt 1997) resulting from the local conjunction (Smolensky 1993) of two micro-constraints (a) and (b). In both cases, (a) corresponds to the initial version of the homonymous constraint in Correspondence Theory, whilst (b) is a constraint requiring that a relationship of positional μ -licensing should obtain. According to Smolensky's definition of local

conjunction, a candidate c violates the macro-constraint (a) \wedge (b) if, and only if, c violates both micro-constraint (a) and micro-constraint (b).²¹ Note that both (2,43) and (2,44) amply satisfy Crowhurst & Hewitt's (1997) 'unique focus' requirement, since, in both cases, the two conjuncts (a) and (b) share the same primary argument:²² μ in the case of DEP^μ, β in the case of IDENT^μ.

Let us now consider in greater detail how the revised version of IDENT^μ works. Conjunct (2,44a) requires that a segment β should be linked to exactly the same number of morae as its input correspondent α . Accordingly, (2,44a) is satisfied whenever α and β are both non-moraic, or both monomoraic, or both bimoraic —and, given Smolensky's definition of local constraint conjunction, whenever (2,44a) is fulfilled, the macro-constraint (2,44) itself is also fulfilled. However, IDENT^μ may be satisfied under an additional set of circumstances captured by (2,44b): namely, when α is non-moraic, and β is prosodically licensed by means of a single moraic attachment. It follows that IDENT^μ, as defined in (2,44), only penalizes those changes to the moraic specification of a segment which lead to either shortening or lengthening; the simple syllabification of an underlyingly non-moraic consonant in the rhyme, as either a weightless or a weight-contributing coda, is not penalized.²³ DEP^μ, as reformulated in (2,43), works in a similar fashion.

2.3.4.2 Strict Surface Orientation and relativized faithfulness

Since (2,43) and (2,44) depart from the standard DEP and IDENT schemas established in McCarthy & Prince (1995), it now behoves us to examine their theoretical status within the OT framework. In particular, we must consider (2,43) and (2,44) in the light of Strict

²¹ Note that, in Smolensky's definition of local constraint conjunction, logical truth is equated with constraint violation. Thus, it is true that x violates (i) \wedge (ii) iff it is true that x violates (i) AND it is also true that x violates (ii). By the same token, it is true that x satisfies (i) \wedge (ii) iff it is true that x satisfies (i), OR it is true that x satisfies (ii), OR it is true that x satisfies both (i) and (ii). Crowhurst & Hewitt (1997), in contrast, equate logical truth with constraint satisfaction; hence, Smolensky's local conjunction is equivalent to Crowhurst & Hewitt's disjunction.

²² Crowhurst & Hewitt (1997: 10) define the primary argument of a constraint as 'the linguistic object upon which some condition of maximum harmony is predicated'; primary arguments are 'always interpreted with universal scope'.

²³ IDENT^μ behaves in a rather interesting fashion in relation to high vocoids, which may alternate between moraic and non-moraic realizations; see (2,16). It penalizes glide formation processes, where an underlyingly monomoraic vowel surfaces as a non-moraic onset glide. The reverse of glide formation is glide vocalization, where an underlyingly non-moraic glide surfaces as a monomoraic vowel. Glide vocalization does not violate IDENT^μ because the resulting vowel meets the criteria for positional μ -licensing.

Surface Orientation, a programmatic principle implicit in Prince & Smolensky's (1993) classical presentation of OT (see Itô & Mester 1997: 419 for a somewhat looser statement):

(2.45) **Strict Surface Orientation**

- (a) The universal constraint set CON comprises constraints of two kinds only: faithfulness constraints and markedness constraints.
- (b) Faithfulness constraints require that the output representation should be identical to the input representation in respect of some designated property.
- (c) Markedness constraints define well-formedness conditions on the output representation.

Strict Surface Orientation implies that CON may not include constraints on input forms (Richness of the Base; see §2.2.2.1), and that output well-formedness is not conditioned by input structure other than through faithfulness. The significance of these restrictions can be best understood by comparison with Cognitive Phonology (G. Lakoff 1993). In Cognitive Phonology, grammars contain so-called 'two-level constructions', i.e. statements establishing correlations between units across two separate levels of representation. Crucially, two-level constructions include statements such that, when an element α occurs in the environment $\gamma_ \delta$ at the Word level, then α must be in correspondence with element β at the Phonetic level:²⁴

(2.46)	W-level:	γ	α	δ
	P-level:		β	

²⁴ In common with Harmonic Phonology (Goldsmith 1993), Cognitive Phonology adopts the 'M/W/P model' of phonology (Goldsmith 1993: 33), which posits three levels of representation: M(orphemic), W(ord) and (P)honic. The M/W/P model has roots in classical structuralism. Interestingly, it has also been influential within the post-SPE tradition, as some versions of rule-based Lexical Phonology propose separate stem, word and postlexical modules (see e.g. Kiparsky 1985, Booij & Rubach 1987, Borowsky 1993). The issue of levels will be extensively discussed in section 3.3 in connection with the problem of opacity and its treatment in interleaved OT.

Crucially, (2.46) will be enforced whenever the sequence $\gamma\alpha\delta$ occurs at the W-level; there is no requirement that the conditioning environment $\gamma_ \delta$ should persist at the P-level. In this respect, a two-level construction such as (2.46) is closely analogous to a transformational rule of the format $\alpha \rightarrow \beta / \gamma_ \delta$, whose application cannot be reversed even if some later rule destroys the frame $\gamma_ \delta$. Thus, both transformational rules and two-level constructions have an inherent ability to trigger opacity effects of the 'overapplication' or 'non-surface-apparentness' variety (see §3.2.1).

The OT equivalent of a two-level construction such as (2.46) would be some kind of 'antifaithfulness constraint', requiring unfaithfulness of the output to the input whenever a certain marked structure occurs in the input. However, the principle of Strict Surface Orientation bans antifaithfulness constraints: according to (2.45b), faithfulness constraints must require input-output identity, not input-output disparity; (2.45c), in turn, restricts the scope of markedness constraints to output representations; and, finally, (2.45a) excludes any constraint type other than faithfulness or markedness constraints. Kirchner (1995) suggests that antifaithfulness devices vastly increase the descriptive power of grammars, leading to overgeneration on as large a scale as in SPE. Of course, the appeal of antifaithfulness constraints lies in the fact that, like SPE rules, they are able to model opaque phonological phenomena in a straightforward way (Kirchner 1996: note 2; see also §3.2.1 below).

It should be clear that DEP^u and IDENT^u, in their revised conjunctive formulations, are *not* antifaithfulness constraints. They differ from the original statements in that they fail to penalize mora-segment associations absent from the input if these arise through positional μ -licensing. Crucially, however, the new constraints do not themselves require that output moraic structure should depart from input specifications. Under Rhyme Exhaustivity, the need for positional μ -licensing arises independently at the behest of other constraints such as PARSE^{seg} (see §2.2.1, §2.3.2 and §2.3.3).

It may therefore be more useful to categorize (2.43) and (2.44) as 'relativized correspondence' constraints. As stated in (2.45b), each faithfulness constraint requires that the output should be identical to the input in respect of some designated property. Against this background, I use the term 'relativized correspondence' to refer to the insight that, in certain cases, the faithfulness requirement is not absolute, but is rather limited in some crucial fashion: e.g. a constraint may penalize disparity between correspondent representations only if it exceeds a certain threshold, or it may require identity only within a certain domain. In the current OT literature, the notion of relativized correspondence has received two instantiations:

First, relativized correspondence has been applied to the analysis of synchronic chain shifts (Kirchner 1995, 1996; Gnanadesikan 1997). Synchronic chain shifts can be intuitively described in terms of phonological scales: phonological units of a certain kind appear to move in unison along the relevant scale. Kirchner (1996: 344) provides an example from Nzebi, where the suffix /-i/ causes stem vowels to move one step up along the height dimension (I ignore the morphological restrictions of the process, and focus on the front vowels only):

- (2,47) (a) Height scale: $i > e > \varepsilon > a$
 (b) Alternations:
- | | <i>unraised</i> | <i>raised</i> | |
|-----------|-----------------|---------------|------------|
| /a/ → [ɛ] | sal | seli | 'to work' |
| /ɛ/ → [e] | βɛ:d | βe:di | 'to give' |
| /e/ → [i] | bet | biti | 'to carry' |

Assume that raising is triggered by some markedness constraint M requiring that the stem-vowel should assimilate in height to the suffix vowel. To account for the fact that there is any raising at all, M must outrank the faithfulness constraint which demands absolute identity with input vowel height specifications. It is nonetheless obvious that M cannot be superordinate, for otherwise we would incorrectly predict /a/ → *[i] and /ɛ/ → *[i]. This suggests that there is some higher-ranked faithfulness constraint, a 'distantial faithfulness' constraint, which tolerates height unfaithfulness, provided that the distance between a vowel and its input correspondent does not exceed one step along the height dimension (2,47a). Interestingly, Kirchner (1995) suggests that such 'distantial faithfulness' constraints can be reduced to elementary absolute faithfulness constraints operating under local conjunction (see §2.3.4.1); Gnanadesikan (1997), in contrast, assumes ternary scales as theoretical primitives. It is apparent, in any case, that, in accordance with Pāṇini's Theorem on Constraint Ranking (Prince & Smolensky 1993: §5.3, §A.3), relativized faithfulness effects occur under the following ranking schema:

(2,48) RELATIVIZEDFAITHFULNESS » MARKEDNESS » ABSOLUTEFAITHFULNESS

The concept of relativized faithfulness also subsumes so-called 'positional faithfulness' (Beckman 1998). Like all other correspondence constraints, positional

faithfulness constraints require identity between correspondent representations in respect of some designated property; they are special, however, in that their application is bound to a certain domain. It has been suggested that the relevant domain may be phonologically or morphologically defined (though see footnote 25). Alderete (1995), for example, proposes a family of head-faithfulness constraints whose scope is restricted to prosodic heads (stressed syllables or feet); see Beckman (1998: ch. 3). Similarly, McCarthy & Prince (1995: §6.2) propose specialized root-faithfulness constraints whose domain of application excludes affixes; see Beckman (1998: ch. 4).²⁵ Interestingly, both types of domain-bound faithfulness can be shown to affect the distribution of complex onsets. In southeastern Brazilian Portuguese dialects, for example, only syllables bearing primary stress can have branching onsets: e.g. *livréto* 'libretto' ~ *liv*[f]o 'book', *práto* 'plate' ~ *p*[f]at*el*éira 'plate rack' (J. Harris 1992: 399-400; data supplied by Thais da Silva). By Pāṇini's Theorem, this pattern implies the following ranking: HEADσ-MAX^{Seg} » *_{CC} » MAX^{Seg}. In Sanskrit, in contrast, branching onsets are circumscribed to roots, and fail to occur in affixes, implying the ranking ROOT-MAX^{Seg} » *_{CC} » MAX^{Seg} (McCarthy & Prince 1995: §6.2).²⁶

In conclusion, the revised formulations of DEP^u (2,43) and IDENT^u (2,44) fall under the category of relativized faithfulness, in that both constraints fail to impose absolute identity between the moraic structure of the input and that of the output; rather, positional μ-licensing is exempt from the identity requirement. Interestingly, I have shown that the appropriate relativization effect is achieved through local conjunction, as suggested by Kirchner (1995, 1996) in respect of distantial faithfulness.

²⁵ Positional faithfulness constraints with morphologically defined domains are probably redundant in an optimality-theoretic model incorporating interleaving. Within such a model, for example, root faithfulness effects would probably be regarded as involving non-derived environment blocking at the stem level. See chapter 3.

²⁶ Licensing asymmetries such as that between the onset and the coda may in principle be described in terms of either positional faithfulness or positional markedness; see e.g. Beckman (1998: ch. 1), Zoll (1998). In a positional faithfulness analysis, it is assumed that onset-specific faithfulness constraints protect onset consonants from the effects of general (non-positional) markedness constraints. In a positional markedness analysis, in contrast, it is assumed that coda consonants are targeted by coda-specific markedness constraints, which override general (non-positional) faithfulness constraints. In relatively complex instances, both positional faithfulness and positional markedness may be required; see the analysis of coda nasals in Spanish in §3.3.2, and also Féry (1998).

2.4 CONCLUSION

This chapter has outlined the version of Mora Theory which underpins the case studies in segmental length change conducted in chapters 3 and 4. I have shown that the matter of adapting Mora Theory to an optimality-theoretic framework is by no means trivial. The principle of Lexicon Optimization, for example, has a direct bearing on how length contrasts are represented in the input. More seriously, the original versions of DEP[#] and IDENT[#] (modelled on the constraint schemas provided by McCarthy & Prince 1995) have been found to penalize the association of any underlyingly non-moraic consonant with a mora, even though length contrasts are not neutralized. This is counterintuitive and entails undesirable typological predictions. I have proposed a solution to the problem by defining a relationship of positional μ -licensing obtaining between an underlyingly non-moraic segment α and a mora μ if μ is the sole prosodic licenser of α . DEP[#] and IDENT[#] have received conjunctive reformulations to take account of positional μ -licensing. The revised constraints have been shown to be compatible with the principle of Strict Surface Orientation, and to fall under the category of relativized faithfulness.

In chapter 3, the concept of positional μ -licensing will prove crucial to understanding a number of problematic opacity effects (§3.5). These arise as a mora μ , initially inserted to positionally license a segment α , is preserved in the output despite the subsequent resyllabification or deletion of α , which cancels the positional μ -licensing relationship. I shall compare two optimality-theoretic approaches to this problem: level segregation and sympathy. Interestingly, West Germanic Gemination (an opaque process not included in Hayes's 1989 list of mora-preserving lengthening effects) will prove instrumental in refuting a key assumption of Sympathy Theory as expounded by McCarthy (1998), and will thereby expose essential flaws in strongly parallel OT.

In chapter 4, the concept of positional μ -licensing will cast light on the system of prosodic preferences driving the ME vowel length adjustment. In particular, we shall see that, although WEAKC (i.e. the constraint requiring final consonant extrasyllabicity; see §2.2.1) was active in the language, the grammatical parse of input /CV[#]C/ was not [°[CV[#]]C], but rather [°[CV[#]C[#]]]. The bimoraic word minimum was achieved by syllabifying the final consonant, rather than lengthening the vowel, because superordinate DEP[#] prevented vowel lengthening, but not positional μ -licensing.

Chapter 3

Opacity

3.1 INTRODUCTION

It is one of the disconcerting facts of human language that phonological processes often misapply: an alternation is triggered outside its proper environment, or it fails where all its conditioning requirements are met. The analyst's skill is not to blame; we know of too many instances in which little room is left to doubt the accuracy of his descriptive generalizations. This counterintuitive quirk of phonological grammars is known as opacity. Why should it exist? This chapter suggests that opacity cannot be adequately understood either in purely synchronic or in purely diachronic terms. Availing myself of Lightfoot's (1999: 111) useful image, I shall claim that opacity is a 'Rube Goldberg feature of grammars': it is not specifically catered for by UG principles (although of course it complies with them), and it is not adaptive or functional; it arises through the way in which the contingencies of acquisition—or rather misacquisition—work upon the interface of phonology with the rest of the grammar.

From a diachronic viewpoint, Neogrammarian *Lautlehre* makes opacity seem the most natural thing in the world. Time and again, a younger sound change is seen to disfigure the output of an older one: apocope removes the trigger of umlaut; hard on the heels of closed syllable shortening, syncope creates new superheavy syllables. If sound changes succeed each other *mit blinder Notwendigkeit*, through the contingent phonologization of automatic phonetic effects, then there is no reason why this game of wrong-footing and one-upmanship should ever cease. In actual fact, phonological change is *not* altogether blind (Kiparsky 1995: §2.1; and see chapter 4 below); but work in experimental phonology has successfully replicated the conditions for hypocorrection in the laboratory. The term 'hypocorrection' designates those stochastic processes whereby the listener may misinterpret an automatic phonetic effect as instantiating a grammatically-controlled feature of the speaker's phonology (Ohala 1989, 1992, 1993).

Nonetheless, an account of opacity which only appeals to the blindness of sound change must obviously be incomplete. For opacity to exist as a specifically synchronic phenomenon, misapplying processes must be learnable. In other words, the language faculty

must enable children to acquire and mentally represent opaque grammars. The Neogrammarians never posed or answered this question, for they were more interested in E-language than I-language and, despite Hermann Paul's psychologizing, worked with an underdeveloped conception of linguistic competence. A full account of opacity must therefore specify what UG resources children draw upon to encode misapplying generalizations in their emergent grammars.

If we are lucky, an answer to this question may provide the key to a larger problem: why are opaque grammars learnable? After all, a prevalent (though not necessarily sound) view holds that opaque generalizations place heavy demands on computing power. Moreover, the idea that a phonological process should apply where its conditioning factors are not met defeats intuitive notions of functional grounding. In this sense, the ability to acquire an opaque grammar may appear maladaptive.

A little thought experiment may be enough to drive this point home. Imagine a world where children cannot learn opaque phonological processes. There is no reason to believe that the communicative efficiency of natural languages would be thereby impaired (claims to the contrary will be dealt with in the course of this chapter). But the acquisition process would be different and, in some crucial respects, simpler. Notably, children would have an easier task of distinguishing the categorical phonological properties of the trigger experience from things such as coarticulation effects, acoustic distortion and the like. Any form of hypocorrection which led to opacity would be grammatically unrepresentable; its very possibility would not even be entertained. Phonological change would therefore be different; the only possible innovations would be those compatible with some transparent reanalysis of the primary linguistic data. As innovation would be held at bay until the opportunity for reanalysis arose, it may even be the case that phonological change would show the same degree of bumpiness as syntactic change (Lightfoot 1999: ch. 4) —instead of its well-known life cycle, whereby innovations are born as Neogrammarian *Lautgesetze*, subsequently acquire morphological sensitivity and even lexical exceptions, and eventually retire as morphological constructions or die as lexical idiosyncrasies (Kiparsky 1984, 1988, 1995; J. Harris 1989; McMahan 1991; Zec 1993).

The SPE solution to our question was as simple as it was drastic: children are born equipped to do internal reconstruction. This is a hackneyed joke, but one that contains a grain of truth. Here, however, I shall adopt an optimality-theoretic approach. In the context of OT the contours of the problem stand out in bold relief because, as has become increasingly apparent, the grammatical apparatus devised by Prince & Smolensky (1993)

does not provide for misapplication effects. Attempts to remedy this flaw have sent phonologists back to the drawing board and have prompted a reconsideration of the essential nature of opacity. As a result, two sharply opposed research programmes have arisen: one holds that misapplication involves the interleaving mechanisms of level segregation and cyclicity, familiar from rule-based LPM; the other resorts to relationships of transderivational correspondence within a strongly parallelist conception of grammar. This chapter is largely taken up with evaluating the relative merits of each proposal. Much of the ensuing debate will seem purely synchronic, concerned with assessing descriptive coverage, typological restrictiveness, theoretical simplicity, and so forth. But I hope to show that, in the event, diachronic considerations must play a significant rôle in the choice of a solution, bearing out our motto: *History matters*.

3.2 DEFINITIONS OF OPACITY: INTENSION AND EXTENSION

3.2.1 Intensional definitions: opacity as misapplication

The concept of opacity first came to the fore during the post-SPE abstractness wars (Kiparsky 1982b [1971]: 75; 1973: 79). Within that context, the term 'opaque' described the application of a phonological rule in relation to the the output of the grammar:

(3.1) Opacity (rule-based definition)

A rule \mathcal{K} of the form $\alpha \rightarrow \beta / \gamma_ \delta$ is opaque to the extent that there are surface representations in the language having

- either (i) α in the environment $\gamma_ \delta$
- or (ii) β derived by \mathcal{K} in an environment other than $\gamma_ \delta$.

In the situation described by clause (3.1i), the grammar generates surface representations which meet the structural description of \mathcal{K} (i.e. $\gamma\alpha\delta$), but which nonetheless fail to undergo the structural change; in such cases, \mathcal{K} is said to 'underapply' (Wilbur 1973) or to be 'non-surface-true' (McCarthy 1998: §1). In the situation described by clause (3.1ii), the grammar generates surface representations which do not meet the structural description of \mathcal{K} (in that the conditioning environment $\gamma_ \delta$ is absent), but which nevertheless have undergone the structural change (in that they contain β derived from α): in such cases, \mathcal{K} is said to

'overapply' or to be 'non-surface-apparent'.¹ Opacity thus arises whenever a phonological rule seems to *misapply*, in the sense that the cause of its application (or non-application) cannot be detected through the mere inspection of underlying and surface forms.

Within the SPE paradigm, opacity is regarded as an epiphenomenon of rule ordering. \mathcal{R} becomes non-surface-true when the application of some later rule \mathcal{S} recreates the structural description of \mathcal{R} . The appearance of underapplication arises because \mathcal{R} , having already been scanned for application, cannot reapply to the output of \mathcal{S} .² In such cases, \mathcal{S} is said to 'counterfeed' \mathcal{R} . McCarthy (1998) distinguishes two types of counterfeeding derivation:

(3,2) Counterfeeding derivations (McCarthy 1998: §4.3)

(a) Counterfeeding in the focus

UR		$\alpha\beta\gamma$
rule \mathcal{A}	$\delta \rightarrow \varepsilon / \alpha_$	—
rule \mathcal{B}	$\beta \rightarrow \delta / _ \gamma$	$\alpha\delta\gamma$
SR		$\alpha\delta\gamma$

(b) Counterfeeding in the environment

UR		$\alpha\beta\gamma\#$
rule \mathcal{C}	$\beta \rightarrow \delta / _ \varepsilon$	—
rule \mathcal{D}	$\gamma \rightarrow \varepsilon / _ \#$	$\alpha\beta\varepsilon\#$
SR		$\alpha\beta\varepsilon\#$

In (3,2a), rule \mathcal{A} targets segment δ in the environment $\alpha_$. UR $\alpha\beta\gamma$ contains the environment $\alpha_$, but not the target δ , and so \mathcal{A} fails to apply. Subsequently, rule \mathcal{B} introduces the target δ in the environment $\alpha_$; but now it is too late for \mathcal{A} to apply. \mathcal{A} becomes non-surface-true. In (3,2b), in turn, rule \mathcal{C} targets segment β in the environment $_ \varepsilon$. In UR $\alpha\beta\gamma\#$, the target β occurs in the non- \mathcal{C} -triggering environment $_ \gamma$, and so \mathcal{C}

¹ Kiparsky's (1973: 79) definition has a third clause, according to which \mathcal{R} is also opaque if the language has surface representations containing β not derived by \mathcal{R} in the environment $_ \delta$. I ignore this clause because, as Kiparsky himself notes, it is not relevant to the status of \mathcal{R} *per se*, but rather addresses the difficulty of determining the underlying correspondent of β in the neutralization environment $_ \delta$; see also Kisseberth (1973: 423), Kaye (1974: 137ff.) and McCarthy (1998: §4.1).

² Post-SPE research in rule-based phonology has found it necessary to relax this constraint on rule application and to admit so-called 'persistent rules', which apply at any point in the derivation where their structural description is met. Notably, syllabification and footing rules are often persistent (see e.g. Itô 1986: ch. 3; Hayes 1995: §5.4.1). Of course, rule persistence has provided an important argument for constraint-based theories of phonology.

fails to apply. Subsequently, rule \mathcal{D} creates the environment $_ \varepsilon$ around the target β , but by now it is too late for \mathcal{C} to apply. \mathcal{C} becomes underapplying in respect of the SR $\alpha\beta\varepsilon\#$.

As regards overapplication, a rule \mathcal{R} becomes non-surface-apparent when some later rule \mathcal{S} destroys the conditioning environment of \mathcal{R} . This is possible because \mathcal{S} cannot consult the derivational history of its own input in order not to destroy the conditioning environment of an earlier rule; similarly, \mathcal{R} cannot look forward so as not to apply in environments which are subject to later transformations. Rules, in other words, are not 'global' (cf. Kisseberth 1973). In cases such as this, \mathcal{S} is said to 'counterbleed' \mathcal{R} . The following abstract illustration comes from McCarthy (1998: §4.2):

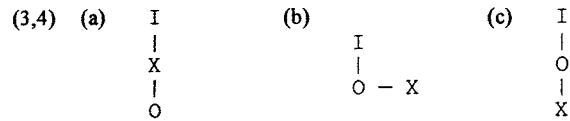
(3,3) Counterbleeding derivation

UR		$\alpha\beta\gamma\#$
rule \mathcal{E}	$\beta \rightarrow \delta / _ \gamma$	$\alpha\delta\gamma\#$
rule \mathcal{F}	$\gamma \rightarrow \varepsilon / _ \#$	$\alpha\delta\varepsilon\#$
SR		$\alpha\delta\varepsilon\#$

Here, \mathcal{E} rewrites β as δ before γ . Subsequently, \mathcal{F} removes the conditioning environment $_ \gamma$. As a result, the SR contains a segment δ derived from β , but the reason for this transformation is not apparent through mere inspection of the SR.

There is a clear strategy guiding this approach to misapplication: if a phonological generalization \mathcal{R} (whose accuracy we assume *ex hypothesi*) appears to misapply in a surface form O , then O must be supposed to reflect the properties of some related representation X to which \mathcal{R} applies appropriately; otherwise, one would simply have to reject the premise that the statement of \mathcal{R} is correct. All responses to opacity in generative phonological theory share this basic assumption. Thus, short of denying the very existence of opaque phonological phenomena (see McCarthy 1998: 9), one is forced to conclude that, for any linguistic expression e , the phonology may require reference to representations other than input(e) and output(e). In other words, misapplication indicates that the phonological mapping of a linguistic expression has access to an *array of representations* encompassing more than the input and the output representations of that expression. Goldsmith (1993, 1996) and G. Lakoff (1993) express the same intuition in terms of 'levels of representation': phonologists, they claim, must countenance a hierarchy of no less than three such levels. However, I have not yet presented any evidence to suggest that the representational array required to account for a misapplying generalization must be

organized serially; at this stage, one could contemplate either arrangement (3,4a) or arrangement (3,4b) or, possibly, arrangement (3,4c):



The choice between these various possibilities lies at the heart of the debate addressed in this chapter. As I shall show below, interleaved OT adopts an approach to opacity which is close to (3,4a), whilst strongly parallel OT propounds a solution close to (3,4b).³

In OT, a negative definition of opacity can be formulated to bring out the link between misapplication and the need to consider an array of representations wider than $input(e)$ and $output(e)$:

(3,5) Opacity (optimality-theoretic definition)

Let there be a language L where a set of input representations $I = \{I_1, I_2, I_3 \dots I_n\}$ is mapped onto a set of output representations $O = \{O_1, O_2, O_3 \dots O_n\}$.

The mapping between I and O is opaque if, and only if, there is no constraint hierarchy \mathcal{M} , where CON is bound by Strict Surface Orientation and the GEN-EVAL function is strongly parallel, such that O is the optimal output with respect to I under \mathcal{M} .

Under this definition, the category of opacity comprises those phonological systems which cannot in principle be described by any optimality-theoretic grammar under the twin assumptions of Strict Surface Orientation (2,45) and Strong Parallelism (3,6).

Strict Surface Orientation, as stated in (2,45), limits the composition of CON to output markedness constraints and IO-faithfulness constraints. From this restriction, it follows that unfaithfulness to the input, leading to phonological alternations, can only be triggered by markedness constraints when the appropriate conditioning factors are present in the output; recall my observations on 'aatiffaithfulness' in §2.3.4.2. Conversely, an output representation may violate a markedness constraint M only if M is dominated by an

³ Arrangement (3,4c) is reminiscent of the sequence Numeration/Spell-out/LF in minimalist syntax.

antagonistic markedness or IO-faithfulness constraint. Possible correspondence relationships with representations other than the input play no rôle; (3,4b) is accordingly out. Under Strict Surface Orientation, therefore, a single pass through GEN and EVAL cannot possibly generate opacity.

Strong Parallelism, in turn, decrees that the pairing of underlying and surface forms must be effected in a single step, with no intermediate representations of any sort:

(3,6) Strong Parallelism

The functions GEN and EVAL are non-recursive; for any linguistic expression e , the phonological mapping between $input(e)$ and $output(e)$ is determined by means of a single pass through GEN and EVAL.

Let us represent a pass through GEN and EVAL as a phonological function ρ . Strong Parallelism, as formulated in (3,6), rules out a grammatical arrangement such as (3,4a), where the mapping $I \rightarrow O$ requires two applications of ρ : i.e. $X = \rho(I)$ and $O = \rho(X)$. Note, in contrast, that arrangement (3,4b) is permissible under Strong Parallelism, for EVAL can check the relationship of O with both I and X simultaneously: i.e. $O = \rho(I, X)$.

The existence of opacity as defined in (3,5) turns out to be an empirical fact; the logical conjunction of Strict Surface Orientation and Strong Parallelism is, by the same token, false. This implies that classical OT, as presented in Prince & Smolensky (1993), is excessively restrictive and requires substantial revision: minimally, either Strict Surface Orientation or Strong Parallelism must be relaxed.

3.2.2 Two optimality-theoretic strategies against opacity

Unsurprisingly, OT researchers have divided into two camps over the solution to opacity: some endeavour to retain Strong Parallelism at the expense of Strict Surface Orientation, whilst others adopt exactly the opposite strategy.

3.2.2.1 Upholding Strong Parallelism

Strong Parallelism is widely regarded as a highly desirable property for phonological theory, on the grounds that serial derivations are 'cognitively implausible' and 'computationally intractable' (see Orgun 1996a: §2.2 and references therein). Within OT, the research programme devoted to upholding Strong Parallelism is therefore extremely

active. The impetus propelling this line of inquiry comes from the concept of correspondence:

(3.7) **Correspondence (McCarthy & Prince 1995: §2.1)**

Given two strings S_1 and S_2 , **correspondence** is a relation \mathfrak{R} from the elements of S_1 to those of S_2 . Elements $\alpha \in S_1$ and $\beta \in S_2$ are referred as **correspondents** of one another when $\alpha \mathfrak{R} \beta$.

Correspondence is a generalization of the notion of faithfulness: correspondence constraints can demand identity between any pair of related phonological representations —not just between the input and the output. Accordingly, Correspondence Theory makes it possible to establish and regulate any number of parallel identity relationships between the output and other representations in a given array. In the light of §3.2.1, the relevance of such a formal tool to the problem of misapplication is self-evident. At the same time, the enrichment of CON with correspondence constraints other than IO-faithfulness transgresses the stringent limits set by Strict Surface Orientation, as defined in (2,45).

The first moves within strongly parallel OT were directed at ‘paradigm effects’.⁴ These can be loosely defined as phonological misapplication effects associated with morphological operations of affixation, compounding, truncation, etc. In British Received Pronunciation, for example, there is a generalization whereby post-*ŋ* *g* is deleted unless it occupies an onset position. This generalization overapplies in the suffixed deverbal noun *banger* [bæŋə] (cf. *Bangor* [bæŋgə] ‘name of a town in Wales’); in the verb *bang* [bæŋ], from which *banger* is derived, *g*-deletion applies correctly. Similarly, coronal non-continuants are alveolar in Belfast English, but become dental when followed by a rhotic (J. Harris 1989: 40-41); this generalization underapplies in the noun *tenner* [tɛːnəɹɪ] ‘£10 note’, derived from the numeral *ten* [tɛːn] through suffixation; cf. *tenor* [tɛŋəɹɪ]. Strongly parallel proposals in the area of paradigm effects include Generalized Alignment Theory and Transderivational Correspondence Theory (henceforth TCT). In Generalized Alignment Theory (McCarthy & Prince 1993b), the constraints of the ALIGN family are allowed access to morphological categories, whose edges may be required to coincide with prosodic boundaries of a designated type; paradigm effects may accordingly be triggered as a

morphological domain becomes phonologically ‘closed off’ (for ‘prosodic closure’, see Benua 1997: §5.3 and references therein). TCT (see e.g. Benua 1995, 1997; Itô & Mester 1997a; Kager 1995; Kenstowicz 1996) launches a more direct attack on morphologically induced misapplication by setting up correspondence relationships between the output representations of morphologically related lexical items. These relationships are monitored by OO-correspondence constraints. In Belfast English *tenner* [tɛːnəɹɪ], for example, it is assumed that pre-rhotic dentalization underapplies in order to preserve identity with the place of articulation of the nasal in the output representation of the base *ten* [tɛːn] (Benua 1997: §5.3.1.2).

Proponents of strongly parallel OT acknowledge, however, that TCT fails to provide a full answer to the opacity problem. In a subset of opacity effects, phonological misapplication fails to correlate with the transparent presence or absence of the relevant conditioning environment in the output representation of a related lexical item (see McCarthy 1998: 7 and references therein). In response to these recalcitrant difficulties, Sympathy Theory (McCarthy 1998) proposes yet a further extension of correspondence relationships. McCarthy posits a novel type of constraint requiring that the optimal output should be faithful to a suboptimal co-candidate, the ‘sympathy candidate’ or \otimes -candidate, which is in turn selected by a designated low-ranking IO-faithfulness constraint.

The general strategy of strongly parallel OT is, in sum, to define new relationships of correspondence between the output representation and other representations independently required by the theory: either paradigmatically related output representations, or representations present in the set of co-candidates supplied by GEN. The function EVAL checks all these relationships in parallel.

3.2.2.2 Upholding Strict Surface Orientation

Strongly parallel OT has been vigorously challenged by more conservative optimality-theoretic models such as LPM-OT (Kiparsky 1998a) and SBPM (Orgun 1996a, b). These models accept Prince & Smolensky’s (1993) basic conclusion that constraint ranking is superior to either rule ordering or parameter setting as a means of stating phonological generalizations. However, they are relatively conservative in stressing that interleaving is

⁴ I borrow this term from Kiparsky (1998a).

a fundamental property of the interface between phonology and morphosyntax.⁵ Essentially, interleaving can be defined as the recursive application of phonology within a nested hierarchy of morphosyntactically defined domains; it is instantiated in grammars through level segregation and cyclicity. In interleaved OT, surface misapplication effects emerge from the input-output relationship holding between the phonological representations associated with phonological domains at different levels of embedding.

Interleaving is strictly incompatible with Strong Parallelism. At the same time, both LPM-OT and SBPM retain Strict Surface Orientation and reject transderivational correspondence relationships of any sort. Indeed, previous research in the LPM paradigm bequeaths to interleaved OT the programmatic goal of minimizing, or altogether eliminating, any overt reference to morphological categories in the phonology. LPM aimed to reduce all morphological conditioning effects to interleaving, where the only way in which morphological structure influences the phonology is by creating a levelled hierarchy of (recursive) domains; see Anderson (1992: ch. 9, §10.2.4). Hence, interleaved OT can be expected to prevent constraints from referring to the internal morphological structure of phonological domains.

In sections 3.2.2.1 and 3.2.2.2, I have outlined the claims of strongly parallel OT and interleaved OT very much in their own terms. These claims, as well as the metatheoretical arguments with which they are supported, will be put to the test and compared below, in sections 3.3, 3.4 and 3.5.

3.2.3 The extension of opacity in Optimality Theory

Definitions (3,1) and (3,5) are *intensionally* similar; implicit in both is a characterization of opacity in which phonological generalizations appear to misapply in surface forms because the latter reflect properties of representations other than the input. At the same time, the denotation of the term ‘opacity’ is bound to vary according to the technical vocabulary in which phonological generalizations are formulated. In this connection, the expressive possibilities of rules and constraints are significantly different. In terms of

⁵ The reader is warned that, throughout this dissertation, ‘morphosyntax’ is used in a wide sense as a cover term for both syntax and morphology. The term ‘morphosyntax’ also has a narrow sense, in which it denotes those syntactic phenomena whose expression is morphological (e.g. agreement); this meaning will not be relevant here.

extension, therefore, rule-based and constraint-based definitions of opacity are bound not to coincide. Indeed, certain phonological phenomena which were previously regarded as opaque can be described in OT in terms compatible with both Strict Surface Orientation and Strong Parallelism; under (3,5), these phenomena must be regarded as transparent. Thus, the class of opaque phenomena is *smaller* in OT than in rule-based phonology. This point must be borne in mind, for it has a direct bearing on the empirical evaluation of strongly parallel OT and interleaved OT, as we shall see in §3.3.3.1.

Like Kager (1997: 497), therefore, I see no good grounds for deciding *a priori* that optimality-theoretic responses to opacity should provide a unified account of the set of phenomena picked out by a rule-based definition.

The most serious problem [for stratal responses to opacity in OT —RBO] [...] is that OT+serialism does not provide a uniform theory of non-surface-apparentness and non-surface-trueness. Rule-based serialism has a single explanation for why phonological generalizations can be non-surface-true or non-surface-apparent: subsequent rules have hidden the generalization or the conditions leading to its applicability. OT+serialism has disparate explanations, in a thoroughly unprincipled way: constraint domination accounts for some non-surface-true generalizations, while stratal ordering must be called on to account for the remaining non-surface-true generalizations and all the non-surface-apparent ones. In this respect, OT+serialism is clearly inferior, explanatorily, to rule-based serialism. [...] OT's account of opacity should be no worse than rule-based serialism's.

McCarthy (1998: 10)

Essentially, McCarthy charges phonological models such as interleaved OT with an opportunistic use of serialism: whilst rule ordering was the only answer to opacity in SPE, interleaved OT deals with a subset of the same phenomena through constraint ranking, and with another subset through interleaving. However, McCarthy's reasoning implies that phonological generalizations can be expressed in a metalanguage *M* independent from both rule-based theory and OT, where *M* provides criteria for (i) deciding whether or not a generalization is ‘linguistically significant’, and (ii) labelling each significant generalization as ‘opaque’ or ‘transparent’. It would then be the task of an explanatorily adequate phonological theory to provide a unified account for the set of generalizations labelled as opaque in *M*.

At the present time, however, there is no such thing as *M*; there is no pretheoretic or metatheoretic procedure to determine the denotation of the category ‘opacity’. Whilst the concepts of opacity deployed in rule-based theory and OT are intensionally analogous, in

the sense specified above, their extension can only be determined in terms of the technical vocabulary of each theory. It is therefore to be expected that the class of opaque phenomena will vary in size from rule-based theory to OT. McCarthy thus makes a category mistake: a theory-neutral *intensional* definition of opacity (McCarthy 1998: 1) is erroneously assumed to pick out the same set of *denotata* in any phonological theory. Of course, one could conceivably construct a valid argument to show that loss of generalization is incurred by splitting the class of opaque rules in (say) SPE, but such an argument would rest on the premise that all the members of that class share some logically independent empirical property. As we shall see presently, McCarthy chooses to deny this premise.

Let us consider a (deliberately oversimplified) example. In rule-based theory, the Nzebi data in (2,47) could be analysed as involving a series of three rules *A*, *B* and *C*, with each rule counterfeeding the preceding rule in the focus:

(3,8)	UR	sali	βe:di	beti
	rule <i>A</i>	e → i / _C ₀ i	—	—
	rule <i>B</i>	ε → e / _C ₀ i	—	βe:di
	rule <i>C</i>	a → ε / _C ₀ i	sali	—
	SR	sali	βe:di	biti

In OT, in contrast, this chain shift can be described in a strongly parallel, strictly surface-oriented (and therefore transparent) fashion, thanks to the notion of relativized faithfulness (§2.3.4.2), instantiated through constraint conjunction; see tableau 4 in Kirchner (1996: 346).⁶ Whether or not Kirchner's solution is right must be judged by the answers to questions such as the following: does the constraint conjunction analysis generalize to all synchronic chain-shifts?; can the constraint conjunction analysis express the concepts of push-chain and drag-chain?; more generally, does local conjunction make correct

⁶ I deliberately ignore rule-based alternatives to (3,8). It is likely, for example, that *A*, *B* and *C* could be collapsed using segmental features in combination with Greek letter variables or some other notational convention, thereby avoiding the loss of generalization inherent in repeating the same conditioning environment in three rules. Similarly, rule opacity could conceivably be eliminated by resorting to multivalued features, or to spreading of [j] or deletion of [a] in some version of Particle Phonology (see e.g. Schane 1995). This is irrelevant to my present purposes. The critical point is that Kirchner (1995, 1996) does not resort to Greek letter variables, or to multivalued features, or to particle-based analyses; his transparent reanalysis of the Nzebi shift crucially relies on a conceptual tool (viz. locally conjoined faithfulness constraints) which is available to OT, but not to rule-based theory.

typological predictions?; are phonological grammars incorporating conjoined constraints learnable?; etc. It really makes no difference that the Nzebi chain-shift is opaque under analysis (3,8). But, by the same token, if Kirchner's transparent analysis proves correct, one cannot justify the demand that OT should deal with data-sets such as the Nzebi shift under the heading of opacity. More generally (and somewhat tautologically), an optimality-theoretic response to opacity need only address those phenomena which are intractably opaque in classical OT, i.e. those covered by definition (3,5).

In rule-based theory, moreover, the concept of opacity has been criticized as being empirically empty. Notably, Kaye (1974) challenges Kiparsky's (1971: §3) claim that opaque phenomena are diachronically unstable; according to Kaye, certain types of derivation covered by (3,1) are highly natural. McCarthy shares this opinion:

One problem is that some opaque generalizations have exactly the same character as transparent generalizations, except for being opaque. Thus, the claim that opaque generalizations have a distinct status is an empty one, since nothing correlates with this putative distinction.

(McCarthy 1998:9)

However, McCarthy overstates the case. At present, the concept of opacity is not entirely unproductive: despite the anomalies pointed out by Kaye, historical linguists still regard opacity as the main contributory factor in the morphologization of phonological processes (see e.g. Anderson 1992: §13.2). In this sense, a revision of the extensional boundaries of opacity may be instrumental in refining the proposals of diachronic phonology and morphology. Paraphrasing McCarthy's quotation, the hope is that those putatively opaque generalizations whose character proves identical, both synchronically and diachronically, to that of transparent generalizations will cease to be covered by our definition. This provides an additional reason not to hypostatize either pretheoretical or rule-based characterizations of any given phonological process as transparent or opaque.

Research in classical OT has resorted to a wide range of strategies to reduce the number of intractably opaque phonological phenomena; McCarthy (1998: §2) offers a comprehensive survey. To assess the extent of opacity as defined in (3,5), we need only be concerned with those solutions which observe both Strict Surface Orientation and Strong Parallelism. Some of these strategies are in fact not specific to OT, but rely on resources, representational or otherwise, also available to rule-based theory: e.g. one-step coalescence analyses of assimilation with deletion of the triggering segment.

(3,9) Vowel nasalization with deletion of the triggering nasal consonant⁷

(a)	<i>Opaque (counterbleeding) analysis</i>		<i>input</i> VN] _σ
	<i>A</i>	V → \tilde{V} / $__N]_{\sigma}$	$\tilde{V}N]_{\sigma}$
	<i>B</i>	N → \emptyset / $\tilde{V}__$	$\tilde{V}]_{\sigma}$
(b)	<i>Transparent (coalescence) analysis</i>		
		VN → \tilde{V} / $__]_{\sigma}$	$\tilde{V}]_{\sigma}$

In other cases, however, the conceptual tool required to dispel opacity was unavailable to previous theoretical frameworks: e.g. relativized faithfulness.

One must also note that certain solutions have been proposed which, though respecting Strong Parallelism and Strict Surface Orientation, weaken OT in other undesirable ways. A case in point is Zec (1995a), who analyses a set of opaque compensatory lengthening phenomena in Pāli; see (3,66) and (3,67). Assuming a Containment model of IO-faithfulness (McCarthy & Prince 1993a: 20; 1993b: §2), Zec assumes that stray segments need not be phonologically inert, but may be visible to markedness constraints requiring mora projection. Zec's account of the Pāli phenomena proves empirically inadequate; it cannot cope with compensatory lengthening affecting rising-sonority coda-onset clusters (see Zec 1995a: note 13). More seriously, allowing unparsed material to be phonologically active would wreak havoc in CON: blindness to stray material would have to be explicitly written into a host of constraints, including alignment as well as OCP-related constraints (see McCarthy & Prince 1995: §2.2).

3.2.4 Conclusion

In this section, I have characterized opacity in terms of the apparent misapplication of a phonological regularity on the surface. This implies that the phonological mapping between the underlying and surface forms of a given linguistic expression may require access to other representations. I have proposed an optimality-theoretic definition of opacity which

⁷ Kaye (1974: 141) upholds the counterbleeding analysis; coalescence has been defended by, among others, Kawasaki (1986) and Ohala (1992: 328). Kaye suggests that, historically, the appearance of $\tilde{V}]_{\sigma}$ is always preceded by a period with $\tilde{V}N]_{\sigma}$, but this assertion proves incorrect if one distinguishes categorical nasalization (i.e. spreading of a nasal autosegment in the phonology) from mere gestural overlap. It is therefore interesting that assimilation with deletion of the triggering segment should provide Kaye (1974) with his main argument against Kiparsky's (1971: §3) claim that opaque phenomena are diachronically unstable. This fits in rather well with my suggestion that a change (more specifically, a reduction) in the denotation of opacity may increase the empirical yield of the concept.

reflects this insight by stating that opaque phonological phenomena cannot be described under the logical conjunction of Strict Surface Orientation and Strong Parallelism. In other words, if we only countenance output markedness and IO-faithfulness constraints, and at the same time we assume that the mapping between underlying and surface forms is effected by means of a single pass through GEN and EVAL, then misapplication phenomena become rigorously intractable. This state of affairs has given rise to two competing research programmes in OT: interleaved OT flatly rejects Strong Parallelism; strongly parallel OT, in contrast, tackles misapplication by positing transderivational correspondence relationships, in violation of Strict Surface Orientation.

However, a few generalizations which appear opaque in rule-based theory become transparent when expressed in optimality-theoretic terms. This shift in the denotation of the term 'opaque' has a direct bearing on the empirical remit of optimality-theoretic responses to the problem. It may also have an important impact on the empirical productivity of the concept itself, particularly for diachronic linguistics.

The following sections are taken up with a comparison between interleaved OT and strongly parallel OT. An important methodological point will emerge from this debate: the evaluation of these two models must not only rely on synchronic criteria such as simplicity, generative power, typological adequacy and learnability; rather, it must also take into consideration what each programme has to offer in relation to the diachronic processes fuelling the rise and fall of opaque phonological phenomena. The reason for bringing historical issues to bear on the debate is twofold. First, phonological change constitutes an *explanandum* for linguistic theory in its own right. Secondly, and more pressingly, we shall have reason to believe that certain recurrent, though contingent, patterns of change (notably the tendency for phonological generalizations to become increasingly sensitive to morphology) play a key rôle in limiting the complexity of opaque effects in natural language. If, in the pursuit of restrictiveness, a grammatical model attempts to attribute the simplifying effects of contingent diachronic processes to formal UG principles, it is bound to come up against crucial counterexamples where its putative UG principles break down.

As hinted in §3.1, a related question looms on the horizon: why does the language faculty provide the computational resources to learn and use misapplying phonological generalizations, thereby opening the door for change to create opacity? Is opacity functionally grounded, as most markedness constraints are thought to be? Or is it simply a side effect (possibly maladaptive, possibly not) of more essential design features of grammars?

3.3 OPTIMALITY THEORY WITH INTERLEAVING

3.3.1 Conceptual issues

Numerous authors have, in different circumstances, suggested the possibility of incorporating grammatical mechanisms such as level segregation and cyclicity into OT (e.g. Baković 1995; Black 1994; Booij 1996, 1997; Kenstowicz 1995; Kiparsky 1998a; Lin 1997; McCarthy & Prince 1993a; Orgun 1994, 1996a, b; Potter 1994; Rubach 1997; Sprouse 1997). In this dissertation, however, the term 'interleaved OT' will be used in a relatively specific sense, denoting those optimality-theoretic phonological models which explicitly adopt interleaving as the fundamental principle governing the interface between phonology and morphosyntax: e.g. LPM-OT (Kiparsky 1998a) and SBPM (Orgun 1996b). Interleaving can admittedly be implemented in a wide variety of ways: compare, for example, the proposals of interactionist LPM (e.g. Pesetsky 1979; Kiparsky 1982a, 1985; Mohanan 1982, 1986), non-interactionist LPM (e.g. Halle & Vergnaud 1987a, Halle & Kenstowicz 1991, Odden 1993), prosodic LPM (e.g. Inkelas 1989, 1993), and SBPM (Orgun 1994, 1996a,b); see Kaisse & Hargus (1993) for enlightening discussion. However, such particulars of implementation can by and large be ignored in the context of a comparison between interleaved OT and strongly parallel OT.

3.3.1.1 Interleaving, level segregation, cyclicity

In general terms, the phonology of a language can be conceptualized as a function ρ —or, as we shall see presently, a set of functions $\rho = \{\rho_1, \rho_2 \dots \rho_n\}$ —which associates (an infinite set of) input representations with their corresponding output representations. In rule-based theory, ρ is conceived of as a battery of ordered rules; in OT, as a pass through GEN and EVAL under a given hierarchization of CON. A phonological theory incorporating interleaving asserts that, given a linguistic expression e with the phonological input representation I , the corresponding output representation O is not defined by a single application of ρ to I ; rather, ρ applies recursively within a nested hierarchy of phonological

domains defined upon I by the morphosyntactic constituent structure of e .⁸ Thus, if the morphosyntactic structure of e defines the hierarchy of phonological domains $[[[x][[y]z]]]$ on I , then the claim is that $\rho(I) = \rho(\rho(x), \rho(\rho(y), z))$. In a serialist interpretation of this notion, ρ is assumed to apply first to the smallest, most deeply embedded domain(s) within I , and then to reapply iteratively to progressively larger, less deeply embedded domains. As we shall see in §3.3.1.3, interleaving does not preclude a declarative approach to phonology. The view that interleaving governs the phonology-morphosyntax interface is nevertheless strictly incompatible with Strong Parallelism, as defined in (3.6).

In generative phonology, interleaving has been instantiated in two distinct ways: through level segregation, and through cyclicity. The term 'level segregation' refers to the idea that the phonology of a language does not consist of a single function ρ , but rather of a set of distinct functions $\rho = \{\rho_1, \rho_2 \dots \rho_n\}$; the specific function $\rho_i \in \rho$ applying to a domain δ_i correlates with the type of morphosyntactic construction which defines δ_i .⁹ In rule-based terms, Anderson (1992: 244) characterizes level segregation as follows:

Rules of the phonology can be divided into subsets, not necessarily disjoint. To each such subset, we can assign a phonological 'type'. The rules of the morphology can be similarly divided into (corresponding) subsets, where the type of a morphological rule is determined by the set of phonological rules that are applicable to the result of applying that rule.

In other words, the morphosyntactic constructions of a language can be assigned to different levels depending on the phonological function which applies to the corresponding phonological domains.

Research in the LPM tradition suggests that phonological systems comprise minimally (typically?) three levels: the stem level and the word level (which together constitute the lexical phonology), and the postlexical level (see e.g. Kiparsky 1985, Booij & Rubach 1987, Borowsky 1993). However, analyses of numerous languages distinguish more than two lexical levels: according to Inkelas & Orgun (1995), for example, the lexical phonology of Turkish comprises four distinct levels. If such analyses are correct, we may

⁸ Observe that the concept of phonological domain, defined as a portion of linguistic material within the scope of a single application of ρ , is logically and empirically distinct from the notion of *prosodic constituent*, ranging over the categories μ , σ , Σ , ω , etc.; see (2.2). For insightful discussion, see Inkelas (1993).

⁹ I use 'level segregation' in deliberate contrast with the more familiar term 'level ordering', which I reserve for the combination of level segregation with the Affix Ordering Generalization; see §3.3.1.2.

reserve the terms 'stem level' and 'word level' for the first and last lexical levels, respectively.¹⁰ In the case of English, however, the need for more than two lexical levels, as advocated in Halle & Mohanan (1985), has never been convincingly demonstrated; see Kiparsky (1985: note 2). Indeed, I am not aware of any language requiring more than three phonological levels under any possible analysis. The issue of the number of levels hinges to a large extent on the treatment of non-derived environment blocking, and will not prove crucial here.

In a three-level model, the phonology of the stem level applies to domains with the lowest degree of morphological complexity: notably, those defined by monomorphemic lexical items.¹¹ In interleaved OT, the properties of underlying representations (which in SPE were the province of problematic 'Morpheme Structure Constraints') will be derived from stem-level constraint rankings through Lexicon Optimization (see §2.2.2.1);¹² of course, this includes whether or not a given feature is phonemically contrastive (Kirchner 1995, Pulleyblank 1997: §5). Additionally, many languages have morphological constructions which are subject to the same phonological well-formedness conditions as monomorphemic lexical items (plus possible derived environment effects):¹³ e.g. class-I affixation in PDE. J. Harris (1994: §1.4.3) and Kaye (1995: §2.2) would describe such morphological constructions as 'non-analytic'; in interleaved OT they are assumed to define stem-level phonological domains (see §3.3.2 below). One may refer to these domains as 'phonological stems' or 'P-stems'; see Sproat (1985).

¹⁰ For observations on the meaning of the words 'first' and 'last' in this context, see footnote 15.

¹¹ A lexical item is understood here as a member of a lexical category: i.e. Noun, Verb, Adjective, Adverb (Kiparsky 1998a: 30). Lexical items constitute the minimal phonological domains: a bound root such as Spanish /desdeɰ/ 'disdain' (whence the noun *desdén* and the verb *desdeñar*; see §3.3.2) does not define a phonological domain because it is not a member of a lexical category (J. W. Harris 1983, Kiparsky 1982a).

¹² Observe that, under interleaving, input-output mappings are recursive; I therefore assume that, in interleaved OT, Lexicon Optimization will be a recursive process. The input to the last phonological level L_n will be projected from surface representations by Lexicon Optimization; of course, this will involve the factoring out of any alternations attendant on morphosyntactic constructions ascribed to L_n . But input(L_n) equals output(L_{n-1}). The next logical step is, therefore, to project input(L_{n-1}) from input (L_n) by Lexicon Optimization, undoing any alternations triggered by L_{n-1} -level morphosyntax. The process iterates until underlying representations are reached. See §3.3.3.2 for discussion.

Kiparsky (1982a: 159-159, 167-168) offered a different solution to the Duplication Problem (Clayton 1976, Kenstowicz & Kisseberth 1977). He established the connection between underlying representations and stem-level phonology in terms of underspecification.

¹³ Determining the appropriate treatment of derived environment effects in interleaved OT falls beyond the scope of this dissertation. For recent proposals in this area, see e.g. Kiparsky (1993), Inkelas & Orgun (1995), Orgun (1996b), and Lubowicz (1998).

The phonology of the word level, in turn, governs the phonological structure of words, understood here as syntactically free (inflected) lexical items. Crucially, a number of morphosyntactic constructions (including cliticization) may define phonological domains subject to the same well-formedness conditions as words, but not to constraints holding at earlier lexical levels such as the stem level: e.g. class-II affixation in PDE. We say that such morphosyntactic constructions create word-level phonological domains. Again, one may refer to such domains as 'phonological words' or 'P-words', on the understanding that this is not the same concept as that of the metrical constituent ω ; see footnote 8.

Finally, the postlexical level deals with the phonological properties of domains arising over complex syntactic constructions. Such domains may be collectively referred to as 'P-phrases'.

The crucial point about level segregation is, in sum, that different levels impose different well-formedness conditions upon the corresponding phonological domains. In optimality-theoretic terms, each level is associated with a different hierarchization of CON. Levels can, in this sense, be conceptualized as autonomous cophonologies (Orgun 1996b: ch. 4).

Alongside level segregation, interleaving can also be instantiated through cyclicity. Indeed, cyclicity may simply be defined as intralevel interleaving; it arises when the phonological domain structure of a given level is recursive.¹⁴ In PDE, for example, the stem level is cyclic; accordingly, P-stems are recursive: e.g. [_{P-word}[_{P-stem}[_{P-stem}[_{P-stem} *origin*]*al*]*ity*]]. In early versions of LPM, cyclicity was regarded as an intrinsic property of certain levels: see e.g. Booij & Rubach (1987), where the stem level is characterized as 'cyclic', and the word level as 'postcyclic'. Difficulties with this approach have led other authors to suggest that cyclicity must be stipulated for individual levels on a language-particular basis (e.g. Halle & Mohanan 1985). Orgun (1994, 1996a,b) argues that, within the lexicon, cyclicity follows from independently motivated morphological structure (in particular, see Orgun 1996a: §5.3, 1996b: ch. 2); this proposal, though extremely appealing, has not yet been adequately tested.

¹⁴ The term 'cyclicity' has also been used to refer to situations where the same phonological generalization applies at two consecutive levels: i.e. 'interstratal cyclicity' (Odden 1993: 115). For the sake of clarity, I reserve the term 'cyclicity' for 'stratum-internal cyclicity'.

3.3.1.2 The arbitrariness of cophologies

Cophologies are powerful descriptive devices; from a methodological standpoint, it makes sense to attempt to impose some kind of formal check on them. One possible strategy is to require that the assignment of specific morphosyntactic constructions to particular phonological levels should be backed up with non-phonological criteria. In this respect, the verbal system of the Bantu language KiRundi (Goldsmith 1990: 241-243) is satisfyingly well-behaved. In KiRundi, the treatment of VV sequences differentiates three phonological levels: the stem level has fusion (e.g. $a + i \rightarrow e$), the word level has assimilation (e.g. $a + i \rightarrow i$), and the postlexical level has deletion (e.g. $e + a \rightarrow a$). Significantly, the stem-level phonology applies within a domain defined by the verbal root ('radical') plus any derivational suffixes ('extensions'); word-level phonological domains comprise the verbal stem together with a series of inflectional prefixes (subject, tense, and object markers); and, finally, the postlexical phonology applies over higher syntactic units.

Disappointingly, such narrow convergence between the phonological and morphosyntactic properties of specific constructions is by no means the rule. In PDE, for example, it has been observed that stem-level morphological processes tend to be less productive and semantically transparent than word-level ones (see Benua §5.6.1 and references therein). Similarly, early work in the LPM paradigm assumed that level membership controlled the distribution of affixes, with stem-level affixes always standing nearer the root than word-level affixes —this is the Affix Ordering Generalization of Siegel (1974). However, whilst these generalizations embody interesting tendencies, none has proved tenable as a formal principle of grammar (for a refutation of the Affix Ordering Generalization, see e.g. Fabb 1988).¹⁵

This state of affairs (which we may call the 'cophology arbitrariness' problem) might be taken to provide a methodological argument against level segregation.

¹⁵ Observe that, in a three-level system such as is described in §3.3.1.1, it makes sense to assert that the stem level 'precedes' the word level even if the Affix Ordering Generalization is disobeyed, for a root must necessarily be assigned to a lexical category before it can acquire functional features. Consider, for example, the derivation of the singular noun *desdén* 'disdain' from the root /desdeɲ/ in Spanish velarizing dialects (see §3.3.2 for full details). First, the root is assigned to the lexical category N and, in so doing, undergoes stem-level phonology, including depalatalization of the coda nasal: i.e. [ɲdes.'den]. Next, the nominal stem is assigned inflectional features, such as singular number, and this triggers the application of word-level phonology, including velarization of the coda nasal: i.e. [ɲdes.'deɲ] 'disdain' sg. Thus, even if Spanish exhibited violations of the Affix Ordering Generalization such as are found in PDE, there would still be a basis for regarding the stem level as 'higher' than the word level.

[... T]he Lexical Phonology notion of stratum (Kiparsky 1982[a], Mohanan 1982) is trivialized by the kind of stratum that, say, Tiberian Hebrew or Levantine Arabic would require — a stratum of convenience rather than a meaningful correlation of phonological and morphological factors.

McCarthy (1998: 10)

The morphosyntactic arbitrariness of cophologies, however, does not reveal a flaw of interleaving as a theoretical principle, but is rather a fact of natural language with which any phonological model must reckon. Work in TCT, for example, has acknowledged the need to segregate morphological constructions into different correspondence classes on purely phonological grounds, each class being controlled by a specific set of OO-correspondence constraints (e.g. Benua 1997). Interestingly, Benua resigns herself to imposing no formal restriction on the proliferation of correspondence classes; such limitations as may exist stem from external factors:

Under the present proposal, arbitrarily many distinct [OO-]faithfulness relations can coexist in the same language. In the limit each morpheme in the lexicon can be associated with a unique correspondence relation. This freedom is constrained by learnability: morphemes class together to limit the demands on the language learner.

Benua (1997: 109-110)

(For analogous claims within interleaved OT, see Orgun 1996b: §4.4.7.) Similarly, Benua acknowledges the existence of tendencies correlating phonological classhood with productivity, semantic transparency, ordering, etc.; but she suggests that such tendencies may require external explanations:

I suggest that the connections between phonological transparency, semantic compositionality and productivity are either functional, or ruled formally by some component of the grammar other than the phonology. The constellation of phonological transparency, productivity and compositionality is interesting and important, but the semantic and morphological facts have no formal status in the phonological component.

Benua (1997: 207-208)

In sum, phonological models incorporating interleaving have failed to discover reliable non-phonological diagnostics of the level to which specific morphosyntactic constructions are assigned; similarly, there is uncertainty as to the number of levels which can coexist within one grammar. However, this arbitrariness does not provide an argument

in favour of strongly parallel OT against interleaved OT. Both theories acknowledge the need to postulate cophonologies whose morphosyntactic membership may be formally arbitrary, and both rely on external constraints to check cophonology proliferation.

3.3.1.3 Derivational vs declarative implementations of interleaving

Interleaved OT has been charged with *derivationalism*; it is suggested that interleaved phonological models cannot avoid serial computation, which is regarded as psychologically implausible and as antithetical to the programmatic goals of OT research (see e.g. Itô & Mester 1997a: §1.1). However, this argument rests on a misconception: recursion of the function ρ (defined, in the case of OT, as a pass through GEN and EVAL under a given hierarchization of CO:N) need not imply serialism.¹⁶ Orgun (1994; 1996a, b) has made it absolutely clear that interleaving can be straightforwardly implemented in a declarative fashion.

To begin with, Orgun (1996b: 16) pointedly notes that “any computation that can be done, can be done nonderivationally”. It is quite possible, for example, to implement a traditional SPE derivation in a fully parallel fashion: one must simply reinterpret the output of each rule as a separate level of representation, and rules themselves as correspondence constraints holding between adjacent levels. In terms of psychological plausibility, there is obviously no advantage to be gained by implementing SPE derivations in this manner. However, Orgun's thought-experiment reminds one of something fundamental: grammars are mathematical functions; “It is not incumbent upon a grammar to compute, as Chomsky has emphasized repeatedly over the years” (Prince & Smolensky 1993: 197). It is only at the level of computational implementation that the problem of parallel vs serial processing becomes relevant. However, if we are interested in assessing the generative power of a phonological theory, it is more useful to focus on the number of distinct representations postulated by the theory, and on their status in relation to other components of grammar. As Orgun (1996b: 17-8) puts it,

¹⁶ In this connection, it is significant that, within strongly parallel OT, the so-called ‘primacy of the base’ in paradigm effects compels TCT to postulate recursions of the constraint hierarchy \mathcal{M} : i.e. $\mathcal{M}_{\text{base}} \gg \mathcal{M}_{\text{derivational}}$ (Benua 1997). Thus, Strong Parallelism (i.e. non-recursion of GEN and EVAL) is achieved at the expense of recursion of the constraint hierarchy within a single application of ρ . See §3.4.1.2 for discussion.

If nonderivationalism is achieved only by proliferating levels in an ad-hoc manner, then something is wrong with the theory. A derivational theory might be undesirable, but so is one that uses ad-hoc representations that have no independent motivation.

This criterion proves particularly relevant to a comparison between strongly parallel OT and interleaved OT, as both models postulate an impoverished array of representations in comparison with rule-based theory.

In interleaved OT, all phonological representations have independent morphosyntactic justification, for all of them arise over domains defined by some morphosyntactic construction. Thus, every phonological representation at the lexical level corresponds either to a root or to a morphologically complex lexical entry (a stem or a word); see Orgun (1996b: 18). Interestingly, interleaving allows phonological representations to be assigned to *derived bound stems*. These phonological representations are neither underlying nor surface forms, but their existence is motivated by their association with an independently required lexical entry. In turn, the existence of derived bound stems as morphological entities is beyond dispute: e.g. they are referred to in the subcategorization frames of affixes.¹⁷

From a morphosyntactic viewpoint, in contrast, the array of phonological representations postulated by strongly parallel OT models is at once too small and too large. First, TCT attempts to reduce all paradigm effects to correspondence between morphologically related output forms; derived bound stems are therefore denied independent phonological status.¹⁸ Yet, at the same time, the \emptyset -candidates postulated by Sympathy Theory have no standing at all in morphological or syntactic terms; this seems a throwback to the representational profligacy and adhocery of SPE (see §3.4.2.1).

In interleaved OT there is, in sum, a satisfyingly close match between the array of representations postulated by the phonology and those required by the morphosyntax. In

¹⁷ As an example, consider the noun stem [_Ndes.'den] ‘disdain’ in Spanish velarizing dialects (see §3.3.2 for full details). It is not an underlying form, for it is derived by assigning the root /desdeɾ/ to the lexical category N (more specifically, to the class of athematic nouns); cf. the verb *desdeñar* and the adjective *desdeñoso*, which retain the underlying palatal. Yet [des.'den] cannot be an output form either, for Spanish velarizing dialects prohibit alveolar nasals in word-final position; cf. the inflected forms [des.'den] (sg.) and [des.'de.nes] (pl.).

¹⁸ Benua's (1997) version of TCT attempts to smuggle in derived bound stems by allowing inflected forms to be designated as ‘bases’ in OO-correspondence relationships: morphological constraints or positional identity (e.g. STEM-IDENT_{OO}; see §2.3.4.2) are then invoked to ensure that only the stem portion of the ‘base’ form exerts a paradigmatic influence. In practice, however, there is no guarantee that any surface form will faithfully reflect the crucial properties of the stem; see §3.4.1.3.

contrast, the set of phonological representations posited by strongly parallel OT is, on the whole, morphosyntactically arbitrary: strongly parallel OT fails to countenance derived bound stems, but incorporates \otimes -candidates, which are neither input nor output forms, nor lexical entries.¹⁹ This mismatch is significant because, as Orgun (1996b) observes, any serial phonological model can be reinterpreted nonderivationally: assuming that the model achieves descriptive adequacy, its theoretical success will be assessed by the extent to which it avoids *ad hoc* phonological representations. It goes without saying that, in this respect, optimality-theoretic implementations of interleaving (such as LPM-OT and SBPM) outperform rule-based models (such as traditional LPM), in that they avoid intermediate representations within level or cycles.

Orgun's declarative implementation of interleaving in SBPM relies on unification-based approaches to grammar (e.g. Pollard & Sag 1994). In essence, the morphosyntactic structure of a complex linguistic expression is represented as a tree, each of whose nodes bears a full set of syntactic and semantic specifications. Crucially, the set of specifications associated with a node *N* is a function \mathcal{F} of the specifications associated with the immediate daughters of *N*; a tree is well-formed if, and only if, for every *N* the specifications of *N* fulfil \mathcal{F} . Orgun's innovation consists in asserting that each node is associated with a full phonological representation, complete with metrical and autosegmental structure; by the same token, the function \mathcal{F} relating the specifications of *N* to those of its immediate daughters includes a phonological component ρ . Level segregation is implemented by associating different phonological functions with different node-types. Cyclicity, as mentioned in §3.3.1.1, is modelled by means of recursive, strictly binary node-types, as opposed to *n*-ary branching node-types. The theory is declarative in the sense that the well-formedness of all nodes, including their phonological specifications, can be evaluated in parallel.

3.3.2 The place of articulation of coda nasals in Spanish

Having discussed the concept of interleaving, let us now see how opacity may arise in an interleaved grammar. This is often illustrated with examples of 'segmental schizophrenia', to use Orgun's (1996a: §3.1) colourful term: i.e. situations where a segment is subject to

¹⁹ Pace McCarthy (1998: 11), \otimes -candidates cannot be legitimized by pointing out that they are independently required as members of the set of representations furnished by GEN; see §3.4.2.1 for discussion.

constraints on a syllabic position which it does not occupy on the surface. Schizophrenia routinely arises when postlexical resyllabification into the onset counterbleeds word-level neutralization processes affecting word-final codas; see Booij (1997: §3.1) for a case in Dutch. In certain dialects of Spanish, however, resyllabification creates even more spectacular counterbleeding effects which cascade across levels: resyllabification before word-level vowel-initial suffixes obscures constraints on coda nasals holding at the stem level, whilst postlexical resyllabification before vowel-initial words obscures constraints on coda nasals holding at the word level. For the data-set and certain aspects of my analysis, I rely on J. W. Harris (1983, 1984).

First of all, the allophony of rhotics supports the assumption that Spanish word-final consonants are in the coda at the word-level and that, postlexically, they are resyllabified as onsets when immediately followed by a vowel-initial word (Kenstowicz 1994: 281). In the lexicon, the opposition between /r/ and /r̄/ is neutralized in the coda. If word-final consonants *are* in the coda, it follows that there can be no contrast between the tap and the trill word-finally, as is indeed the case. Postlexically, however, a coda rhotic may optionally be trilled in rather deliberate speech. Crucially, trilling is disallowed before vowel-initial words, indicating that word-final rhotics are attracted into the onset (J. W. Harris 1983: 70-71):

(3,10)

<i>humor</i>	[u.'mor]≈[u.'mor]	'humour'
<i>humor fino</i>	[u.'mor.'fi.no]≈[u.'mor.'fi.no]	'subtle humour'
<i>humor agudo</i>	[u.'mo.ra.'ɣu.ðo], not *[u.'mo.ra.'ɣu.ðo]	'clever humour'

(For the view that word-final consonants occupy onset positions licensed by empty nuclei, see e.g. Kaye 1990 or Harris 1994.)

Let us now turn to the behaviour of nasal stops in the coda. In Spanish, coda nasals are unable to license their own place features, and so must be homorganic with a following onset consonant. Domain-finally, however, assimilation is impossible, as there is no following onset consonant; in such environments, coda nasals are assigned place specifications by default.

At the stem-level (which I assume to be non-cyclic; see below), the default specification for coda nasals is alveolar. Take, for example, the behaviour of underlying /ɲ/

in the various lexical items derived from the root /desdeɲ/²⁰ 'disdain': e.g. *desdén* (noun) 'disdain', *desdeñar* (verb) 'disdain', *desdeñoso* (adjective) 'disdainful'. In the case of the noun *desdén*, the stem is formed by assigning the root to the athematic noun class: i.e. [N[desdeɲ]]. The underlying palatal nasal is accordingly syllabified in a domain-final coda position, where, unable to license its underlying place specifications, it becomes alveolar by default: i.e. [N[desdeɲ]] → [Ndes.'den]. In the case of the verb *desdeñar*, in contrast, the stem is formed by suffixing the thematic vowel /-a-/ (the phonological marker of the first conjugation) to the root: i.e. [V[desdeɲ]a]. Underlying /ɲ/ is therefore syllabified in onset position, where its place features are licensed: e.g. [V[V[desdeɲ]a]o] → [Vdes.'de.ɲo] 'disdain' 1sg.pres.ind.²¹

There are some exceptions to the claim that, at the stem level, coda nasals either assimilate to the place of a following onset consonant or else become alveolar by default. In a number of idiosyncratic lexemes, domain-final coda /m/ retains its underlying labial specification: e.g. *áibu*[m], *Sia*[m]; cf. *Adán* [Nadam] → [Na.'dan] 'Adam', but *adámico*

²⁰ A brief comment on the underlying representation of /d/, although not crucial to the matter at hand, is perhaps in order, as it is of interest in connection with Lexicon Optimization. In Spanish, as in Catalan, /d/ and /t/ are non-contrastively realized as dental. Since Lexicon Optimization demands that predictable non-alternating structure should be fully specified in the lexicon (see table (2,15) above), the correct URs would appear to be /d/ and /t/. In Catalan, however, phonological alternations show that all coronal anterior stops (i.e. /d t n/) are alveolar at the stem level, and hence, by Lexicon Optimization, in UR. When final /t/ in *vint* 'twenty' is deleted at the word level, the preceding homorganic nasal emerges as coronal, rather than dental: i.e. [bin], not *[bin]; cf. *venc* /beNk/ → [ben] 'I sell' (see Kiparsky 1985: §1.5; Goldsmith 1990: 223ff.). The dentalization of Catalan /d/ and /t/ must accordingly be late, either word-level or postlexical. (Note that, as argued by Borowsky 1993, the word level is not bound by Structure Preservation; in LPM-OT, this follows as a corollary from the claim that underlying inventories emerge from stem-level rankings through Lexicon Optimization—see §3.3.1.1.) To the best of my knowledge, there is no positive empirical evidence requiring dentalization to be late in Spanish; but I leave open the possibility that late dentalization may be necessary on other grounds.

Another question is the alternation between the plosive and approximant allophones of /d/: [d] occurs postpausally, and after nasals or laterals; [ð] (or, more accurately, [ʝ]) occurs elsewhere. We have robust evidence that the alternation is postlexical, but there is no agreement on whether it should be treated synchronically as a fortition or a lenition. Conservatively, I assume lenition, but see Baković (1995) for an optimality-theoretic defence of the opposite view.

²¹ In Spanish, inflectional verb endings are introduced at the stem level. This we know because they trigger mid-vowel diphthongization alternations such as *cuenta* ['kʏen.ta] 'count' 3sg.pres.ind vs *contamos* [kon.'ta.mos] 'count' 1pl.pres.ind.; cf. the word-level diminutive suffix in *miel* ['mijel] 'honey' vs *mielecita* [mje.le.'θi.ta] 'honey' dim. (the literature on this alternation is copious; see e.g. Cole 1995: §6.2, §6.7 and references therein). Thematic vowels delete before vowel-initial inflectional endings, as in [V[V[desdeɲ]a]o] → [Vdes.'de.ɲo]. In such cases, the presence of the thematic vowel is not strictly necessary to prevent depalatalization of underlying /ɲ/ if the stem level is non-cyclic. I nevertheless include the thematic vowel here to facilitate comparison with other analyses below.

[A[A[Nadam]ik]o] → [Aa.'da.mi.ko] 'pertaining to Adam or the Adamites'. There is moreover a set of systematic exceptions arising from the failure of /m/ to assimilate to a following /n/: e.g. *colu*[m.n]na 'column', *hi*[m.n]o 'hymn', *sole*[m.n]e 'solemn'. I shall deal with these below.

For my present purposes, the crucial fact is that resyllabification of stem-final nasals before word-level suffixes, such as the plural noun inflection *-(e)s*, counterbleeds stem-level default alveolarization: e.g. *desdenes* [N[Ndes.'den]s] → [Ndes.'de.nes] 'disdain' pl. noun.

At the word level, nasals other than exceptional /m/ continue to be subject to the coda condition, and so assimilate to a following onset consonant, as in the following adjectives containing the negative prefix *in-*: *imposible* [Ain[po.'si.ble]] → [Aim.po.'si.ble] 'impossible', *ingrato* [Ain[gra.to]] → [Ain.ɲ.'gra.to] 'ungrateful, not gratifying'; cf. *ineficaz* [Ain[Ae.fi.'kaθ]] → [Aine.fi.'kaθ] 'inefficient'.²² However, in numerous varieties of Spanish, including several Northern Peninsular accents, the default place feature for coda nasals at the word level is not alveolar, but rather velar (for the dialectology of velarization, see García de Diego 1959 and Zamora & Guitart 1982). This can be observed in the singular form of the noun *desdén*, which is derived from the stem without the addition of an affix. Stem-final [n], now appearing in word-final position, becomes velar: e.g. sg. [N[Ndes.'den]] → [Ndes.'den]; cf. pl. [N[Ndes.'den]s] → [Ndes.'de.nes]. Again, postlexical resyllabification counterbleeds word-level default velarization: e.g. *desdén absoluto* [N[Ndes.'den][Ab.so.'lu.to]] → [Ndes.'ðe.ɲaβ.so.'lu.ʝo] 'absolute disdain'.

In the velarizing dialects of Spanish, therefore, resyllabification into the onset counterbleeds the assignment of default place features to coda nasals not once, but twice: word-level resyllabification counterbleeds stem-level default alveolarization, and postlexical resyllabification counterbleeds word-level default velarization.

²² One knows that *in-* is a word-level prefix because, when added to /n/-initial stems, it creates novel clusters which are banned tautomorphemically; see (3,18) below.

(3,11) Cascading counterbleeding in Spanish velarizing dialects

	<i>desdeño</i>	<i>desdenes</i>	<i>desdén</i>	<i>desdén absoluto</i>
UR	/desdeɲ/	/desdeɲ/	/desdeɲ/	/desdeɲ/
<i>Stem level</i>				
Morphology	[_V [_V {desdeɲ}a]o]	[_N {desdeɲ}]	[_N {desdeɲ}]	[_N {desdeɲ}]
Phonology	[_V des.'de.ɲo]	[_N des.'den]	[_N des.'den]	[_N des.'den]
<i>Word level</i>				
Morphology	—	[_N [_N des.'den]s]	[_N [_N desden]]	[_N [_N desden]]
Phonology	[_V des.'de.ɲo]	[_N des.'de.nes]	[_N des.'deɲ]	[_N des.'deɲ]
<i>Postlexical</i>				
Syntax	—	—	—	[_{NP} [_N des.'deɲ][_A ab.so.'lu.to]]
Phonology	[_V des.'ðe.ɲo]	[_N des.'ðe.nes]	[_N des.'ðeɲ]	[_{NP} des.'ðe.ɲaβ.so.'lu.ʔo]

Developing Orgun's amusing metaphor, underlying /ɲ/ could be said to suffer from multiple schizophrenia in these examples: it has a palatal surface correspondent in *desdeño*, an alveolar surface correspondent in *desdenes*, and a velar surface correspondent in *desdén absoluto*, even though in all three cases the output correspondent is in a non-neutralizing, place-licensing onset position.

Analysts with a skeptical attitude towards opacity might object that the alveolar nasal of the noun stem [des.'den] is not synchronically derived from an underlying /ɲ/; they could suggest, following J. Harris (1994: §1.4.3) and Kaye (1995: §2.2), that the morphological structure of [des.'den] is phonologically 'non-analytic' (see §3.3.1.1). In this view, the relationship of *desdén* with the verb *desdeñar* and the adjective *desdeñoso* would not involve derivation from a common root, but rather lexical look-up (Kaye 1995: 311-312). In fairness to this argument, one must observe that, even if one counts laterals together with nasals (see footnote 24), stem-level coda depalatalization triggers at best only a handful of alternations: alongside *desdén* ~ *desdeñar*, one finds *Don* 'Mr' ~ *Doña* 'Mrs', *e*[l] 'he' ~ *e*[ʎa] 'she', *donce*[l] 'lad' ~ *donce*[ʎa], and very few more. Yet, as J. W. Harris (1984: 71-72) has insightfully observed, the relative rarity of such alternations does not detract from the fact that no nasal consonant other than /n/ is possible stem-finally in athematic nouns and adjectives; this generalization is quite exceptionless.

(3,12) Possible and impossible athematic paradigms (J. W. Harris 1984: 71-72)

(a) Non-velarizing dialects

Possible

[pan]~[pa.nes]

'bread'

[ko.mun]~[ko.mu.nes]

'common'

Impossible

*[pan]~[pa.nes]

*[ko.mun]~[ko.mu.nes]

(b) Velarizing dialects

Possible

[paŋ]~[pa.nes]

'bread'

[ko.muŋ]~[ko.mu.nes]

'common'

Impossible

*[paŋ]~[pa.nes]

*[ko.muŋ]~[ko.mu.nes]

In other words, the fact that *desdén(es)* has /n/ rather than /m/ or /ɲ/ in stem-final position is not a lexical accident, but rather reflects a robust phonological constraint on Spanish stems; moreover, the evidence of velarizing dialects shows that this constraint is specific to stem-level domains and does not generalize to words. One's phonological theory must be able to accommodate these phenomena regardless of one's position on the synchronic relationship between *desdén* and *desdeñar* or *desdeñoso*. This is particularly true of OT, which is committed to the idea that the properties of underlying representations emerge from constraint rankings through Lexicon Optimization. So, by assuming a look-up relationship between *desdén* and *desdeñar* and stipulating that stem-final nasals must be alveolar, one merely replaces opacity with the Duplication Problem.

In a different connection, note that the behaviour of coda nasals in Spanish requires level segregation, but not cyclicity. Evidence from primary stress assignment and mid vowel diphthongization leads Cole (1995: §6.2, §6.7) to a similar conclusion. In J. W. Harris (1983) and Kiparsky (1982a, 1985), the stem level was assumed to be intrinsically cyclic (see §3.3.1.1) and subject to the Strict Cycle Condition (Kean 1974, Mascaró 1976). Under this assumption, the contrast between *desdeño* and *desdenes* proved somewhat problematic. J. W. Harris (1983: 73) and Kiparsky (1982a: 144-145, 168) made the dubious suggestion that syllabification creates derived environments for the purposes of the Strict Cycle Condition. This allowed the plural suffix in [_N[_N{desdeɲ}]s] to be introduced at the stem level, as the nasal would undergo coda depalatalization after syllabification on the inner cycle. In contrast, underlying /ɲ/ in [_V[_V{desdeɲ}a]o] was syllabified into the onset on the very first cycle, thereby escaping depalatalization (recall that that the root /desdeɲ/

is not a member of a lexical category, and therefore fails to define a cyclic domain; see footnote 11). However, Kiparsky (1985: §1.3) reformulated the Strict Cycle Condition in such a way that prosodification could no longer be assumed to create derived environments. In order that depalatalization should then be able to apply to underived *desdén*, Kiparsky moved the relevant rule to the (non-cyclic) word level. Now, if depalatalization applies across the board at the word level, we expect $[_N[_N[\text{desde}]]_s] \rightarrow *[\text{des.'de.}[_N\text{es}]]$. To avoid this undesirable result, Kiparsky (1985: 91) was forced to resort to extrinsic rule ordering within the word level. He asserted that word-level syllabification counterbleeds depalatalization: underlying /j/ is in coda position in the output of the stem level, and undergoes word-level depalatalization *before* being resyllabified into the onset.

J. W. Harris (1993: 183) adopts a similar trick in his treatment of *desdén* and *desdén absoluto*. For reasons which are not relevant in the present context, Harris assumes that velarization applies postlexically (more specifically, in the *P2* module; see §3.3.3.1). In order to avoid ungrammatical $*[\text{des.'ðe.na}\beta.\text{so.'lu.}\text{to}]$ alongside $[\text{des.'ðe}]\text{ng}$, however, the velarization rule must precede postlexical resyllabification. (Kenstowicz 1994: 223-224 espouses this analysis in order to safeguard Structure Preservation at the word level; but see footnote 33 below). In this connection, it is important to emphasize that such manoeuvres are strictly impossible in interleaved OT. Since interleaved OT espouses Strict Surface Orientation, input-output mappings effected by a single pass through GEN and EVAL (e.g. in a noncyclic level) are entirely transparent; see §3.2.1 and §3.3.3.1.

Let us now find out what constraint rankings obtain at each level, beginning with the stem level. I assume, first of all, that coda nasals fail to license their own place features because the constraint SONCODACOND outranks IDENT^{Place}.

(3,13) SONCODACOND

Coda sonorants do not license place features.²³

²³ One wonders whether or not SONCODACOND should be reduced to the local conjunction of simpler constraints (see §2.3.4.1). SONCODACOND could for example be analysed as CODACOND \wedge *SONC, where CODACOND states that codas in general do not license place features, and *SONC forbids sonorant consonants; the question here (to be decided on the basis of factorial typology) would be whether or not *SONC exists. CODACOND, however, cannot be decomposed into NOCODA and *PLACE, for the constraint does not penalize consonants which are in the coda *and* bear place specifications; coda consonants whose place features are licensed by a following onset segment are just fine. Rather, CODACOND forbids place nodes which are not dominated by some non-coda root node. In this sense, SONCODACOND cannot be decomposed in the same way as the constraint responsible for coda devoicing in German, which Féry (1998), following Itô and Mester, equates with NOCODA \wedge *VOICE/OBS.

Any coda sonorant whose place specifications are not licensed (in the sense of Itô 1986) by a neighbouring consonant violates (3,13). The ranking SONCODACOND \gg IDENT^{Place} thus enforces place assimilation for all coda sonorants—including, of course, nasals.²⁴ When there is no place licenser available in the vicinity of a coda nasal, e.g. word-finally, the nasal is nonetheless not allowed to remain placeless (unlike the *anusvara* of Pāli; see Zec 1995a), as the interpretability constraint N \rightarrow PLACE dominates SONCODACOND. For the notion of interpretability, see Pulleyblank (1997: 91).

(3,14) N \rightarrow PLACE

A nasal consonant must have oral place specifications.

Under pressure from N \rightarrow PLACE, unassimilated coda nasals will therefore receive a default place feature.

How are default specifications chosen? According to the current consensus, the selection of defaults is controlled by the constraint hierarchy: given a hierarchization \mathcal{H} of the set of relevant markedness constraints, the default specification of a phonological feature in a given environment is simply the least marked, or optimal, relative to \mathcal{H} (McCarthy & Prince 1994; cf. the postphonological filling-in of epenthetic material in Prince & Smolensky 1993). In Spanish, we must accordingly assume that, at the stem level, *N/ALV is the lowest ranked of all constraints on place of articulation for nasals: i.e. $\{*\text{N/VEL}, *\text{N/PAL}, *\text{N/LAB}\} \gg *\text{N/ALV}$. Under this ranking, unassimilated nasals become alveolar by default. Observe that IDENT^{Place} must be placed at the very bottom of the hierarchy, so that unassimilated nasals will not satisfy superordinate N \rightarrow PLACE by simply surfacing with their input place specifications: i.e. $\{*\text{N/VEL}, *\text{N/PAL}, *\text{N/LAB}\} \gg *\text{N/ALV} \gg \text{IDENT}^{\text{Place}}$.

²⁴ The Spanish coda licenses place constraints for obstruents: e.g. *acto* ['ak.to] 'act', *apto* ['ap.to] 'apt', *molibdeno* [mo.liβ.'ðe.no] 'molybdenum', *magdalena* [may.ða.'le.na] 'small sponge cake'; *apnea* [ap.'ne.a] 'apnoea', *étnico* ['e.θi.ni.ko] 'ethnic', *acné* [ak.'ne]; *abnegación* [aβ.ne.ɣa.'θi.ɔn] 'abnegation', *ignorancia* [iγ.no.'raŋ.θja] 'ignorance'; *lapso* ['lap.so] 'lapse', *taxi* ['tak.si]; *capaz* [ka.'paθ] 'able', *jamás* [xa.'mas] 'never', *reloj* [re.'lox] 'clock, watch'; etc.

In the case of sonorants, place specifications are always fully predictable in the coda. However, whereas nasals can assimilate to any place of articulation, laterals assimilate only to coronals: e.g. *alza* ['a.θa] 'ascent', *cadalso* [ka.'θal.so] 'scaffold'; cf. *algo* *['a.l.go] 'something' (on velar laterals, see e.g. Ladefoged & Maddieson 1996: 190-191). I assume that *LAT/LAB and *LAT/DOR dominate SONCODACOND. Unassimilated laterals become alveolar by default. Stem-finally, this triggers depalatalization: e.g. *doncel* [_N[doŋθe.ʎ]] \rightarrow [_Ndoŋ.θe] 'lad'; cf. *doncella* [_N[doŋθe.ʎa]] \rightarrow [_Ndoŋ.θe.ʎa] 'lass'. The rhotics /r/ and /r/ are predictably alveolar everywhere.

Despite the low ranking of $\text{IDENT}^{\text{Place}}$, Spanish onsets support a three-way place contrast for nasal stops: /m/~/n/~/ɲ/, e.g. *cama* /kama/ ‘bed’, *cana* /kana/ ‘white hair’, *caña* /kaɲa/ ‘reed, rod’. I assume that the place features of nasals are protected in the onset by the high-ranking positional faithfulness constraint $\text{ONS-IDENT}^{\text{Place}}$ (3,15b); for the notion of positional faithfulness (a type of relativized faithfulness), see §2.3.4.2. More specifically, $\text{ONS-IDENT}^{\text{Place}}$ must dominate $*\text{N/PAL}$, $*\text{N/LAB}$, and $*\text{N/ALV}$. In line with the idea that contrastivity in underlying representations is an epiphenomenon of stem-level constraint rankings (see §3.3.1.1), the fact that [ɲ] is not a phoneme of Spanish may be captured by ranking $*\text{N/VEL}$ above $\text{ONS-IDENT}^{\text{Place}}$. This implies that, at the stem level, velar nasals can occur only in the coda through assimilation to a following velar onset consonant. By Lexicon Optimization (see §2.2.2), the predictable (allophonic) distribution of [ɲ] becomes a property of underlying representations.

From all the above, we obtain the following partial constraint hierarchy:

(3,15)

(a) **Stem-level hierarchy: all dialects (initial version)**

$\text{N} \rightarrow \text{PLACE} \gg \text{SONCODA COND} \gg * \text{N/VEL} \gg \text{ONS-IDENT}^{\text{Place}} \gg \{ * \text{N/PAL}, * \text{N/LAB} \} \gg * \text{N/ALV} \gg \text{IDENT}^{\text{Place}}$

(b) **$\text{ONS-IDENT}^{\text{Place}}$**

Let α be a segment in the input.

Let β be a correspondent of α in the output.

If β is in the onset, then the place specifications of α and β are identical.

Under the present analysis, the fact that [ɲ] is not the default realization of coda nasals at the stem level, not even in velarizing dialects, is not merely *described*, but is actually *explained* by ranking logic. As argued above, the inventory of nasal phonemes of Spanish forces one to establish the stem-level ranking $* \text{N/VEL} \gg \text{ONS-IDENT}^{\text{Place}} \gg \{ * \text{N/PAL}, * \text{N/LAB}, * \text{N/ALV} \}$. From this ranking, it automatically follows that velar cannot be the least marked place of articulation for nasal consonants at the stem level. Hence, if default selection is indeed controlled by the ranking of markedness constraints, velar place will not be assigned to coda nasals by default. This is an extremely valuable result from a diachronic viewpoint, for it accounts for the fact that default velarization has not percolated

upward from the word to the stem level in any dialect of Spanish (see §3.3.3.2): as we now realize, the child has independent access to robust evidence which prevents him from projecting the word-level specifications of unassimilated coda consonants onto the stem level.

On the following page, tableau (3,16) illustrates hierarchy (3,15a) with the stem-level mapping of the nouns *banco* ‘bank, bench’ and *desdén*.²⁵

²⁵ In tableau (3,16), the UR of *banco* has a placeless nasal for purely expository reasons. Since the velar realization of the nasal is non-alternating, Lexicon Optimization demands the UR /baŋko/.

(3,16)

Stem Level	N→PLACE	SONCODA COND	*N/VEL	ONS-IDENT ^{Place}	*N/PAL	*N/LAB	*N/ALV	IDENT ^{Place}
baNko	*							
		*			*			*
		*				*		*
		*					*	*
				*				*
desdeŋ	*							*
								*
								*
								*

As we descend from the stem to the word level, we must note the following changes in the constraint hierarchy of Spanish. First, IDENT^{Labial} is promoted over SONCODA COND. This reranking not only protects the small handful of idiosyncratic lexical exceptions to stem-level default alveolarization like *álbum* and *Siam*, but also prevents the consistently unassimilated /mn/ cluster from merging with the novel [nn]–[ɲn] cluster which appears in the word module; see (3,18). Secondly, and most importantly, velarizing dialects reshuffle the markedness constraints on place features for nasals: *N/VEL is demoted below *N/ALV, with the effect that velar becomes the default place of articulation for coda nasals. The reader is invited to check that, under the word-level constraint hierarchy given in (3,17), stem-level [_Ndes.'den] maps onto word-level [_Ndes.'deŋ] if no affixes are added.

(3,17) **Word-level hierarchy: velarizing dialects (initial version)**

{N→PLACE, IDENT^{Labial}, ONS-IDENT^{Place}} » SONCODA COND » {*N/PAL, *N/LAB, *N/ALV} » *N/VEL » IDENT^{Place}

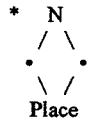
At this point, one must address the problem of the systematic absence of stem-level assimilation in /mn/ clusters. (As I have just suggested, this absence of assimilation is preserved at the word level through the promotion of IDENT^{Labial}.) Unexpectedly, a vital clue lies in the behaviour of the word-level prefixes *in-* /in-/, *en-* /en-/, and *con-* /kon-/ before stems beginning with a nasal consonant. As J. W. Harris (1984: 68) observes, the final consonant of these prefixes fails to assimilate; instead, it acquires the appropriate word-level default feature for place of articulation, i.e. alveolar in non-velarizing and velar in velarizing dialects:

(3,18)

	<i>Non-velarizing</i>	<i>Velarizing</i>	
[_Λ in[_Λ mo.'ral]] →	[_Λ in.mo.'ral]	[_Λ iŋ.mo.'ral]	'immoral'
[_Λ in[_Λ 'no.ble]] →	[_Λ in.'no.ble]	[_Λ iŋ.'no.ble]	'ignoble'

I impute the failure of assimilation to the following superordinate constraint:

(3,19) *DOUBLESARE



Two root-nodes may not simultaneously share a nasal autosegment and a place node.

Although apparently flying in the face of Steriade's (1982) Shared Features Convention, (see also Clements 1985a: 240), *DOUBLESARE may be regarded as germane to the principle, familiar from feature geometry studies, that phonological processes can spread no more than one feature node (Clements & Hume 1995: 250).²⁶ Let us now assume that OCP^{Nasal} is superordinate, compelling both members of an $N_c][L_cN$ cluster to be linked to the same nasal autosegment. If *DOUBLESARE dominates SONCODACOND, then the coda nasal will be prevented from also sharing the place features licensed by the following onset. Superordinate $N \rightarrow PLACE$ will accordingly force the coda nasal to support its own place specifications, which, as we know, will be the least marked under prevailing rankings. If we assume, moreover, that OCP^{Place} is relegated to the bottom of the hierarchy, then the default specification will be chosen without regard for the place features of adjacent segments. The behaviour of the negative prefix, in sum, implies the following word-level rankings:

(3,20) Word-level hierarchy (final version)

(a) *Non-velarizing dialects*

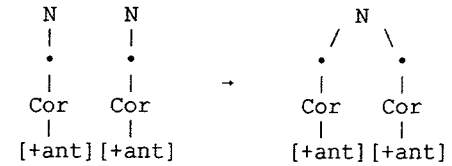
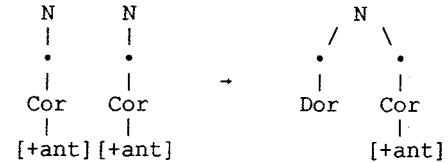
{ $N \rightarrow PLACE$, $IDENT^{Labial}$, $ONS-IDENT^{Place}$, OCP^{Nasal} , *DOUBLESARE} »
 SONCODACOND » {*/N/PAL, */N/LAB, */N/VEL} » */N/ALV »
 { $IDENT^{Place}$, OCP^{Place} }

(b) *Velarizing dialects* — cf. (3,17)

{ $N \rightarrow PLACE$, $IDENT^{Labial}$, $ONS-IDENT^{Place}$, OCP^{Nasal} , *DOUBLESARE} »
 SONCODACOND » {*/N/PAL, */N/LAB, */N/ALV} » */N/VEL »
 { $IDENT^{Place}$, OCP^{Place} }

²⁶ The statement in (3,19) is only a first approximation, but will suffice for my present purposes. A more comprehensive treatment of the Spanish data would take into consideration the behaviour of obstruent sequences and would address the rôle of voice. The gaps in the set of clusters exemplified in footnote 24 may be taken as roughly representative.

Under these rankings, an /n-n/ input to the word-level phonology maps as follows:

(3,21) (a) *Non-velarizing dialect*(b) *Velarizing dialect*

The situation at the stem level is by and large the same: nasal + nasal clusters do not constitute place assimilation environments. There is, however, one crucial difference. We know that the stem-level default place feature for coda nasals is alveolar. However, not only do underlying /mn/ clusters fail to undergo default alveolarization, but /nn/ does not occur at all; with the exception of just one word, Spanish lacks tautomorphic /nn/ clusters.²⁷ This situation can be easily described if, at the stem level, SONCODACOND is dominated not only by *DOUBLESARE and OCP^{Nasal} , but also by OCP^{Place} . With superordinate OCP^{Nasal} , the members of an underlying /Nn/ sequence are forced to share the same nasal autosegment; *DOUBLESARE, in turn, prevents them from simultaneously sharing the place node. The coda will again have to support its own place features, in violation of subordinate SONCODACOND; but, crucially, the alveolar default will simply not do in this case, for alveolar specifications on adjacent segments would violate superordinate OCP^{Place} . Accordingly, a coda nasal immediately followed by /n/ at the stem level can neither assimilate, nor be specified as alveolar by default. The only way out is for the coda

²⁷ As J. W. Harris (1984: 70) observes, *perenne* 'perennial' is the only word with tautomorphic /nn/ in general use in Spanish; three other words (*pinnado*, *pinnipedo* and *estannifero*) are rare scientific terms. *Perenne* is pronounced [pe.'ren.ne] in velarizing and [pe.'ren.ne] in non-velarizing dialects. The velarized pronunciation shows this one instance of underlying /nn/ to behave like a cluster, not a geminate. This is not surprising, for Spanish has no length contrasts; we can assume that, at the stem level, there is a superordinate constraint against root-nodes with two prosodic attachments.

to adopt the next best default specification, which turns out to be labial; in other words, [mn] is the stem-level default realization for any underlying /Nn/ sequence (J. W. Harris 1984: 71, 80).²⁸

The fact that labial is the most harmonic stem-level default specification for nasals after alveolar can be predicted without stipulation. To establish this point, I shall proceed by elimination. The second best default could never be velar: in order to capture the fact that [ŋ] is not a phoneme of Spanish, it is necessary to adopt the stem-level ranking *N/VEL » ONS-IDENT^{Place} » {*N/PAL, *N/LAB, *N/ALV}, but this renders [ŋn] less harmonic than [mn] at the stem level. By the same token, UG can be assumed to supply the stem-level ranking *N/PAL » N/LAB for free, for the presence of the phoneme /n/ in a language universally implies that of less marked /m/ and /n/ (Maddieson 1984: 69). This implies, in turn, that stem-level [ŋn] is less harmonic than stem-level [mn].

The stem-level hierarchy of Spanish should, in sum, incorporate the following refinements; cf. (3,15a):

(3,22) **Stem-level hierarchy: all dialects (final version)**

{N→PLACE, OCP^{Nasal}, OCP^{Place}, *DOUBLESHARE} » SONCODACOND »
*N/VEL » ONS-IDENT^{Place} » *N/PAL » *N/LAB » *N/ALV » IDENT^{Place}

3.3.3 Restrictions on opacity in interleaved Optimality Theory

3.3.3.1 Internal restrictions: the architecture of the model

Whilst displays such as (3,11) are familiar from traditional LPM literature, it must be stressed that interleaved OT places much tighter restrictions upon opacity than rule-based LPM ever did. In the latter, misapplication effects had two possible sources: interleaving and extrinsic rule ordering. Given interleaving, a generalization holding at some level or cycle could be obscured by the application of phonology at a later level or cycle; given extrinsic rule ordering, a rule \mathcal{R} could be rendered opaque by a later rule \mathcal{S} within the same level or cycle. In interleaved OT, however, Strict Surface Orientation eliminates the

²⁸ J. W. Harris (1984: §2) also observes that there are no tautomorphic N₁N₂ clusters where N₂ is non-alveolar; at present, I have no explanation for this fact.

second possibility: the application of phonology to any one domain consists of a single pass through GEN and EVAL, and will therefore be fully transparent.

To appreciate the significance of this innovation, consider a grammar G where all levels are non-cyclic (e.g. Spanish). According to rule-based LPM, opaque interactions in G may have unlimited depth and involve any number of extrinsically ordered rules. In interleaved OT, in contrast, opacity is bound by the number of levels distinguished in G : i.e. given a mapping such as UR ^{\mathcal{A}} → stem-level representation ^{\mathcal{E}} → word-level representation ^{\mathcal{C}} → postlexical representation (=SR), \mathcal{A} may be rendered opaque by \mathcal{B} or \mathcal{C} , and \mathcal{B} may only be rendered opaque by \mathcal{C} , whilst \mathcal{C} itself must perforce be transparent. Even if any of the levels in G were cyclic, opaque derivations would remain relatively shallow, as the possibility of opaque interactions between cycles is limited by the depth of morphological embedding. In interleaved OT, therefore, one can derive precise predictions about opacity from the overall architecture of grammars.

In this connection, let us look at the postlexical module in greater detail. Postlexical processes are typically non-cyclic (Kaisse & Shaw 1985: §2.3), but a case has been made for cyclicity over nested phrasal domains in some languages (see e.g. Dresher 1983, Kaisse 1985, 1987; Mohanan 1986). Following this line of inquiry, Kaisse (1985, 1987) and Mohanan (1986) suggest that the postlexical module may comprise two separate levels, which Kaisse calls 'postlexical 1' ($P1$) and 'postlexical 2' ($P2$); the former would be cyclic (and therefore sensitive to syntactic information), whilst the latter would be non-cyclic. Rather than discuss the specific merits of this proposal, I wish to focus here on the properties of those across-the-board postlexical generalizations which unequivocally belong to the $P2$ type. In interleaved OT, such generalizations must be fully transparent: misapplication cannot arise because the $P2$ level is bound by Strict Surface Orientation, it is non-cyclic, and its output undergoes no further phonological operations. This is a strong empirical prediction which rule-based LPM cannot derive without stipulation. I shall illustrate its import with two across-the-board postlexical processes in Spanish: secondary stress assignment and hiatus-removal (Roca 1986).

In Spanish, the domain of secondary stress is the intonation group (henceforth IG). Since IGs encompass word sequences of indefinite length, one can safely infer that secondary stress is assigned postlexically.²⁹ Within this domain, a grid-mark is assigned to

²⁹ Analyses relying on iterative footing for the assignment of primary stress in the lexicon must resort to a house-keeping operation of conflation to delete unwanted secondary stresses: see e.g. Halle & Vergnaud (1987b: 93-96, 98-100), J. W. Harris (1995: 875).

every second unstressed syllable to the right of a primary stress. Clash is avoided, hence the absence of stress on the syllable [mas] in (3,23a). There is also an initial dactyl effect, as shown in (3,23b).

(3,23) (a) *los problemas de los adolescentes* 'teenagers' problems'

* * * * *
* * * * * * * * * * * *
los.pro.βle.mas.ðe.lo.sa.ðo.les.θeɲ.tes

(b) *el beneplácito de las autoridades* 'the authorities' approval'

* * * * * *
* * * * * * * * * * * *
el.be.ne.pla.θi.ɰo.ðe.la.saɯ.ɰo.ri.ða.ðes

To account for this pattern, I propose that the following constraint ranking holds postlexically in the phonology of Spanish:

(3,24) { IDENT^{StressPeak}, TROCHEE, FTBIN, ALIGN(IG, L; Σ, L) } » PARSE^σ »
ALIGN(Σ, R; IG, R)

Spanish assigns primary stress lexically; IDENT^{StressPeak} ensures that primarily stressed syllables retain their prominence (see Pater 1995: 21). The feet are strictly binary trochees constructed over syllables; the ranking FTBIN » PARSE^σ rules out clash by banning degenerate feet. In turn, the ranking PARSE^σ » ALIGN(Σ, R; IG, R) triggers exhaustive right-to-left footing, whilst superordinate ALIGN(IG, L; Σ, L) takes care of the initial dactyl effect. For the formulation and effect of the alignment constraints employed here, see McCarthy & Prince (1993b: §3) and section 3.4.1.1 below.³⁰

At the same time, the postlexical phonology of Spanish complies with a further generalization whereby immediately adjacent vowels not bearing primary stress must be tautosyllabic. Sequences of adjacent atonic vowels of different quality normally satisfy this

³⁰ Subhierarchy (3,24) predicts that, when final within the intonation group, the last syllable of proparoxytonic words will not bear secondary stress: e.g. *pánico* 'panic' [pa.ni.ko]; the alternative parse [pa.ni][.ko] violates FTBIN. Some dialects, however, are reported as allowing the prominence countour *pánico*. This may not be a stress effect at all, but rather an epiphenomenon of intonational or durational boundary phenomena. For discussion, see Stockwell, Bowen & Silva-Fuenzalida (1956), Roca (1986), and J. W. Harris (1991e); for a similar problem in Icelandic, see Hayes (1995: §6.2.2).

constraint by forming a diphthong or even a triphthong (whereas identical atonic vowels coalesce). Diphthongization across word boundaries, traditionally known as 'synaloepha', is thus extremely frequent in Spanish: e.g. *Su ayuda elimina obstáculos* [sua.'zu.ðaɛ.'li.mi.naɔβs.'ta.ku.los] 'His help removes obstacles'. Interestingly, two of the diphthongs which occur in this example, [aɛ] and [aɔ], are ill-formed in the lexical phonology, where two adjacent vowels cannot form a diphthong unless one of them is unstressed and high.³¹ Postlexical hiatus-removal creates such novel diphthongs also word-internally ('synaeresis'): e.g. *coartada* [kɔa.'ta.ða] 'alibi'.

Secondary stress assignment and hiatus-removal can easily be stated as phonological rules. To obtain the right output, however, one must stipulate that hiatus-removal precedes secondary stress assignment. Under the reverse order, the alternating pattern of secondary stress would incorrectly be rendered opaque by synaloepha and/or synaeresis.

(3,25) *ánimo a los de la oposición* 'encouragement to those in opposition'

(a) **Opaque ordering**³²

input ('a.ni.)mo.a.los.ðe.la.o.po.si.(θi)on
secondary stress ('a.ni.)(mo.a.)(los.ðe.)(la.o.)(po.si.)(θi)on
hiatus-removal * ('a.ni.)(mɔa.)(los.ðe.)(laɔ.)(po.si.)(θi)on

(b) **Transparent ordering**

input ('a.ni.)mo.a.los.ðe.la.o.po.si.(θi)on
hiatus-removal ('a.ni.)mɔa.los.ðe.laɔ.po.si.(θi)on
secondary stress ✓ ('a.ni.)(mɔa.los.)(ðe.laɔ.)(po.si.)(θi)on

³¹ Alongside the intuitions of native speakers (including poets), there is evidence from stress assignment to support this statement. In Spanish, primary stress is assigned lexically, and is confined to a trisyllabic window located at the end of the word (see e.g. Roca 1991, J. W. Harris 1995). In words such as *náutico* 'nautical' and *terapéutico* 'therapeutic', the sequences *au* and *eu* constitute falling diphthongs, and so the position of the stress does not violate the trisyllabic window: i.e. ['naɯ.ti.ko], [te.ra.'peɯ.ti.ko] (Roca 1997: 248ff.). In contrast, observe how the addition of the suffix *-ico* causes the stress to shift in the adjective *oceánico* [o.θe.'a.ni.ko] 'oceanic', derived from the noun *océano* [o.'θe.a.no] 'ocean'. The alternative **océánico* violates the trisyllabic window, since, at the lexical level, *ea* is an ill-formed diphthong.

³² Although here the application of hiatus-removal renders the alternating pattern of secondary stress opaque, this is not, strictly speaking, an instance of counterbleeding order, for both rules get a chance to apply anyway; see Kisseberth (1973: 423ff.).

In post-SPE rule-based theory, opaque rule ordering has often been characterized as relatively marked, but it has never been successfully eliminated. Accordingly, rule-based LPM has no principled grounds to exclude (3,25a); it can *describe* the transparent interaction between secondary stress assignment and hiatus-removal in the postlexical phonology of Spanish, but it cannot provide an *explanation*. One might suggest that extrinsic ordering is confined to lexical levels, but this would be an *ad hoc* hedge. In interleaved OT, in contrast, the ungrammaticality of *ánimo* à *lòs de là opòsición* can be deduced from first principles. The transparent parse *ánimo* à *los de la opòsición* is optimal because subhierarchy (3,24) applies in parallel with the constraint against hiatus between atonic vowels.

The assertion that across-the-board postlexical regularities are universally transparent has undeniable empirical content. It is in respect of such predictions that the issue of the extension of opacity, discussed in §3.2.3, acquires critical significance. Suppose, for example, that there is a language with a chain-shift process of vowel raising similar to that of Nzebi in (2,47) and (3,8). Suppose, however, that this chain shift applies postlexically across the board: i.e. including the environment $V\#i$. In such a case, rule-based theory could still resort to a battery of counterbleeding rules, but interleaved OT would be compelled by its own architecture to provide a transparent (i.e. strictly surface-oriented, strongly parallelistic) constraint-based description. The adequacy of transparent analyses such as Kirchner's would then be of criterial theoretical importance. The same applies to processes of assimilation with deletion of the trigger, such as yod-coalescence (i.e. $tjV \rightarrow t[V]$) and vowel nasalization (i.e. $VN \rightarrow \tilde{V}$): if such processes are found to apply postlexically across the board, then interleaved OT predicts that the transparent one-step coalescence analysis is the correct one; see (3,9).

Opacity effects arising from cyclicity are also better constrained in interleaved OT than they are in rule-based theory. As Cole (1995: §8) points out, the combination of cyclicity with extrinsic rule ordering creates the possibility of apparent 'rule-ordering paradoxes'. Consider, for example, a battery of two rules \mathcal{A} and \mathcal{B} , where \mathcal{B} counterfeeds \mathcal{A} . If \mathcal{A} and \mathcal{B} apply cyclically, it is possible in principle for an application of \mathcal{B} to counterfeed \mathcal{A} on cycle n , only to then feed \mathcal{A} on cycle $n+1$: $[\mathcal{A} \rightarrow \mathcal{B}]_{\text{cycle } 1} \rightarrow [\mathcal{A} \rightarrow \mathcal{B}]_{\text{cycle } 2} \rightarrow \text{etc.}$ As Kiparsky (1985: 88–89) emphasized, one of the main functions of the Strict Cycle Condition in LPM was precisely to forestall such rule-ordering paradoxes; but the Strict Cycle Condition has been jettisoned on independent grounds, after proving incapable of

dealing with non-derived environment blocking effects (Kiparsky 1993). According to Cole, there is no reliable evidence that cyclic application of extrinsically ordered rules is ever required; see Booij (1997: 277ff.) for a concurring view.

3.3.3.2 External restrictions: Lexicon Optimization in acquisition and change

As the previous section has shown, interleaved OT imposes stringent limitations on the complexity of misapplication effects. Significantly, these restrictions follow as theorems from the combination of interleaving with Strict Surface Orientation; no *ad hoc* clauses need be written into the theory. However, Benua (1997: §3.5.4.2) and McCarthy (1998: 9–10) suggest that level segregation retains excessive generative power. In interleaved OT, there is no formal UG principle requiring constraint ranking to be homogeneous across the various levels of a phonological system.³³ This —it is argued— leads to incorrect typological predictions, for in attested languages constraints hierarchies at different levels do not differ from one another in a random and unlimited fashion: “the actual differences between strata in a single language are quite limited” (McCarthy 1998: 9). I shall henceforth call this the ‘cophonology divergence’ problem.

Proponents of interleaved OT acknowledge that cophonology divergence is not unrestricted in the languages of the world; indeed, they accept that ranking uniformity across levels may be ‘unmarked’ or ‘optimal’ (in a sense to be specified below). They suggest, however, that the actual similarities and differences between levels within one grammar are an epiphenomenon of processes of acquisition and change; UG imposes no specific restrictions on the reranking of constraints across levels. Orgun (1996b: 122) boldly

³³ As Orgun (1996b: §4.4.2, §4.4.3) and Benua (1997: §5.7.2) observe, in rule-based LPM there were several attempts at curbing cophonology divergence by adding *ad hoc* clauses to the theory: notably Structure Preservation (Kiparsky 1985: 92) and the Uniform Domain Hypothesis (Mohan 1982, Halle & Mohan 1985).

Structure Preservation asserted that markedness conditions on underlying representations remain in force throughout the lexical phonology. However, later research has found that only the stem level is structure-preserving (see e.g. Borowsky 1993). In interleaved OT, this fact need not be stipulated, but follows as a corollary from the assumption that the content of underlying representations is determined by stem-level rankings through Lexicon Optimization (see §3.3.1.1).

The Uniform Domain Hypothesis held that phonological rules must be active within a contiguous set of levels. Again, there is evidence that this statement is redundant as a principle of UG. Under a specific view of sound change to be discussed below, the state of affairs described by the Uniform Domain Hypothesis is an expected result of the diachronic evolution of phonological innovations.

The Uniform Domain Hypothesis had a more stringent counterpart, the Strong Domain Hypothesis (Kiparsky 1984), which asserted that, if a rule applies at level n , then it also applies at level $n-1$. The Strong Domain Hypothesis is false as a typological generalization (Orgun 1996b: §4.4.2), but, as I shall argue below, it can be seen as embodying the default assumptions of the learner in the acquisition process. In this rôle, however, the Strong Domain Hypothesis can be deduced in interleaved OT from Lexicon Optimization.

asserts, "When learnable, considerable differences between cophologies are possible, and are attested." The problem of cophology divergence is therefore to be approached in the same way as the difficulties raised by cophology proliferation and cophology arbitrariness (see §3.3.1.2).

The appeal to history in relation to the cophology divergence problem is more than a promissory note. Kiparsky (1998a: 66) makes succinct but explicit proposals concerning the course of phonological change in an interleaved optimality-theoretic system. In essence, he suggests that constraint rankings have a 'life cycle' analogous to that of phonological rules in rule-based LPM (Kiparsky 1995: 657).³⁴ Phonological innovations first arise through the promotion of markedness constraints at the postlexical level; this results in Neogrammarian sound change. Subsequently, innovative constraint rankings tend to percolate upwards to higher levels so as to restore ranking uniformity throughout the grammar; this corresponds to so-called analogical change. (For Kiparsky's 1995 treatment of lexical diffusion in rule-based LPM, and its adaptation to an optimality-theoretic framework, see chapter 4.) If phonological change follows this pattern, then the scope for ranking divergence between levels is severely limited: ranking differences between cophologies will in the main reflect the historical lag of higher levels relative to lower levels. Paraphrasing Givón (1971), Kiparsky's proposals entail that today's lexical phonology is yesterday's postlexical phonology.³⁵

The idea that phonological innovations first arise as contingent, morphologically insensitive, Neogrammarian changes has received support from research in laboratory phonology. Ohala (1989, 1992, 1993) argues that phonological innovation is fuelled by acquisition errors which arise through stochastic processes of 'hypocorrection': when a learner fails to compensate adequately for distortions in the incoming signal, a phonetic effect may be misinterpreted as a categorical pattern of the target grammar. The release of stop consonants, for example, is noisier before high than before low vowels, as air velocity increases in a narrowed channel; through hypocorrection, however, this purely acoustic effect can give rise to an innovative grammatical process of affrication —as Ohala (1989:

³⁴ The idea that phonological patterns have a life cycle has a long tradition. Pioneering proposals include those of Baudouin de Courtenay and Kruszewski regarding morphologization (Anderson 1985: ch. 3). Work within LPM has focused on the hypothesis that rules tend to ascend from lower to higher levels, becoming subject to different conditions on application in the process; see, in particular, Kiparsky (1984, 1988, 1995), J. Harris (1989), McMahon (1991), and Zec (1993).

³⁵ Moreover, the state of affairs described by the Uniform Domain Hypothesis will obtain whenever a constraint ranking percolates upwards without skipping levels (see footnote 33).

186) puts it, a 'rule of physics' is mistaken for a 'rule of grammar'.³⁶ Crucially, innovative phonological regularities arising by hypocorrection are necessarily surface-true (the rules of physics do not misapply); in an interleaved system, they will therefore involve the promotion of the appropriate markedness constraints at the postlexical level (see §3.3.3.1). Modelling the absence of morphological conditioning requires no special provision, for word-internal domains are unavailable at the postlexical level. More importantly, new postlexical rankings may interact opaquely with rankings holding at higher levels; this accounts for the recurrence of locality effects in language history, where older sound changes are routinely obscured by younger ones. In strongly parallel OT, in contrast, all constraints interact globally within a single hierarchy; thus, the promotion of a markedness constraint must often be accompanied by a wholesale alteration of other ranking relations in order to avoid undesirable global effects (see Idsardi 1997: 389, Kiparsky 1998a: 66). Interleaved OT predicts global effects only within the level where reranking takes place (see chapter 4).

According to Kiparsky (1998a: 66), innovative postlexical rankings tend to percolate to higher grammatical levels. In this process, the corresponding phonological generalizations become morphologically sensitive, as word-internal domains become available within the lexical phonology; this produces the appearance of 'analogical change'. Kiparsky asserts that, optimally, constraints rankings are uniform across levels; thus, the upward percolation of innovative constraint rankings results in the simplification or optimization of the grammar. Although Kiparsky (1998a) does not elaborate on the theoretical status of this assertion, I shall contend that it can be deduced as a theorem from the principle of Lexicon Optimization.

Recall that, according to Lexicon Optimization, learners project input representations algorithmically from the corresponding output forms. In particular, for any morpheme *M*, the learner selects that unique input representation *I* which produces the correct set *S* of output alternants with the least serious violations of the constraint hierarchy. In the case of a non-alternating morpheme, the input representation will be identical with the output in order to avoid gratuitous violations of faithfulness constraints (see §2.2.2.1). In an

³⁶ Note that the theory of hypocorrection does not purport to explain the sociolinguistic process of propagation whereby an innovative phonological pattern becomes part of the community norm (for the distinction between innovation and propagation, see e.g. Milroy & Milroy 1985, Milroy 1992). Nonetheless, different types of hypocorrection process have different probabilities of occurrence as a result of purely physical factors. This can contribute to explaining the rate of diffusion of the corresponding grammatical variants across a population (see Nigoyi & Berwick 1997 for groundwork in dynamical systems theory, and Pulleyblank & Turkel 1996: §5 for a worked-out phonological simulation).

interleaved system, however, input-output mappings are recursive; the learner must bootstrap the grammar from surface data whilst negotiating a hierarchy of levels, each with its own ranking of constraints. Therefore, an interleaved OT model which takes Lexicon Optimization seriously must assume that Lexicon Optimization applies recursively, optimizing input-output mappings over each level. As I shall demonstrate, this has momentous consequences for the relationship between constraints rankings across levels.

As I have just pointed out, the learner must bootstrap the entire grammar from surface forms. In an interleaved system, however, the surface representation consists of the output of the last phonological level L_n . At the outset of the acquisition process, therefore, the learner will use surface forms to determine the constraint hierarchy of L_n ,³⁷ the input to L_n will be projected from surface forms in line with Lexicon Optimization. Naturally, the learner must at this stage undo precisely those alternations which are generated at L_n , and this implies knowing or finding out what morphosyntactic constructions are ascribed to L_n in the target grammar. In LPM, L_n is the postlexical level; I assume therefore that, in the first recursion of Lexicon Optimization, the learner aims to undo phrasal alternations and to obtain invariant word-level representations for single words.³⁸ In line with Lexicon Optimization, these word-level representations will differ minimally from surface forms in order to avoid gratuitous violations of postlexical faithfulness constraints. Having determined the input to L_n , we can move on to the next stage of the learning process, for the input to L_n equals the output of L_{n-1} . The learner will consequently use the input to L_n to determine the constraint hierarchy of L_{n-1} , and will apply Lexicon Optimization to derive the input to L_{n-1} from the input to L_n (factoring out alternations associated with L_{n-1} morphological constructions; see footnote 38). The process iterates until underlying representations (i.e. input representations of unanalysable morphemes) are reached. In this sense, the mechanics of constraint hierarchization and Lexicon Optimization are largely the same in interleaved OT as in strongly parallel OT, with two significant exceptions. First, learners in interleaved OT deal with a relatively impoverished constraint set which excludes

³⁷ The details of the constraint hierarchization process are orthogonal to the argument that I am developing here. For opposing views of the initial state of CON, compare Gnanadesikan (1995) and Smolensky (1996) with Hale & Reiss (1997a, b).

³⁸ Within the lexicon, one can plausibly hypothesize that the learner is 'greedy', in the sense that he attempts to assign the maximum possible number of morphological constructions to the current level before ascending to the next higher level. This assumption is logical if, in the interests of economy, grammars are to distinguish the fewest possible levels. An interesting corollary, though, is that, in the course of history, productive morphological constructions will tend to be demoted to the word level. This may be a key to explaining the clustering of phonological and morphological properties discussed in §3.3.1.2.

OO-correspondence and sympathy constraints. Secondly, the learner of an interleaved grammar does not tackle all alternations *en masse*, but rather peels away levels of morphosyntax one by one like onion layers.

Assume now that the marked configuration m is banned from the output of L_n . In response to this fact, the learner will rank the relevant markedness constraint M over any antagonistic faithfulness constraint F in the constraint hierarchy of L_n . Guided by Lexicon Optimization, however, the learner will also make the default assumption that m is absent from the input to L_n . If this assumption proves tenable, not only superordinate M but also subordinate F will remain unviolated through the mapping $\text{input}(L_n) \rightarrow \text{output}(L_n)$. Crucially, the learner will stick to this assumption unless some phonological alternation generated at L_n forces him to posit an input with the property m . If such alternations do not arise in the learner's trigger experience, and m is therefore absent from $\text{input}(L_n) = \text{output}(L_{n-1})$, then the learner will again rank M over F in the constraint hierarchy of L_{n-1} , and the process will start all over again to discover $\text{input}(L_{n-1}) = \text{output}(L_{n-2})$. In sum, the default assumption of a learner trying to acquire an interleaved phonological system under recursive Lexicon Optimization will be that, if a constraint ranking \mathcal{K} holds at level L , then \mathcal{K} also holds at all levels higher than L ; this assumption will be maintained unless the learner's triggering experience contains positive counterevidence from phonological alternations.³⁹

Let us take a more concrete example. Imagine a language where [h] is banned from the coda in surface forms: i.e. the constraint *CODA/h outranks MAX^{Seg} in the postlexical constraint hierarchy. What will the situation be at the word level? As I explained in the previous paragraph, a learner who has successfully established the postlexical ranking *CODA/h » MAX^{Seg} will make the default assumption that coda [h] is also prohibited at the word level. Under this assumption, word-level representations will vacuously satisfy *CODA/h when submitted to the postlexical phonology, and no violations of MAX^{Seg} will have to be incurred. What sort of evidence could prompt the learner to reject this default assumption? Let us assume, for example, that in this language word-final coda consonants are resyllabified into the onset when followed by a vowel-initial word; in these circumstances, word-final [h] will escape postlexical deletion when the next word begins with a vowel. If the learner is exposed to words which alternate between [h]-final and non-

³⁹ In this sense, Kiparsky's (1984) Strong Domain Hypothesis can be seen as characterizing the default assumptions of the learner during the acquisition process. Note, however, that, in this rôle, the Strong Domain Hypothesis is a simple corollary of recursive Lexicon Optimization. See footnote 33.

[h]-final realizations depending on the initial segment of the following word (and if these alternating items occur alongside non-alternating ones), then he will infer that the presence or absence of coda [h] is a contrastive feature at the word level, at least in word-final position; by the same token, he will rank MAX^{Seg} over *CODA/h in the word-level constraint hierarchy. In this sense, the learner relies on the positive evidence of alternating [h]-final words to arrive at the correct word-level ranking. If for any reason this cue should fail to appear in the learner's trigger experience with sufficient robustness, then the default assumption will be adhered to, and the postlexical ranking *CODA/h » MAX^{Seg} will percolate upwards to the word level. (For broadly similar suggestions regarding the connection between the trigger experience, default grammatical settings, and phonological change, see e.g. Dresher 1993: 132-133.) Conversely, we have seen that nasal velarization has failed to percolate from the word to stem level in Spanish velarizing dialects because the absence of a phoneme */ŋ/ provides robust evidence for the correct stem-level rankings; see §3.3.2.

To recapitulate, I suggest, following Kiparsky (1998a: 66), that innovative constraint rankings normally arise at the postlexical level by Ohalian hypocorrection. Historically, constraint rankings tend to percolate from lower to higher levels because, in the acquisition process, the learner's initial or default assumption is that constraints holding for the output of a level L_n will also hold for the output of level L_{n-1} . This assumption is dictated by the principle of Lexicon Optimization, which, for any output representation, instructs the learner to select that input representation which triggers the least serious violations of the relevant constraint hierarchy. The assumption of ranking uniformity between L_n and L_{n-1} is only abandoned if the triggering experience contains sufficiently robust counterevidence from phonological alternations at L_n . This model of phonological acquisition and change predicts that ranking differences between levels will in the main reflect the historical lag of higher, relatively conservative levels with respect to lower, relatively advanced levels.⁴⁰ Thus, interleaved OT solves the problem of cophonology divergence without stipulations of any kind. I therefore assert, with Orgun (1996b: 122), that no formal UG restriction on

⁴⁰ In this sentence, the qualification 'in the main' acknowledges that other forms of reanalysis are possible: e.g. so-called rule inversion (Vennemann 1972). Note, however, that even rule inversion may be driven by Lexicon Optimization. Imagine that a level L_n has a very extensive process of vowel epenthesis. Under the default assumption that the input and the output of L are identical, the child may take the epenthetic vowels to be already present in the output of L_{n-1} , and posit a restricted process of syncope at L . Here, the innovative process of syncope has not percolated upward from L_{n+1} , but Lexicon Optimization was still involved in its birth.

the reranking of constraints across levels is either necessary or appropriate; I shall henceforth refer to this statement as the 'principle of free reranking'.

3.3.3.3 Synchronic markedness reversal

According to Benua (1997), however, the principle of free reranking is false. She contends that, within one grammar, the relative ranking of markedness constraints always remains invariable across levels; cophonologies differ only in respect of the ranking of faithfulness constraints. In attested languages there would be, in other words, no synchronic markedness reversal. Benua states this claim in the strongest possible terms: "Promotion of faithfulness is the key to serial OT analyses of cyclic effects—in fact it is only faithfulness, and never markedness, that changes its ranking position between levels" (Benua 1997: 90). As she observes, the absence of markedness reversal would deliver a severe blow to interleaved OT, for this generalization (unlike the observation that rankings tend to percolate from lower to higher levels) could not be deduced from first principles:

Markedness rankings are fixed across the grammar. But this is simply a stipulation, that does not follow from anything else in the subgrammar theory. If re-ranking between levels is possible at all, it is unclear why faithfulness constraints are mobile and markedness constraints are not.

Benua (1997: 90)

[T]o the extent that subgrammar theory is predicated on differences between levels of derivation, it is committed to explaining the similarities between them. In particular, it should explain why relative markedness rankings do not differ.

Benua (1997: 218)

Why is it always the case that markedness relations do not change, but faithfulness is promoted as the derivation proceeds?

Benua (1997: 225)

The importance of Benua's challenge to the principle of free reranking cannot be stressed enough: if empirically corroborated, the absence of synchronic markedness reversal would render interleaved OT massively overgenerating. All the other arguments against interleaving that we have seen so far have proved ineffectual. In §3.3.1.2, we saw Benua concede that UG imposes no formal restrictions on the number of cophonologies or the assignment of specific morphological constructions to particular cophonologies. In §3.3.1.3,

the charge of derivationalism against interleaving was shown to be ill-founded, with interleaved OT displaying a satisfyingly close match between the array of representations postulated by the phonology and those required by morphosyntax. In §3.3.3.1, I deduced formal constraints on the complexity of misapplication effects from the architecture of interleaved grammars under Strict Surface Orientation. And in §3.3.3.2, the cophonology divergence problem was solved by showing how the diachronic life cycle of constraint rankings determines the similarities and differences between constraint hierarchies across levels; in this connection, I proved that the upward percolation of innovative constraint rankings requires no special stipulation, but is driven by the recursive application of Lexicon Optimization. The putative impossibility of synchronic markedness reversal would nonetheless lay interleaved OT open to the charge that, when it comes to constrain opacity, it asks processes of grammar ontogenesis to do too much, whilst it has UG do too little.

The rejection of synchronic markedness reversal must be seen, in contrast, as a fundamental postulate of strongly parallel OT. As I pointed out in §3.2.2.1, this research programme is defined by the attempt to render misapplication phenomena compatible with the view that GEN and EVAL are non-recursive functions. Faced with this conundrum, proponents of Strong Parallelism assert that the output enters into a variety of parallel correspondence relationships, each of which is monitored by a specialized set of correspondence constraints: "Thus multiple sets of faithfulness constraints, proper to different types of linguistic relations, coexist in the grammar, interacting with one another and with the hierarchy of phonological markedness constraints" (Benua 1997: 4). Misapplication is therefore regarded as an epiphenomenon of the multiplicity of parallel transderivational correspondence relationships. In this context, the relative ranking of markedness constraints emerges as the backbone of the grammar, the one constant factor which unifies phonological mappings within one system.

This idea informs Benua's otherwise extremely liberal version of TCT. Affixes may be divided into an indefinite number of correspondence classes, where each class is monitored by its own set of independently ranked OO-identity constraints (see §3.3.1.2); additionally, each affix is allowed to stipulate the choice of surface 'base' in its subcategorization frame (see §3.4.1.3). Nonetheless, all affixes are bound together by the same hierarchy of markedness constraints, regardless of class membership. Similarly, McCarthy's (1998) outline of Sympathy Theory insists that \otimes -candidates must be selected by IO-faithfulness constraints, thereby ensuring that \otimes -candidates comply with the same ranking of markedness constraints as output forms. The analogy between Benua's TCT and

McCarthy's (1998) proposals is brought out by thinking of \otimes -candidates as 'virtual bases', i.e. virtual output forms with no real morphosyntactic status (see §3.4.2.1). Finally, Fukazawa (1998) and Fukazawa, Kitahara & Ota (1998), refining proposals made by Itô & Mester (1995), argue that the various Japanese sublexica are controlled by separate, differently ranked sets of IO-faithfulness constraints, interspersed among a common hierarchy of markedness constraints.

Nonetheless, synchronic markedness reversal does occur in natural language, nor is it an extraordinary feature of grammars. This dissertation adduces two instances: default place assignment to unassimilated coda nasals in Spanish velarizing dialects (§3.3.2), and West Germanic Gemination (to be discussed in §3.5). Recall that, in Spanish velarizing dialects, [n] is selected as the default realization of unassimilated coda nasals at the stem level because, in the constraint hierarchy of that level, *N/VEL dominates *N/ALV; see (3,22). At the word level, however, *N/ALV is promoted over *N/VEL, with the result that unassimilated coda nasals become velar; see (3,20b). Since both *N/ALV and *N/VEL are markedness constraints, the existence of synchronic markedness reversal is empirically demonstrated. By the same token, if the principle of free reranking is correct, one surmises that a theory of grammar which forbids markedness reversal must lead to *ad hoc* analyses or miss important generalizations; in §3.4.1.3, I return to my Spanish data to show that this is precisely what does happen. Furthermore, the case of West Germanic Gemination will add to a mounting fund of evidence showing that markedness constraints must be allowed to select \otimes -candidates (Itô & Mester 1997b, de Lacy 1998).

The proposition that all levels must share the same markedness hierarchy appears particularly unlikely in the light of the historical evidence and of what is known about the mechanisms of phonological innovation. Hypocorrection processes, as described by Ohala (see §3.3.3.2), are rooted in performance and are, as such, essentially stochastic. At the same time, non-global interactions are extremely frequent in the known history of languages, including countless cases where a phonological innovation causes an erstwhile surface-true generalization to be violated; this is enough to suggest that, even though innovations may well be selectively 'primed' by existing structure, priming effects can be

overridden (cf. Kiparsky 1995: §2.1).⁴¹ Given such evidence, phonological innovations requiring markedness reversal are to be expected.

Interleaved OT, in contrast, has no trouble accommodating markedness reversal. Yet, as I have shown, the principle of free reranking does not imply that anything goes. Interstratal divergence is not arbitrary, but is always compatible with, and constrained by, a lawful ontogenetic scenario. As Anderson (1992: 336) puts it,

[... A]n adequate theory of what structures are possible in natural languages probably cannot be provided on a purely synchronic basis. Rather, we must view "what there is" as the product of "what is possible" with the range of possible diachronic developments.

History does matter.

3.4 STRONGLY PARALLEL OPTIMALITY THEORY

A strong claim of section 3.3 is that interleaved OT requires no theory of misapplication; its model of the phonology-morphosyntax interface solves the problem of opacity without further stipulation. Moreover, its synchronic account of misapplication is closely integrated with a long-established theory of phonological change, now strengthened, as I was able to show in §3.3.3.2, by the independently motivated principle of Lexicon Optimization. In this section I will endeavour to show that strongly parallel OT cannot lay claim to a comparable degree of generality. Strongly parallel approaches to the rôle of morphosyntax in phonology, notably TCT, have failed to provide comprehensive coverage of opaque phonological phenomena; special theoretical provision for opacity, in the shape of Sympathy Theory, has eventually proved indispensable. As Strong Parallelism struggles thus to cope with misapplication, its prospects of integrating the synchronic and diachronic aspects of the problem appear dim.

In the area of paradigm effects (see §3.2.2.1), the bulk of strongly parallel research has focused on OO-correspondence. As the undesirable implications of this concept have become apparent, proponents of TCT have to an increasing extent withdrawn their most

⁴¹ To exemplify language-specific top-down constraints on phonological innovation, Kiparsky (1995: 656) chooses the case of compensatory lengthening: de Chene & Anderson (1979) claim that consonant deletion can only trigger compensatory vowel lengthening in languages where a vowel length contrast already exists. The fortunes of this claim are discussed in chapter 4.

important empirical claims and fallen back on stipulation. I shall chart this retreat in §3.4.1, following lines of argument developed by Orgun (1996a, b) and Kiparsky (1998a):

(i) OO-correspondence models itself on base-reduplicant identity (henceforth, BR-identity; see McCarthy & Prince 1995). However, BR-identity is a symmetrical relationship, whereas OO-correspondence is demonstrably asymmetrical: i.e. phonological generalizations may misapply in a derived form to achieve identity with their output 'base', but not vice versa. To enforce the priority of the base, TCT must resort to costly stipulations: Benua's (1997) solution, for example, involves a simulacrum of recursive phonological mapping within the constraint hierarchy. This result refutes the claim that the mechanisms of Correspondence Theory are fully general and provide a unified account of misapplication.

(ii) In an interleaved régime, a morphological operation \mathcal{M} can render opaque only those phonological properties which are present in its input (Kiparsky 1998a: 26). If bases are understood as surface forms, however, OO-identity constraints will meet two interrelated obstacles: first, they will have access to irrelevant phonological information (properties introduced after the application of \mathcal{M} , e.g. at the postlexical level); secondly, the crucial properties of the input to \mathcal{M} may not surface transparently anywhere at all. In response to these masking effects, proponents of TCT have adopted increasingly looser and more stipulative definitions of 'base'.

It will thus be shown that TCT can only account for a subset of paradigm effects: more specifically, a phonological generalization can underapply in a given word if there is a morphologically related item, capable of functioning as base, in whose output representation the relevant conditioning environment is absent; conversely, overapplication is possible when a phonological generalization is properly conditioned in the base, but not in the derived item. When none of the possible candidates to base status exhibits the requisite phonological properties, OO-correspondence has nothing to contribute. Additionally, TCT fails to account for misapplication effects which are not obviously paradigmatic; these include instances of opacity which do not involve overt morphological operations (affixation, compounding, truncation, etc.), as well as misapplication effects which are not consistently associated with specific morphological constructions (e.g. Benua 1997: §4.4.3).

The shortcomings of TCT have made it clear that strongly parallel OT must attack opacity with more powerful weapons than OO-correspondence. This realization has prompted the development of Sympathy Theory (McCarthy 1998). But Sympathy Theory suffers from ailments of its own, which are discussed in §3.4.2. First, the asymmetry of

☉O-correspondence relationships is established in a purely stipulative fashion. Secondly, ☉-candidates are morphosyntactically unmotivated (as anticipated in §3.3.1.3). Thirdly, the facts contradict McCarthy (1998) by requiring that ☉-candidates should be chosen not only by IO-faithfulness constraints, but also by subordinate markedness constraints; I shall provide evidence to this effect in §3.5. 'Extended' Sympathy Theory (Itô & Mester 1997b, de Lacy 1998) resigns itself to this state of affairs, but the unacknowledged consequences of this capitulation are dire. If markedness constraints are allowed to select ☉-candidates, then markedness reversal is surreptitiously introduced into strongly parallel OT (cf. §3.3.3.3). Moreover, McCarthy's functional justification of sympathy collapses. Following Kaye (1974, 1975), McCarthy (1998: 12-13, 14, 15) had suggested that sympathy constraints are functionally grounded: they allegedly aid lexical recognition by causing the output to reflect some property of the input indirectly, through the mediation of the ☉-candidate. However, this is only true if the ☉-candidate preserves an otherwise lost property of the input by virtue of being selected by a low-ranking IO-faithfulness constraint.

On the whole, current research in strongly parallel OT suffers from acute *embarras de richesses*. As CON undergoes a spectacular inflation in response to recalcitrant misapplication effects, the extent to which different theoretical devices overlap with or supersede each other becomes increasingly uncertain; e.g. see Itô & Mester (1997b: §4) and McCarthy (1998: 46) for the relationship between TCT and Sympathy Theory. Problems of overgeneration and analytical underdetermination are fast becoming unmanageable. More importantly, it has now become clear that strongly parallel OT provides a purely descriptive approach to opacity, as it fails to deduce the possibility of misapplication from more deeply rooted properties of phonological grammar. Neither OO-identity nor sympathy constraints have any justification beyond the facts they purport to explain. At a purely formal level, it is definitely untrue that the schemas of Correspondence Theory generalize from IO-faithfulness and BR-identity to OO-identity and sympathy without additional stipulation. From a functional viewpoint, attempts to ground ☉-constraints on considerations of parsing efficiency have collapsed as 'Extended' Sympathy Theory supersedes Sympathy Theory. In this sense, strongly parallel OT cannot answer the fundamental question which I posed at the outset of this chapter: why are human beings endowed with a linguistic faculty which enables them to learn and compute misapplying phonological generalizations, thereby allowing processes of change to give rise to opacity?

3.4.1 Paradigm effects

3.4.1.1 Alignment Theory

Indicative of the *embarras de richesses* to which I referred in the previous paragraph is the fact that strongly parallel OT furnishes two alternative tools to deal with paradigm effects: Generalized Alignment Theory (McCarthy & Prince 1993b) and TCT. Although neither provides a comprehensive solution, their empirical remits overlap to a large extent, creating problems of analytical underdetermination; see §3.4.1.3.

According to McCarthy & Prince (1993b), the universal set of markedness constraints includes the ALIGN family, whose members require that the (left or right) edges of designated categories should coincide. Their general format is as follows:

(3,26) **Generalized Alignment (McCarthy & Prince 1993b:30)**

$$\text{ALIGN}(\text{Cat1, Edge1; Cat2, Edge2}) =_{\text{def}}$$

$$\forall \text{Cat1} \exists \text{Cat2} \text{ such that Edge1 of Cat1 and Edge 2 of Cat2 coincide.}$$

Where

$$\text{Edge1, Edge2} \in \{\text{Right, Left}\}$$

This schema can be motivated on the basis of purely phonological phenomena, particularly in the realm of foot construction (McCarthy & Prince 1993b: §3); for an illustration, see the analysis of Spanish secondary stress assignment in (3,24). In these applications of ALIGN, both *Cat1* and *Cat2* are prosodic categories. However, McCarthy & Prince (1993b) claim that alignment constraints can also take morphosyntactic categories as arguments: i.e. in (3,26) they state that $\text{Cat1, Cat2} \in P(\text{prosodic})\text{Cat} \cup G(\text{grammatical})\text{Cat}$. Under this assumption, the location of morphological or syntactic category boundaries can directly determine the construction of prosodic constituents, particularly under the influence of high-ranking constraints of the format $\text{ALIGN}(G\text{Cat, Edge; PCat, Edge})$.⁴²

⁴² In recent work (e.g. McCarthy & Prince 1995; McCarthy 1998: 21), constraints with the format $\text{ALIGN}(G\text{Cat, Edge; PCat, Edge})$ have been replaced by ANCHOR constraints; see §3.4.1.3. The difference is immaterial to our purposes.

Constraints of this type can for example induce phonological misapplication at the edge of designated morphosyntactic domains.⁴³ In Biblical Hebrew, the constraint *CODA/PHAR excludes pharyngeal consonants from the coda (McCarthy & Prince 1993b: 126-127; Hoberman 1995: 841-842). Where necessary, violations of this constraint are avoided by means of epenthesis; see the examples in (3,27a), where epenthetic vowels are underlined (transcription as in the original). However, McCarthy and Prince observe that pharyngeal codas are tolerated in stem-final position; see (3,27b), where] indicates the right edge of the stem. They propose that *CODA/PHAR is outranked by ALIGN(Stem, R; σ, R), which requires that the final segment of the stem should be final in some syllable.⁴⁴

(3,27) Biblical Hebrew (McCarthy & Prince 1993b: 126-127)

ALIGN(Stem, R; σ, R) » *CODA/PHAR » DEP^σ

(a)		(b)	
ya.ʔ <u>a</u> mōd	'he will stand'	šā.maʔ.]	'he heard'
ye.ḥ <u>e</u> .zaq	'he is strong'	šālah.]tī	'I sent'
ya.h <u>a</u> pōk	'he will turn'	yədaʔ.]tem	'you knew'
ye.ʔ <u>e</u> .sōp	'he will gather'	yədaʔ.]tīka	'I knew you'

Similarly, constraints of the format ALIGN(GCat, Edge; PCat, Edge) can close a morphological domain off to prosodification. As an example, McCarthy & Prince (1993b:

⁴³ Anderson (1992: 289), drawing on Stanley (1973) among others, suggests that morphosyntactic structure may in general condition phonological generalizations in three main ways:

- A phonological generalization may hold only within the limits of a given morphosyntactically defined domain.
- A phonological generalization may hold only at the (right or left) edge of a given morphosyntactically defined domain.
- A phonological generalization may hold only within morphosyntactically complex (i.e. non-tautomorphic) domains.

(c) refers to blocking in non-derived environments and need not concern us any further here; see footnote 13. The Biblical Hebrew data in this paragraph illustrate (b), whilst the Polish data in the following paragraph illustrate (a).

⁴⁴ Without attempting an alternative analysis here, I shall merely note that under interleaving the treatment of this sort of edge phenomenon usually involves a combination of extraprosodicity and level ordering. In this case, one assumes that *CODA/PHAR dominates DEP^σ at the stem-level, but domain-final pharyngeals fail to trigger epenthesis by occupying an extrasyllabic position. If the relative ranking of *CODA/PHAR and DEP^σ is then reversed at lower levels, stem-final pharyngeals can lose their extrasyllabicity without triggering epenthesis.

128-129) cite the case of Polish, where resyllabification across the prefix-stem juncture (3,28a) or across word boundaries (3,28b) is forbidden (Rubach & Booij 1990: 442).

(3,28) (a)	roz.[ogniç	'heat'
	*ro.z[ogniç	
(b)	mechanizm.][obronny	'defense mechanism'
	*mechaniz.m][obronny	

In this case, McCarthy and Prince propose that ALIGN(Stem, L; ω, L) dominates ONSET.⁴⁵

In an interleaved régime, in contrast, effects of this nature are achieved indirectly. It is assumed that phonological generalizations need not refer to the edges of specified morphosyntactic categories; rather, constraints simply access the (edges of) the visible domain, where the notion 'visible domain' bears no morphosyntactic specifications such as 'root', 'stem', etc. Level segregation and (if appropriate) cyclicity independently ensure that only morphosyntactic categories of the correct type provide phonological domains for the relevant phonological generalization (see e.g. Anderson 1992: 245, 249, 289). Thus, whereas strongly parallel OT assumes that all morphological and syntactic structure is simultaneously accessible to the phonology (notably through hybrid alignment constraints), interleaved OT adheres to the generative tradition of attempting to limit access to non-phonological information. In particular, ALIGN constraints will have very limited access to non-phonological categories; I assume that ALIGN(Stem, R; σ, R) and ALIGN(Stem, L; ω, L) will have no place in the theory.

[Note, however, that there seems to be no good reason why interleaved OT should reject morphological constraints of the format ALIGN(Affix, Edge; Cat, Edge), which convey the type of information usually encoded in affixal subcategorization frames. These constraints have proved particularly useful in the analysis of phenomena such as prosodically conditioned infixation; see McCarthy & Prince (1993b: §4, §6).]

⁴⁵ Under interleaving, this putative alignment effect can be reduced to IO-faithfulness. Let us posit the following constraint:

IDENT^ω

Let α be a segment in the input; let β be a correspondent of α in the output.

Let ω_i be a prosodic word in the input; let ω_j be a correspondent of ω_i in the output.

If α is initial in ω_i , then β is initial in ω_j .

Now all one need assume is that (i) Polish prefixes are word-level, and (ii) Polish ranks ONSET below IDENT^ω at both the word and the postlexical levels; for the first assumption, see Rubach and Booij (1990: 459-461).

3.4.1.2 The priority of the base

TCT (e.g. Benua 1995, 1997; Itô & Mester 1997a; Kager 1995; Kenstowicz 1996) asserts that paradigm effects are caused by OO-correspondence constraints. These require that the surface representations of morphologically related lexical items should be identical in respect of some designated phonological property. OO-correspondence relationships are usually assumed to hold between *pairs* of forms, with a relatively simple lexical item standing as the 'base' of a morphologically more complex form: e.g. in §3.2.2.1, RP [bæŋ] and Belfast English [tɛːən] would be the bases of [bæŋə] and [tɛːənəɪ], respectively. Whilst the notion of 'base' as a surface form presents particular difficulties (see §3.4.1.3), this preliminary characterization will suffice to bring into focus another of the major drawbacks of OO-correspondence. Since correspondence relationships are inherently symmetrical, it is to be expected that OO-identity could be achieved through misapplication in either the base or the derived form. The empirical evidence suggests, however, that OO-correspondence relationships have in fact to be asymmetrical, in the sense that the derived phonological properties of a lexical item never trigger misapplication in its base. TCT's account of this observation, which Benua dubs the 'priority of the base', is inherently stipulative (Orgun 1996a: §5.1). Let us see why.

A correspondence constraint C is fulfilled whenever two correspondent elements α and β are identical in respect of a designated phonological property P . Correspondence relationships are symmetrical in the sense that, should α and β differ in respect of P , the constraint C does not specify which element — α or β — must be modified in order to enforce identity; rather, the choice depends on the nature of the mapping process. In the case of IO-faithfulness, this property of symmetry has proved to be a major theoretical asset. During phonological acquisition, output representations take precedence over the input: the learner takes surface forms as fixed and, in compliance with Lexicon Optimization, constructs underlying representations which differ minimally from the triggering data; see §3.3.3.2. During production, in contrast, the relative priority of input and output representations is reversed: now input forms remain inalterable, whilst the selection of output candidates minimizes unfaithfulness to the input (subject, of course, to the best satisfaction of the constraint hierarchy). In this sense, it is because they are symmetrical that IO-faithfulness constraints are equally relevant in acquisition as they are in production.

In production, however, positing correspondence relationships between two separate *output* representations γ and δ creates an interesting twist. In such circumstances, neither the content of γ nor that of δ is fixed beforehand; rather, the phonology computes the properties of both representations by reference to their respective inputs according to the requirements of the constraint hierarchy. Moreover, since neither representation is a constituent of the other, the two must be assumed to be generated in parallel. Therefore, if achieving identity between both representations turns out to require misapplication, there is no reason, in the absence of extraneous assumptions, to expect either γ or δ to take precedence in determining the outcome. What type of misapplication occurs (over- or under-application) and where (γ or δ) will rather be determined globally by the entire constraint hierarchy, including markedness, IO-faithfulness and OO-identity constraints.

In McCarthy & Prince's (1995) theory of reduplication, these predictions are strikingly borne out. In reduplicative morphology, one might expect misapplication effects to be circumscribed to the reduplicant, which is underlyingly unspecified and acquires its phonological content by copying the base. However, the facts prove otherwise: a process whose conditioning environment is only present in the reduplicant can be forced to overapply in the base so as to preserve BR-identity ('over-copying'; McCarthy & Prince 1995: §3.8). BR-identity is thus symmetrical, as predicted. McCarthy and Prince illustrate over-copying with the Chumash form /k-RED-ʔaniʃ/ 'my paternal uncles', where /k/ is a prefix and /ʔaniʃ/ is the base. In Chumash, an oral stop coalesces with an immediately following ʔ or *h*, respectively yielding a glottalized or an aspirated segment. Normal application would therefore yield *[k'anʔaniʃ], where [k'] results from the coalescence of prefixal /k/ with reduplicant-initial ʔ. Nonetheless, the correct output form is [k'ank'aniʃ], with overapplication of coalescence in the base. As McCarthy and Prince explain, over-copying preserves BR-identity whilst avoiding the unfused cluster *[kʔ], presumably forbidden by high-ranking *COMPLEX.

(3,29)

/k-RED-?aniʃ/	*COMPLEX	IDENT _{BR}	FAITH _{IB}	FAITH _{IR}
k-?an-?aniʃ	*!			
k'an-?aniʃ		*!		*
k'an-k'aniʃ			*	*

If OO-correspondence relationships were symmetrical like BR-identity, then one would expect to find instances of paradigm levelling in which a phonological process misapplies in the base in order to preserve identity with a non-basic form. According to Kenstowicz (1996: §4.3), relying on research by Flemming and Steriade, Latin rhotacism confirms this prediction. In Latin, certain tokens of underlying /s/ surface as [r] in intervocalic position: cf. (3,30b) and (3,30c). In third declension nouns, however, rhotacism applies to nom.sg. forms even in the absence of the proper conditioning environment; see (3,30a).

(3,30) Latin rhotacism (Kenstowicz 1996: 385)

(a) <i>nom.sg.</i>	(b) <i>gen.sg.</i>		(c)	
honor	honōris	'honour'	honestus	'honest'
arbor	arbōris	'tree'	arbustus	'wooded'
angor	angōris	'constriction'	angustus	'tight'

In the paradigm of *honor*, the *nom.sg.* should presumably be treated as the base, since, unlike the rest of forms, it bears no overt inflectional marker: cf. *acc.sg. honōrem*, *dat.sg. honōrī*, etc. (see §3.4.1.3). Thus, rhotacism seems to overapply in the base on the analogy of the affixed forms.

It is significant that, if one assumes the misapplication of rhotacism to be caused by an OO-identity constraint, then both the type of misapplication (viz. overapplication, as opposed to underapplication) and the direction of levelling (from affixed forms to the base) can be predicted without further stipulation. The right results fall out from the symmetrical nature of correspondence constraints, and from the fact that the optimal output candidates

for both the base and the derived form are selected in parallel. This is shown in tableau (3,31):

(3,31)

base form /hono:s/	derived form /hono:sis/	*VsV	IDENT _{OO}	FAITH _{IO}
honor	hono:sis	*!	*!	*
hono:s	hono:ris		*!	*
hono:s	hono:sis	*!		
honor	hono:ris			**

For rhotacism to be active in the language, *VsV must outrank FAITH_{IO}. As long as this ranking obtains, no possible reranking of IDENT_{OO} can generate underapplication (i.e. **honōs* ~ **honōsis*): the ranking of IDENT_{OO} relative to *VsV is non-crucial, whilst demotion of IDENT_{OO} with respect to FAITH_{IO} (i.e. *VsV » FAITH_{IO} » IDENT_{OO}) would trigger normal application (i.e. **honōs* ~ *honōris*).

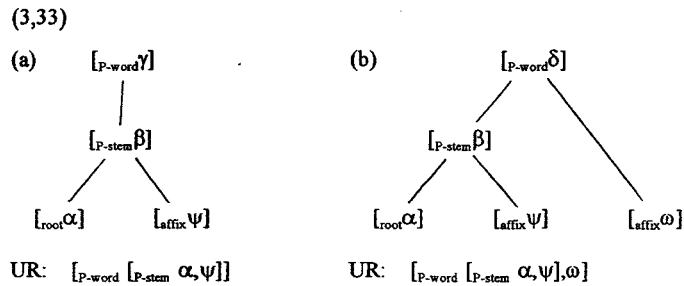
More generally, the case of Latin rhotacism illustrates two major predictions of symmetrical TCT, highlighted in table (3,32). First, the requirement that base and derived form should be identical can trigger overapplication in base forms, analogous to 'over-copying' in reduplicative morphology; see (3,32k). Secondly, underapplication, whether in base or derived forms, is impossible; see (3,32i) and (3,32j), respectively. This, I must stress, is what one would expect to happen if OO-constraints behaved like IO-faithfulness or BR-identity constraints, i.e. if Correspondence Theory attained the degree of generality which proponents of strongly parallel OT attribute to it.

Key	Base form		Derived form		Application pattern	Ranking schema	Prediction	
	Conditioning	Application	Conditioning	Application				
a	present	no	present	no	double misapplication	X	impossible	
b	absent	yes	absent	yes				
c	present	no	absent	yes				
d	absent	yes	present	no				
e	present	yes	present	no	(single) misapplication without OO-identity	X	impossible	
f	present	no	present	yes				
g	absent	no	absent	yes				
h	absent	yes	absent	no				
i	present	no	absent	no	underapplication in base	X	impossible	
j	absent	no	present	no				underapplication in derivative
k	absent	yes	present	yes				overapplication in base
l	present	yes	absent	yes	overapplication in derivative	M, IDENT _{OO} » FAITH _{IO}	possible	
m	present	yes	absent	no	normal application (with allomorphy)			
n	absent	no	present	yes	normal application (without allomorphy)	M » FAITH _{IO} » IDENT _{OO}	possible	
o	present	yes	present	yes				
p	absent	no	absent	no				

† Ranking of IDENT_{OO} indeterminate.

In fact, both predictions turn out to be wrong. First of all, Kiparsky (1998a) and Orgun (1996a, b) have persuasively argued that all known instances of derivative-to-base levelling can be reanalysed as involving restructuring of the underlying representation, so that the assumption of misapplication in the grammar becomes unnecessary; this is now accepted by proponents of TCT such as Benua and, according to the latter (1997: 241), Burzio. In the case of Latin rhotacism, for example, there are at least two possible alternatives to Kenstowicz's overapplicational approach. Benua (1997: 241) flatly denies that the noun *honor*, *-ōris* is synchronically derived from the same root as the adjective *honestus*; the noun stem would simply be /honor:/ . For his part, Kiparsky (1998c), cited in Kiparsky (1998a: 8), seems to suggest that the nom.sg. of *honor*, *-ōris* has acquired the vocalic ending of parasyllabic nouns such as *cūis*, *-is* 'citizen'; this ending triggers rhotacism before undergoing deletion. The available evidence, therefore, fails to confirm the prediction that OO-correspondence, being a symmetrical relationship, can force bases to behave opaquely; misapplication appears to be circumscribed to derived forms. Following Benua, I shall henceforth refer to this observation as the 'priority of the base'. The priority of the base implies that derivative-to-base levelling is not a synchronic grammatical phenomenon, but a type of phonological change giving rise to innovative underlying representations.

Tellingly, the impossibility of synchronic derivative-to-base levelling can be deduced from first principles in interleaved OT; the theory correctly predicts that the phonology of derived forms can only affect the 'base' by prompting the restructuring of its underlying representation. A particularly perspicuous proof of this theorem is possible in Orgun's declarative implementation of interleaving (Orgun 1996a: §5.1). Let $[\gamma]$ and $[\delta]$ be two members of the same inflectional paradigm, with the phonological domain structures represented in (3,33a) and (3,33b) respectively. Moreover, let S and W be the phonological functions respectively associated with stem-level and word-level domains. As can be seen, $[\gamma]$ bears no overt inflectional marker, and so can be assumed to be the output 'base' of $[\delta]$; see §3.4.1.3.



What happens in (3,33a) if, on the analogy of [δ], [γ] is replaced by [γ']? The tree will automatically become ill-formed, for the phonological representation associated with a node *N* is a function ρ of the representations associated with *N*'s daughters (see §3.3.1.3), but [γ'] does not satisfy the word-level function \mathcal{W} in relation to the stem [β]: i.e. [γ'] ≠ $\mathcal{W}([β])$. Imagine now that this is remedied by restructuring the stem: i.e. [β] is replaced with [β'], where $\mathcal{W}([β']) = [γ']$. The tree will nonetheless remain ill-formed, for the restructured stem will not satisfy stem-level phonology in respect of the root [α] and the stem-suffix [ψ]: i.e. [β'] ≠ $\mathcal{S}([α],[ψ])$. However, grammaticality can now be restored by replacing the root [α] with [α'], where $\mathcal{S}([α'],[ψ]) = [β']$.⁴⁶ It follows that a derived form [δ] can only affect the phonological representation of its output 'base' [γ] by triggering the restructuring of [γ]'s underlying representation; phonological misapplication in respect of [γ]'s original underlying representation is impossible.

Let us now turn to the second incorrect prediction of symmetrical TCT: the impossibility of underapplication. It is true that, in compliance with the priority of the base, paradigm-induced underapplication in base forms does not occur synchronically; in derived forms, however, underapplication is well-documented. In §3.2.2.1, for example, I mentioned the dentalization of alveolar non-continuants before rhotic sounds in Belfast English: the process underapplies in the noun *tenner* [tɛːnɔ̃ɹ], allegedly in order to preserve identity with the base *ten* [tɛːən], where the rhotic conditioning environment is absent (cf. [tænɔ̃ɹ] *tenor*). Yet under the ranking schema {M, IDENT_{OO}} » FAITH_{IO}, symmetrical parallel evaluation of both items predicts overapplication in the base: i.e. *[tɛːən]~*[tɛːənɔ̃ɹ]; cf. (3,31) above, and see Benua (1997: 192-193). However, we now know this prediction to be necessarily incorrect, for there is no synchronic derivative-to-base levelling.

There thus emerges a crucial difference between BR-identity and OO-correspondence, striking right at the heart of the claims of Correspondence Theory to provide a unified (and, to that extent, explanatory) account of all misapplication phenomena. Despite the expectation of symmetry, OO-constraints must somehow be prevented from triggering misapplication in 'base' forms (Benua 1997: §2.4.3). Paradigmatic underapplication occurs precisely when base priority blocks overapplication: i.e. in cases where the conditioning environment of some active phonological process is absent in the base but present in the derivative; see (3,32j) and (3,32k). In reduplicative morphology, in contrast, correspondence relationships between the base and the reduplicant prove to be symmetrical, as one would expect, and misapplication in the base is permitted (as e.g. in 'over-copying'). Moreover, reduplicative underapplication is not triggered by some extraneous principle, but follows directly from constraint ranking: it is predicted to occur when a candidate displaying overapplication violates some high-ranking markedness constraint (McCarthy & Prince 1995: §5). Consider the following two examples:

(i) Akan (McCarthy & Prince 1995: §5.1) has a monosyllabic reduplicative prefix whose vowel is obligatorily high. In the mapping /RED-ka/ → [kɪka] 'bite', the reduplicant shows underapplication of a process that palatalizes velars before front non-low vowels. Whilst normal application (*[tɛɪka]) would have disrupted BR-identity, overapplication (*[tɛɪtɕa]) is ruled out by a high-ranking OCP-related constraint against coronal sequences.

(ii) In Chumash (McCarthy & Prince 1995: §5.2), the constraint RED=σ^{mu} is superordinate, compelling all reduplicants to conform with a CVC template. In the mapping /ɟ-RED-tal'ik/ → [ɟaltal'ik] 'his wives', RED=σ^{mu} forces underapplication of pre-coronal /-deletion in the reduplicant; normal application (*[ɟatal'ik]) would result in an ill-formed *CV reduplicant. Note that overapplying *[ɟataik'] is out for the same reason.

Crucially, neither of these underapplication-triggering mechanisms accounts for underapplication in non-reduplicative morphology. Since BR-correspondence is a relationship between co-present constituents (i.e. *in praesentia*), markedness constraints can have scope over an entire reduplicant + base ensemble, as in the Akan case; in OO-correspondence, in contrast, the relationship between basic and non-basic forms is paradigmatic (*in absentia*), and cannot be assessed by markedness constraints (Benua 1997: 48-49). Similarly, prosodic effects such as reduplicant bimorcity do not generalize to ordinary affixes which may trigger underapplication without being the domain of special prosodic conditions.

⁴⁶ Of course, for this restructuring to work it must be the case that $\mathcal{W}([β'],[ω]) = [δ'] \neq [δ]$.

As I pointed out above, these observations indicate that BR-identity and OO-correspondence are fundamentally heterogeneous. The former is a symmetrical relationship between co-present syntagmata, and complies with general principles of harmonic evaluation. In contrast, the latter holds *in absentia* and, given the inherent symmetry of correspondence constraints, requires *ad hoc* theoretical provision to enforce the priority of the base. This result is of the utmost theoretical importance. To the extent that the analogy between reduplicative and non-reduplicative morphology turns out to be misleading, the empirical success of McCarthy & Prince's (1995) approach to reduplication fails to lend support to TCT. At the same time, the fact that the priority of the base requires *ad hoc* theoretical provision undermines the claim that Correspondence Theory can provide a unified account of opacity (cf. §3.2.2.1). McCarthy & Prince (1995: §3.8) assert that, "Correspondence Theory generalizes over reduplicative identity, faithfulness, and other phonological relations, avoiding any reduplication-specific mechanism"; but it is OO-correspondence, rather than reduplicative identity, that turns out to require *ad hoc* elaboration.

Significantly, it is implicit in our discussion that interleaved OT need not sacrifice the insights of McCarthy & Prince (1995) regarding reduplication. Since BR-identity constraints impose conditions of maximum harmony on the output representation of reduplicated words, they can be classed as markedness constraints under (2,45c), and are therefore compatible with Strict Surface Orientation; the fact that markedness is in this case defined as disparity between two constituents of the target representation is irrelevant, for the relationship between these constituents is syntagmatic, rather than transderivational. Thus, an interleaved OT model incorporating BR-identity predicts that symmetrical misapplication effects such as over-copying are possible at the level where reduplication takes place, as base and reduplicant are derived in parallel from the input to that level. (Moreover, assuming some form of Bracket Erasure, it is predicted that BR-identity will cease to be active at lower levels.)

It should now be clear that, if it were to be countenanced at all, TCT would have to adopt some theoretical stipulation specifically designed to enforce the priority of the base. Quite what this stipulation should be is not apparent at first blush. As we have seen, Kenstowicz (1996) fails to reject misapplication in bases, but he is nevertheless aware of the need to have asymmetrical OO-identity constraints. However, he seems to assume that base priority can simply be written into the statement of such constraints. This flies in the face of McCarthy & Prince's (1995) explicit claim that the schemas of Correspondence

Theory (MAX, DEP, IDENT, etc.) generalize over all correspondence relationships. MAX constraints, for example, must conform with the following template:

(3,34) MAX (McCarthy & Prince 1995: Appendix A)

Every element x of S_1 has a correspondent in S_2 .

If we substitute $B(ase)$ for S_1 and $O(utput)$ for S_2 , we obtain a constraint which is every bit as reversible as a MAX_{IO} constraint: if under normal application x is present in B but absent in O , compliance with MAX_{BO} can be brought about in either of two ways, viz. by deleting x in B or by inserting x in O .

Benua's (1997) ingenious solution to the problem deserves our attention because it brings the inherent inadequacies of strongly parallel approaches to paradigm effects into sharp focus. As Benua observes, the priority of the base implies that misapplication in bases is always less harmonic than misapplication in non-basic forms. From this she deduces that any constraint violation incurred by the base is infinitely more serious than any constraint violation incurred by the derivative. She therefore proposes that, in the evaluation of any base-derivative pair, the constraint hierarchy \mathcal{M} is duplicated, yielding two exact copies of itself: one, \mathcal{M}_{base} , indexed to the base form; the other, $\mathcal{M}_{derivative}$, indexed to the derivative. Crucially, the constraint hierarchy indexed to the base dominates the constraint hierarchy indexed to the derivative: i.e. $\mathcal{M} = \{ \mathcal{M}_{base} \gg \mathcal{M}_{derivative} \}$.

Let us see how hierarchy recursion generates underapplication in the case of Belfast pre-rhotic dentalization (Benua 1997: §5.3.1). Recall that in Belfast English coronal non-continuants are generally alveolar, but become dental when followed by a rhotic (§3.2.2.1). Unfortunately, the details of the process are unclear at the subsegmental level: both J. Harris (1989: 40) and Benua (1997: 188) assume that the feature [distributed] is involved, but whereas J. Harris takes Belfast /r/ to be [+distributed] and treats dentalization as a process of assimilation, Benua assumes /r/ to be [-distributed] and treats the process as dissimilation. Given this lack of accord, I will not attempt a precise formulation of the markedness constraints involved; in (3,35) they are simply labelled M^1 and M^2 , and their effects are described informally. The low ranking of $IDENT_{IO}^{[distr]}$ renders the value of [distributed] predictable (Benua 1997: §5.3.1.1).

(3,35) $\mathcal{M} : M^1 \gg M^2 \gg \text{IDENT}_{\text{IO}}^{[\text{distr}]}$

M^1 : A coronal stop followed by a rhotic must be dental.

M^2 : A coronal stop must be alveolar.

In Benua's view, the evaluation of a pair of correspondent forms such as *ten-tenner* calls for a duplicated constraint hierarchy, where constraints indexed to the base *ten* dominate constraints indexed to the derivative *tenner*.

(3,36) $\mathcal{M}_{\text{base}} \gg \mathcal{M}_{\text{derivative}}$

$\text{BASE-}M^1 \gg \text{BASE-}M^2 \gg \text{BASE-IDENT}_{\text{IO}}^{[\text{distr}]} \gg \text{DERIV-}M^1 \gg \text{DERIV-}M^2 \gg$

$\text{DERIV-IDENT}_{\text{IO}}^{[\text{distr}]}$

Since *ten* satisfies $\text{BASE-}M^1$ vacuously (the rhotic trigger is absent), the final consonant of the base must surface as alveolar in compliance with $\text{BASE-}M^2$. In normal application, $\text{DERIV-}M^1$ would in turn trigger dentalization in *tenner*, overriding subordinate $\text{DERIV-}M^2$. Assume, however, that high-ranking $\text{IDENT}_{\text{OO}}^{[\text{distr}]}$ requires that the consonants of *ten* and *tenner* should agree for the value of [distributed]. It is now clear that misapplication will have to take place in the derivative, for $\text{BASE-}M^2$ dominates $\text{DERIV-}M^1$ by virtue of hierarchy recursion:

(3,37)

base	derivative	$\text{IDENT}_{\text{OO}}^{[\text{distr}]}$	$\text{BASE-}M^2$	$\text{DERIV-}M^1$
tɛːən	tɛːənɚɪ	*!		
tɛːən	tɛːənɚɪ		*!	
tɛːən	tɛːənɚɪ			*

Consider now an analysis based on interleaving. Assuming level segregation, one need only stipulate that dentalization is active at the stem level, but ceases to apply at the

word level (J. Harris 1989).⁴⁷ Thus, whilst simplex stems and class-I derivatives undergo dentalization normally (e.g. [tɛːnɚɪ] *tenor*, *elɛmɛ[nɪ]ary*), the addition of word-level suffixes such as *-er* counterfeeds dentalization.

(3,38) Belfast English *tenner*

UR: [P-word[P-stem tɛːən]ɚɪ] ⁴⁸				
STEM LEVEL		M^1	M^2	$\text{IDENT}_{(\text{OO})}^{[\text{distr}]}$
[tɛːən]	a) tɛːən		*!	*
	b) tɛːən			*
WORD LEVEL		$\text{IDENT}_{(\text{OO})}^{[\text{distr}]}$	M^1	M^2
[[tɛːən]ɚɪ]	c) tɛːənɚɪ	*!		*
	d) tɛːənɚɪ		*	

The priority of the stem-level representation (3,38b) over word-level (3,38d), with misapplication occurring in the latter but not the former, is a straightforward consequence of the input-output relationship holding between the two; this, in turn, is determined by morphological constituency. In this respect, the contrast with Benua's analysis is illuminating. In compliance with Strong Parallelism, Benua asserts that phonology is non-recursive; accordingly, she must deny that the innermost set of brackets in [[tɛːən]ɚɪ]

⁴⁷ In rule-based LPM, assigning pre-rhotic dentalization to the stem level seemed problematic. Since the feature [distributed] is non-contrastive in Belfast English, it was assumed to be underlyingly unspecified. However, this meant that applying dentalization at the stem level would infringe Structure Preservation, as the rule would violate constraints on underlying representations by specifying values for [distributed] (J. Harris 1989: 41; Borowsky 1993: 209-210). For this reason, Borowsky insisted that dentalization must be word-level. In interleaved OT, in contrast, these qualms are shown to be unfounded. Under Richness of the Base, there is no assumption of underspecification of allophonic features (§2.2.2.1); the content of underlying representations is determined by stem-level rankings through Lexicon Optimization (§3.3.1.1, §3.3.2, §3.3.3.2). Thus, [distributed] can be present in underlying representations without creating contrasts; its (predictable) distribution will be that which best satisfies the stem-level constraint hierarchy.

⁴⁸ Here, I leave the nasal underlyingly unspecified for [distributed] just in order to illustrate the operation of stem-level rankings; under Lexicon Optimization, the optimal underlying representation has an alveolar nasal (see the previous footnote). In contrast, the nasal of *tenor* [tɛːnɚɪ] is non-contrastively dental underlyingly.

defines a phonological domain (in fact, *the* domain of dentalization). However, this strong stance against recursion soon proves untenable, for the surface form *ten*, enlisted to do the job of (3,38b), must take priority over *tenner*, if **te[ŋ]~*te[ŋ]er* is to be avoided. This turns out to require recursion of the constraint hierarchy, for correspondence constraints are inherently symmetrical.

The general implications are clear. If phonology and morphology are interleaved, then the surface representation of a word may become opaque in relation to its underlying representation because a phonological subdomain (corresponding to the base of the word in the ordinary sense) undergoes phonology on its own and then provides the input to a subsequent construction (see §3.3.1.1 and §3.3.2). Thus, under the hypothesis that phonology is recursive, the priority of bases over their derivatives requires no stipulation: it is implicit in the fact that, in production, inputs take precedence over outputs. I shall henceforth refer to bases in this ordinary sense as '*i*-bases', to distinguish them from bases in the sense of TCT or '*o*-bases'. Strong Parallelism, however, denies access to *i*-bases as phonological entities, and forces TCT to substitute the surface representation of some related lexical item (the *o*-base) as a source of opacity. But, unlike an *i*-base and its derivative, an *o*-base and its correspondent do not stand in an input-output relationship to each other, for the *o*-base is not a morphological constituent of the derived form. Given the inherent reversibility of correspondence constraints, the relationship between *o*-bases and derivatives is therefore erroneously expected to be symmetrical. To circumvent this obstacle, TCT must adopt some stipulative theoretical provision. Benua (1997) opts for a simulacrum of recursive mapping, with the recursion of \mathcal{M} giving absolute ranking priority to constraints on the *o*-base over constraints on the derivative. However, this manoeuvre has no parallel in BR-identity; its *ad hoc* nature belies the claim that Correspondence Theory can provide a fully general solution to the problem of misapplication.

3.4.1.3 Masking

All the difficulties under which TCT labours spring from a single cause: trying to get *o*-bases to do the job of *i*-bases. For some of these difficulties there are workable, albeit stipulative, solutions: as we have seen, Benua is able to assert the priority of *o*-bases over their correspondents by cloning the constraint hierarchy, and by indexing and ranking the resulting clones. However, other problems are ultimately insurmountable. These relate to 'masking' effects and will be discussed in this section.

In an interleaved régime, a morphological construction \mathcal{M} can create opacity in two ways: in the derived form, the conditioning environment of some phonological process undergone by the *i*-base may be destroyed (overapplication); alternatively, the derivative may contain the conditioning environment of a process which would have been applicable to the *i*-base but is no longer active (underapplication). In either case, \mathcal{M} can only render opaque those phonological properties which are present in the *i*-base. Thus, paradigm effects never involve postlexical phonological generalizations: since the postlexical phonology applies after all the morphology (or, in declarative terms, P-stems and P-words never dominate P-phrases), *i*-bases will never bear postlexical properties. As Kiparsky (1998a: 26) puts it,

Paradigmatic transfer of phonological properties from bases to derivatives involves only those properties which are represented at the level at which the morphology is introduced.

Moreover, both the *i*-base and the derivative may be subject to further grammatical operations, whether morphosyntactic or phonological. If, for example, the *i*-base is a bound stem, then each of its exponents will have undergone some additional morphological operation before it reaches the surface. Minimally, both the *i*-base and the derivative will in any case pass through the postlexical level. By implication, it is possible for the opaque property of the derivative not to surface anywhere else at all, or to surface only in forms where it has also become opaque (McCarthy 1998: 7).

All this puts TCT in a quandary. TCT can account for a misapplication effect in a word only if its opaque phonological property happens to occur in some related surface form, capable of qualifying as an *o*-base; i.e. TCT predicts no misapplication without surface identity —see (3,32e) to (3,32h). Moreover, the phonological property in question must be transparent in the *o*-base; otherwise, the solution of the problem is simply deferred —see (3,32a) to (3,32d). However, our discussion shows that the crucial properties of the *i*-base may be *masked* in all candidates to *o*-basehood: *o*-bases may be cluttered with irrelevant (e.g. postlexical) phonological information which cannot undergo paradigm levelling, or they may no longer exhibit the transferred property in the requisite transparent state, if at all. OO-identity constraints are thus doomed to failure, for they can access at once too much information and not enough. In a vain effort to remedy these flaws, the concept of *o*-base has been tweaked and twisted a great deal in the short history of TCT.

To the extent that an overall trend may be discerned in the evolution of the concept, it unsurprisingly leads towards increasingly lax definitions, as advocates of OO-correspondence sidestep the problem of overgeneration and focus on recalcitrant paradigm effects. Tellingly, all definitions of *o*-basehood incorporate *ad hoc* criteria for morphological relatedness between surface forms, with blithe disregard for word-formation theory. The trend reaches its logical conclusion in Benua (1997), where the *o*-base of an affixed form is assumed to be specified in the subcategorization frame of the affix (Benua 1997: 30); in this extremely liberal view, the *o*-base is whatever surface form is stipulated to be the *o*-base.

As regards its logic, the charge of overgeneration against TCT is impregnable. Since *o*-bases are output forms, they will display phonological properties acquired in their pass through the postlexical level; these properties are thereby rendered accessible to OO-identity constraints. But paradigm effects never transfer postlexical features; therefore, OO-correspondence must overgenerate. Some scholars have challenged the empirical validity of the premises; but such challenges have not proved decisive. A representative case is that of /t/ flapping in American English. The distribution of the flapped allophone is a prototypical example of a postlexical phonological generalization (Kaisse & Shaw 1985: 4). Yet, remarkably, Withgott (1983) observes that the contrast between [t] and [ɾ] in the words *militaristic* [ˌmɪlətəˈɪstɪk] and *capitalistic* [ˌkæpəˈɪstɪk] mirrors the allophony of their respective bases: *military* [ˌmɪləˈtɛəri] and *capital* [ˌkæpəˈrɪt]. Steriade (1996) infers that postlexical features are indeed available for paradigmatic transfer. Nonetheless, Kiparsky (1998a: §2.2.3) points out that there is an alternative analysis. The foot structure of *militaristic* is arguably different from that of *capitalistic* as a result of the interaction between stress assignment and morphology in the lexical phonology. If so, postlexical flapping need not be assumed to misapply, but can be viewed as properly conditioned by foot structure (see also Pater 1995: appendix C for references and related issues).

Inevitably, therefore, the concept of *o*-base casts the net too wide, as it renders an excessive amount of phonological information available for paradigmatic transfer through OO-identity. However, I have also suggested that the concept of *o*-base will always prove too narrow, for an opaque phonological property in a derivative may be 'masked' (i.e. lost or rendered opaque) in all the surface forms which could conceivably stand as *o*-base. For this reason, definitions of *o*-base have become looser and looser over time. A permanent feature of such definitions is the requirement that an *o*-base must be a free form, i.e. a

possible output. However, authors have progressively increased the permissiveness of the 'compositionality' criterion, which dictates that the derivative and the *o*-base must be morphologically related. As I shall show, the notion of relatedness which proponents of TCT entertain has little, if anything, to do with word-formation patterns.

Kager (1995: §2.1) provides the following statement (boldface provided):

[...] I will use the notion of 'base' in a specific sense, namely as a form that is *compositionally related* to the affixed word in a morphological and a semantic sense. **(The meaning of the affixed form must contain all grammatical features of its base.)** Moreover, the base is a free form, i.e. a *word*. This second criterion implies that a base is always an output itself.

Kager's interpretation of the compositionality criterion is obviously too strict, for it implies that two words cannot enter into OO-correspondence if their bear different values for an obligatory grammatical feature, even if the phonological expression of this feature is covert in one of the forms (Kiparsky 1998a: §2.4.1). Thus, Kiparsky notes that the English singular noun *pass* cannot, according to Kager's definition, be the *o*-base of the plural form *passes*, for the two disagree in number features, even though SG has no overt phonological exponence. Yet, following Benua (1995), Kager would impute the overapplication of Philadelphia *æ*-tensing in [pɛːəsəz] *passes* to levelling from the singular [pɛːəs] *pass* (cf. [pæsɪv] *passive*).⁴⁹

Kenstowicz's (1996: 370) approach to compositionality is less stringent:

Base-Identity: Given an input structure [X Y] output candidates are evaluated for how well they match [X] and [Y] if the latter occur as independent words.

Assuming *X* and *Y* to stand in each case for a string of phonological input material associated with some morpheme or set of morphemes, Kenstowicz's statement implies that a word *α* can stand as *o*-base to another word *β* if *α* consists of a proper subset of the overt morphemes of *β*, where a morpheme is 'overt' if its phonological exponence is underlyingly non-null. Observe that this is the intuitive definition of *o*-base with which I operated in §3.4.1.2. It picks out Latin *honor* as the *o*-base of *honōris*, even though the two

⁴⁹ The literature of Philadelphia *æ*-tensing is copious. Labov (1989) describes the empirical facts in exquisite detail; for *pass-passes* see page 36.

words disagree in case features; similarly, it allows OO-correspondence between *pass* and *passes*.

As Kiparsky (1998a: 14) observes, Kenstowicz's compositionality criterion causes OO-correspondence to overgenerate in a peculiar way, for it allows words to have multiple *o*-bases: e.g. *politicization* would have both a proximate *o*-base *politicize* (obtained by subtracting the morpheme *-ation*) and a remote *o*-base *politic* (subtract *-iz-ation*). However, identity with a remote base violates well-corroborated locality restrictions, which interleaving captures in a straightforward fashion: since a derivative only stands in a direct input-output relationship with its immediate *i*-base, it will not display faithfulness to remote *i*-bases. Predictably, cyclic secondary stress in *politicizátion* indicates faithfulness to a proximate, rather than a remote, *i*-base (cf. *politicize* and *pólitics*).

More generally, Kenstowicz's compositionality criterion is suspect in that it defines lexical relatedness in terms of sets of (non-null) morphemes, rather than actual word-formation operations; as a result, it will be led astray by morphological processes involving affix subtraction. Consider, for example, the verb *politicize*. If *politicize* is to be regarded as a denominal formation, then its base will be the noun stem *politics*; this would involve a word-formation rule $[_N Xs] \rightarrow [_V Xize]$; see Anderson (1992: ch. 7). If, on the contrary, *politicize* is to be treated as a deadjectival formation, then the semantics suggest that the base is the adjective stem *political*, rather than *politic* (which only means 'political' in archaic expressions such as *the body politic*); the morphological construction involved would in that case be $[_A Xa] \rightarrow [_V Xize]$. Whatever solution we adopt, it is clear that the actual base of *politicize* will not consist of a proper subset of its overt morphemes, for both $[_N Xs] \rightarrow [_V Xize]$ and $[_A Xa] \rightarrow [_V Xize]$ involve affix subtraction as well as addition. Yet the only surface form which satisfies Kenstowicz's compositionality criterion in relation to *politicize* is *politic*, which, if anything, is a remote base.

Kenstowicz's definition of *o*-base also undergenerates because languages often have groups of morphologically related words where no member is exhaustively composed of a proper subset of the overt morphemes of another member. Obvious instances of this state of affairs are inflectional paradigms in which each form bears a different overt inflectional affix; in such cases, the *i*-base is a *bound* stem. According to Kenstowicz's compositionality criterion, such paradigms have no *o*-base. Yet there is profuse evidence that bound stems can trigger (and undergo) misapplication (Orgun 1996a: §5.2, Kiparsky 1998a: §2.3). Drawing on data from Hyman (1994), Orgun provides an interesting example from the Bantu language Cibemba, where the superclose vowel *j* triggers mutation of a preceding

non-nasal consonant, so that labials turn to *f*, and non-labials to *s*. Accordingly, when the causative suffix *-j-* is appended to the root *-leep-* 'be long', the resulting causative stem displays mutation: *-leef-j-* 'lengthen'. Crucially, the applicative-causative stem *-leef-es-j-*, which is derived from the causative stem by infixing the applicative morpheme *-el-*, undergoes overapplication of mutation in the root-final consonant. This phenomenon does not submit to an analysis where the applicative-causative stem *-leef-es-j-* (input *-leep-el-j-*) observes OO-identity with the causative stem *-leef-j-*, for both stems are bound forms; neither occurs as an independent word.

Faced with the problem of derived bound stems, Benua (1997) abandons all attempts to identify *o*-bases by means of a general compositionality criterion. She assumes instead that OO-correspondence relationships are stipulated in the subcategorization frame of affixes:

In addition to their segmental content (if any), affixal morphemes are supplied with a subcategorization frame that specifies idiosyncratic information about the affix, such as its selectional restrictions, and whether it is a prefix or a suffix. **I propose that the affix's subcategorization frame also specifies the OO-correspondence relation that links the affixed output in a paradigmatic identity relation.**

(Benua 1997: 30; boldface provided)

In this view, the *o*-base need not consist of a proper subset of the overt morphemes of the derivative. As an example, let the surface forms *Xα* and *Xβ* be two members of the same inflectional paradigm, where *-α* and *-β* are overt affixes which realize contrasting grammatical features. According to Benua's proposal, *Xα* will stand as *o*-base to *Xβ* if the subcategorization frame of *-β* includes a statement to the effect that *o*-base($\mathcal{P}([X]\beta)$) = $\mathcal{P}([X]\alpha)$, where \mathcal{P} is the phonological function.⁵⁰ Following Kraska-Szlenk (1995: 108ff.) and Kenstowicz (1996: §4.4), Benua (1997: 106-107, 236-237) suggests that precisely this sort of situation obtains in Polish, where gen.pl.fem. diminutive forms such as *kr[u.]wek* 'cow' display overapplication of a process raising [o] to [u] in closed syllables. Misapplication is imputed to OO-identity with nom.sg.fem. diminutive forms such as *kr[u.]w.ka*, where raising is properly conditioned. It is implied that the gen.pl.fem.

⁵⁰ This formal interpretation is my own; Benua (1997) fails to specify how this aspect of subcategorization frames is to be notated.

diminutive suffix subcategorizes for nom.sg.fem. diminutive forms as *o*-bases, even though the latter are overtly marked for incompatible case and number features.⁵¹

It is easy to demonstrate that, by decoupling *o*-basehood from morphological constituency, Benua's version of TCT becomes vastly permissive. To establish this point, consider a hypothetical language *L* where all nouns are overtly inflected for case (nom. vs acc.) and number (sg. vs pl.); assume, moreover, that all derivation is ascribed to the stem level, whereas all inflection takes place at the word level. In an interleaved framework, *L*'s noun paradigms will conform with the following pattern:

(3,39)	<i>sg.</i>	<i>pl.</i>
<i>nom.</i>	[_{P-word} [_{P-stem} X _N] α]	[_{P-word} [_{P-stem} X _N] γ]
<i>acc.</i>	[_{P-word} [_{P-stem} X _N] β]	[_{P-word} [_{P-stem} X _N] δ]

Observe that in (3,39) all forms share the same *i*-base, viz. the noun stem [_{P-stem} X_N]. Opacity can therefore arise in two ways only: inflection may destroy the conditioning environment of a process undergone by [X] at the stem level (overapplication), or create the conditioning environment of a process which [X] failed to undergo at the stem level (underapplication). Consider now the possibilities in Benua's framework. We assume *ex hypothesi* that *L*'s inflectional affixes all belong to the same level; in terms of TCT, this translates into the assumption that they are all members of the same correspondence class (let's call it class *c*), i.e. they subcategorize for the same set of OO-identity constraints (see §3.3.1.2). Still, Benua's proposals allow each affix to stipulate a different surface form as *o*-base. In this sense, the relevant subcategorization frames could be annotated as follows:

⁵¹ One conceptual difficulty for TCT is how to prevent OO-identity constraints from blocking the overt realization of affixes which are absent from the *o*-base (Benua 1997: §2.3.2). This difficulty is compounded by allowing an *o*-base to bear inflectional markers which are incompatible with those of its derivative; in this case, OO-identity constraints could conceivably trigger the transfer of affixal material from the *o*-base to the derivative (Benua 1997: 237). Benua suggests that these effects may be avoided in a variety of ways: one option is to relativize OO-identity to the root or stem portion of the *o*-base (see §2.3.4.2 for relativized correspondence); alternatively, morphological constraints may be called upon to block double or incompatible marking.

(3,40)	<i>nom.sg.</i>	[[_N X]α]		
	<i>acc.sg.</i>	[[_N X]β]	affix class <i>c</i>	<i>o</i> -base(ρ ([[X]β])) = ρ ([[X]α])
	<i>nom.pl.</i>	[[_N X]γ]	affix class <i>c</i>	<i>o</i> -base(ρ ([[X]γ])) = ρ ([[X]β])
	<i>acc.pl.</i>	[[_N X]δ]	affix class <i>c</i>	<i>o</i> -base(ρ ([[X]δ])) = ρ ([[X]γ])

This allows the following possibilities:

- (i) the acc.sg. can become opaque by copying transparent properties of the nom.sg.;
- (ii) the nom.pl. can become opaque by copying properties of the acc.sg., whether these are transparent or levelled from the nom.sg.;
- (iii) the acc.pl. can become opaque by copying properties of the nom.pl., which may be transparently motivated in the nom.pl., opaquely copied to the nom.pl. from the acc.sg., or opaquely copied to the nom.pl. from the nom.sg. via the acc.sg!

In other words, Benua's proposal allows multiple levelling effects within the paradigm, each triggered by a different allomorph of the stem;⁵² moreover, opacity can become increasingly complex by cascading across surface forms. Such promiscuous levelling stands in stark contrast with the predictions of interleaved OT: in the situation described in (3,39), the stem-level representation of [X] is the only possible source of opacity.

Not even in so unconstrained a model as Benua's, however, is it possible to guarantee that the crucial properties of the *i*-base will surface transparently in some candidate to *o*-basehood. To prove this assertion, let me return to the behaviour of coda nasals in Spanish velarizing dialects (§3.3.2). Recall that the paradigm of the noun *desdén* 'disdain', derived from the root /desdeɲ/, comprises two forms: the sg. [ɖes.ˈðeŋ] and the pl. [ɖes.ˈðe.nes]. The plural exhibits overapplication of the stem-level default process which assigns alveolar place to unassimilated coda nasals; cf. transparent *[ɖes.ˈðe.nes].⁵³ As we saw in §3.3.2, this opacity effect is due to the fact that the plural affix *-es* is not introduced until the word level; consequently, underlying /ɲ/ fails to occupy an onset position at the stem level, and is therefore unable to support its palatal specification. For

⁵² Commenting on his TCT analysis of Polish diminutives, Kenstowicz (1996: 389) observes that the stipulated *o*-base, viz. the nom.sg., would commonly be regarded as the morphologically unmarked member of the paradigm. However, neither Kenstowicz nor Benua suggest that the notion of morphological markedness could be formalized in such a way as to constrain the choice of *o*-bases.

⁵³ Of course, [ɖes.ˈðe.nes] is grammatical as the realization of the 2sg.pres.subj. of *desdeñar*.

exposing the *embarras de richesses* which currently besets strongly parallel OT.⁵⁴ Furthermore, the case of Spanish coda nasals shows that, in practice, the deployment of anti-opacity strategies in strongly parallel OT is entirely opportunistic and unprincipled; whether a particular paradigm effect will be tackled through TCT or Generalized Alignment depends on the vagaries of masking and the chance availability of a serviceable *o*-base.

The alignment approach to misapplication in *desdenes* and *desdén absoluto* incurs one further serious problem: massive loss of generalization. The behaviour of resyllabified nasals is intimately connected to that of nasals in other neutralizing non-assimilation environments. Interleaved OT captures these connections in a straightforward manner (see §3.3.2 again):

(i) In velarizing dialects, the default place of articulation for nasals at the word level is velar; see (3.20b). This accounts for place assignment to nasals in *three* environments:

- word-finally in the coda (i.e. preconsonantly or prepausally),⁵⁵
- word-finally in the onset (i.e. in the environment of postlexical resyllabification),
- in heteromorphemic [N.n] clusters (created by word-level suffixes such as *con-*, *en-*, *in-*).

(ii) The default place of articulation for nasals at the stem level is alveolar, provided that no violation of OCP^{Place} is incurred; see (3.22). Stem-final nasals subject to word-level resyllabification into the onset therefore surface as alveolar. In tautomorphemic [N.n]

⁵⁴ Sympathy Theory (McCarthy 1998) compounds the underdetermination problem, as the overapplication of word-level velarization under postlexical resyllabification can be imputed to faithfulness to a \otimes -candidate selected by \star -ANCHOR₁₀(Word, σ , R); see §3.4.2.1 for technical details.

/tren#aθul/ 'blue train'	ONS	CODACOND	IDENT ₁₀ ^{Place}	ONS-IDENT ₁₀ ^{Place}	*n	*ŋ	★-ANCHOR ₁₀
trɛn.a.θul	*!	*!	*!		*		✓
trɛŋ.a.θul \otimes	*!	*!				*	✓
trɛ.na.θul \otimes			*!		*		*
trɛ.ŋa.θul \otimes				*		*	*

⁵⁵ In Spanish, word-final nasals do not phonologically assimilate to the place of articulation of an immediately following consonant, even though the appearance of assimilation may arise as a perceptual epiphenomenon of gestural overlap in the phonetics (see Padgett 1995: 31 and references therein); for evidence that the same situation obtains in English, see Browman & Goldstein (1989: 215-221; 1990). On the surface, therefore, Spanish word-final nasals are alveolar in non-velarizing and velar in velarizing dialects, whether or not a consonant follows.

clusters, however, OCP^{Place} blocks [n.n]. The correct alternative [m.n] is predicted by independently required rankings:

- [ŋ.n] is impossible because the phonemic inventory of Spanish requires the ranking *N/VEL » ONS-IDENT^{Place} » {*N/PAL, *N/LAB, *N/ALV}, which renders [ŋn] highly disharmonic;
- Since UG supplies the stem-level ranking *N/PAL » *N/LAB for free, [m.n] is predictably more harmonic than [ŋ.n].

Of course, tautomorphemic [m.n] must be exempt from word-level velarization: e.g. *colu*[m.n]a 'column', *hi*[m.n]o 'hymn', *sole*[m.n]e 'solemn'; cf. word-level [ŋ.n] from underlying /nn/ in the exceptional word *perenne* 'perennial' (see footnote 27). But even this fact requires no stipulation: the high ranking of IDENT^{labial} at the word level is independently motivated by the need to protect exceptional stems such as *ábu*[m] and *Sia*[m].

In this sense, interleaved OT reveals a beautiful and intricate web of implication. Through level segregation, the theory has access to levels of representation with exactly the right degree of abstractness; it is thereby possible to show that nasals occurring in superficially disparate environments behave uniformly because, at a given level, they all find themselves in the coda and unable to assimilate. In contrast, the macro-constraints [ANCHOR₁₀(Stem, σ , R) \wedge N \rightarrow ALV]_{segment} and [ANCHOR₁₀(Word, σ , R) \wedge N \rightarrow VEL]_{segment} destroy the pattern, for they self-sufficiently determine the place features of resyllabified nasals, dissociating their behaviour from that of nasals in other environments. The facts of Spanish velarizing dialects, though deceptively straightforward, are thus shown not only to be strictly intractable under TCT, but also to pose a formidable challenge to strongly parallel OT in general.

3.4.1.4. Summary of paradigm effects

TCT's attempt at dealing with paradigm effects through OO-correspondence leads, in sum, to unmitigated disaster. Let us briefly review the reasons.

Morphology can cause phonology to misapply because morphological constructions define a hierarchy of phonological domains, where the representations associated with deeply embedded domains provide the input to the application of phonology over higher domains; in this sense, phonological mappings are recursive. Proponents of Strong Parallelism choose to deny this fact. As a result, the input-output relationship between a

complex lexical item and its *i*-base has to be reconstructed as an output-output relationship with some *o*-base. This requires all manner of theoretical contortions:

First, stipulative provisions for base priority are needed, for, unlike *i*-bases, *o*-bases cannot lay claim to the inherent priority of inputs in production. These stipulations drive a wedge between successful applications of Correspondence Theory (such as IO-faithfulness and BR-identity) and anomalous OO-correspondence. The recursion of phonology (i.e. of GEN and EVAL) ends up being aped by means of a duplicated constraint hierarchy (Benua 1997).

Secondly, nothing guarantees that the crucial properties of the *i*-base will surface transparently in any of the candidates to *o*-basehood; on the surface, these properties can be irretrievably masked by word-level or postlexical phonology. Masking places a number of paradigm effects firmly beyond the grasp of TCT. This fact remains however loosely lexical relatedness ('compositionality') is defined, whilst increasing permissiveness in the selection of *o*-bases incurs further problems: OO-correspondence relationships bear little resemblance to word-formation patterns, and levelling is overgenerated on a massive scale. In the event, it is the chance availability of a serviceable *o*-base that decides whether a paradigm effect will be dealt with by TCT or by some other device such as Generalized Alignment. As the case of Spanish velarizing dialects shows, such an opportunistic and atomistic approach to opacity effects is bound to miss important generalizations.

3.4.2 Non-paradigmatic opacity effects: sympathy

3.4.2.1 Sympathy Theory

The opacity effects to which TCT primarily addresses itself are triggered by *overt* morphological constructions. For my present purposes, a morphological construction \mathcal{M} is 'overt' if it involves some phonological operation: e.g. (non-zero) affixation, infixation, truncation, etc. In the usual scenario of paradigmatic overapplication, \mathcal{M} (or a phonological process transparently triggered by \mathcal{M}) destroys the conditioning environment of some phonological generalization \mathcal{P} , which nevertheless applies: e.g. in RP [bæŋə] *banger*, suffixation of *-er* has destroyed the conditioning environment of post-*ŋ* *g*-deletion. In the usual scenario of paradigmatic underapplication, \mathcal{M} (or a phonological process transparently triggered by \mathcal{M}) creates the conditioning environment of \mathcal{P} , which however fails to apply: e.g. in Belfast English [tɛːnəɹ] *tenner*, suffixation of *-er* has created the environment for

pre-rhotic dentalization. In either case, TCT's strategy is to set up an OO-correspondence relationship between the output of \mathcal{M} and some surface form not subject to \mathcal{M} , the *o*-base, to which \mathcal{P} applies normally. As we saw in §3.4.1.3, TCT often fails because no such *o*-base is available.

Additionally, there is an entire class of opacity effects which fall beyond the reach of TCT because they are not triggered by an *overt* morphological construction; these opacity effects are not paradigmatic in the relevant sense. McCarthy (1998) gives an example from Tiberian Hebrew (Malone 1993). Tiberian Hebrew has a process of epenthesis which breaks up consonant clusters in final position (3,43a), as well as a process deleting \mathcal{P} in non-onset position (3,43b). Strikingly, the overapplication of epenthesis in forms such as /deʃʔ/→[deʃe] 'tender grass' (3,43c) shows that \mathcal{P} -deletion counterbleeds epenthesis.

(3,43) Epenthesis and \mathcal{P} -deletion in Tiberian Hebrew (McCarthy 1998: 2)

	(a) 'king'	(b) 'he called'	(c) 'tender grass'
UR	/melk/	/qaraʔ/	/deʃʔ/
Epenthesis	melek	—	deʃeʔ
\mathcal{P} -deletion	—	qara	deʃe
SR ⁵⁶	[melex]	[qara:]	[deʃe]

The survival of the epenthetic vowel in [deʃe] causes opacity, as the conditioning environment of epenthesis, viz. C__C#, has been destroyed. Observe, however, that no overt morphological operation is responsible for this effect: root-final \mathcal{P} has not been resyllabified before a suffix or removed by some morphological process of truncation, but has been deleted by an ordinary phonological rule which is not associated with any particular morpheme. Accordingly, OO-correspondence has nothing to contribute. As McCarthy (1998: 7) points out, there cannot be a transparent *o*-base from which [deʃe] copies its epenthetic vowel: epenthesis is transparent only if underlying \mathcal{P} surfaces as the second C in the frame C__C#, but \mathcal{P} -deletion ensures that \mathcal{P} only occurs in the onset.

⁵⁶ These surface representations bear the effects of phonological processes other than epenthesis and \mathcal{P} -deletion: e.g. spirantization, compensatory lengthening, and pretonic lengthening. For the question whether or not the epenthetic vowel in (3,43c) undergoes compensatory lengthening, see McCarthy (1998: notes 2 and 24).

In this type of situation, interleaved OT adopts a clear strategy: epenthesis is predicted to be active at an earlier phonological level than β -deletion; see tableau (3,44a). This prediction has obvious empirical content and can easily be tested against other facts in the morphology and phonology of Tiberian Hebrew (for relevant references, see Idsardi 1997: 370).

(3,44) (a) Interleaved analysis of Tiberian Hebrew /defʔ/→[deʃe]

LEVEL A					
defʔ	*COMPLEX	MAX ^{Seg}	IDENT ^{α]}	*CODA/?	DEP ^{Seg}
defʔ	*!			*	
deʃ		*!			
de.ʃe		*!			*
deʃ.ʔe			*!		*
de.ʃeʔ				*	*
LEVEL B					
de.ʃeʔ	*CODA/?	IDENT ^{α]}	MAX ^{Seg}		
deʃʔ	*!		*		
deʃ.ʔe		*!	* ⁵⁷		
de.ʃeʔ	*!				
deʃ			**!		
de.ʃe			*		

(b) IDENT^{α]} ⁵⁸

Let α be a segment in the input.

Let β be a correspondent of α in the output.

If α is domain-final, then β is domain-final.

This example highlights a crucial difference between TCT's approach to the phonology-morphology interface and that of interleaved OT. As I suggested in the opening paragraph, TCT can only deal with misapplication effects involving overt morphology. Interleaved OT, in contrast, does not suffer from such limitations because morphological constructions can define phonological domains even if they are non-overt; in other words, an item can undergo the phonology of a level without being subject to any overt morphological operation at that level. The behaviour of nasals in Spanish velarizing dialects again provides a particularly good example; see diagram (3,41a). The word-formation process which derives the noun stem *desdén* from the root /desdep/ is phonologically covert, as the stem is assigned to the athematic noun class; cf. the adjective *desdeñoso*, created by means of the derivational suffix *-os-*, and the verb *desdeñar*, which takes the first conjugation marker *-a-* (though see footnote 21). Nonetheless, this covert morphological operation of noun stem formation defines a stem-level phonological domain, and so underlying /ɲ/ undergoes depalatalization: [_{P-stem}des.'den]. Similarly, the stem [_Ndes.'den] must be assigned inflectional —more specifically, number— features if it is to become a syntactically free lexical item, i.e. a word (see §3.3.1.1).⁵⁹ Again, the assignment of singular number is a covert operation involving no affixation; cf. the plural, which takes the inflectional suffix *-(e)s*. In either case, however, the assignment of number features defines a word-level phonological domain; as a result, the unsuffixed singular form undergoes velarization: [_{P-word}des.'den]. In sum, an item can be subject to the phonological function

⁵⁸ In this analysis, IDENT^{α]} replaces McCarthy's (1998: 21) ANCHOR₀(Root, σ, Final). IDENT^{α]} avoids direct reference to the root as a phonological domain; for the issue of access to morphologically labelled boundaries, see §3.4.1.1. In his analysis of Arabic syllabification, Kiparsky (1998a: 91) uses the following constraint:

Align-0

The right edge of the phonological word coincides with the right edge of the prosodic word. No deletion or epenthesis in word-final position.

⁵⁹ In Spanish, as in Latin, gender is an inherent property of the stem, even though there is a strong correlation between gender features and membership of certain nominal classes: e.g. membership of the *-o-* class is strongly correlated with masculine gender, but cf. *mano* fem. 'hand'; see J. W. Harris (1991a, b) and Aronoff (1994). The athematic noun stem *desdén* is masculine.

⁵⁷ Assuming syncope plus paragoge, rather than metathesis.

associated with a level L even if it does not undergo overt morphological processing at L ; analyses such as (3,44) depend on this result.⁶⁰

The type of opacity effect represented by the overapplication of epenthesis in Tiberian Hebrew proves intractable not only to TCT, but also to Generalized Alignment. Strongly parallel OT must, as a result, adopt theoretical provisions specifically designed to deal with the problem. Recall now that the programmatic assumption of strongly parallel OT is that misapplication effects arise because the output participates in transderivational correspondence relationships, each of which is monitored in parallel by an appropriate set of constraints modelled on the schemas of Correspondence Theory (see §3.2.2.1). Since the opaque phenomena at hand are not triggered by identity with an o -base, a different type of correspondence relationship must be involved; according to McCarthy (1998), this novel correspondence relationship is 'sympathy'.

In McCarthy's proposal, sympathetic correspondence connects the output to a virtual phonological representation, known as the 'sympathy candidate' or '☉-candidate'. The latter is a designated member of the set of output-candidates produced by GEN; i.e. it is a co-candidate of the actual output. This particular suboptimal output-candidate is endowed with special theoretical status on the grounds that it does not violate a designated subordinate constraint: the 'selector constraint' or '★-constraint'; more specifically, the ☉-candidate is the most harmonic member of the set of output candidates which best-satisfies the ★-constraint. Thus, the ☉-candidate under hierarchy (3,45a) is the optimal candidate relative to (3,45b); see (Itô & Mester 1997b: 10).

- (3,45) (a) $C_a \gg \dots \gg C_i \gg \star\text{-}C \gg C_j \gg \dots \gg C_z$
 (b) $\star\text{-}C \gg C_a \gg \dots \gg C_i \gg C_j \gg \dots \gg C_z$

In McCarthy's notation, the ☉-candidate selected by a constraint C within a given hierarchy is referred to as \aleph_C . A set of specific correspondence constraints, known of as ☉O-

⁶⁰ The line of reasoning developed in this paragraph is in broad agreement with Inkelas & Orgun's (1995) principle of Level Economy, which states that a form is only subject to the phonology of those levels at which it undergoes morphological operations. The vital proviso would be that *covert* morphology suffices to trigger phonological processing. The precise empirical predictions of this approach (including whether or not domain prespecification is required) will largely depend on morphosyntactic theory, and particularly on where the line is drawn between obligatory and non-obligatory morphosyntactic features. See also footnote 13 for some references to recent work on derived environment effects.

faithfulness constraints, monitor the relationship between the output and the ☉-candidate; the former is required to be faithful to the latter.

McCarthy's proposal incorporates two elaborations of this basic idea:

(i) First, an input may enter into a sympathetic relationship with an indefinite number of ☉-candidates: i.e. given two different ★-constraints A and B , there will be two parallel ☉-candidates: \aleph_A and \aleph_B . The faithfulness of the output to each ☉-candidate is monitored in parallel by a separate set of independently ranked ☉O-constraints: e.g. \aleph_A O-FAITH and \aleph_B O-FAITH.⁶¹ This idea is analogous to Benua's suggestion that a language may divide its affixes into any number of OO-correspondence classes, where each class is monitored by its own specific set of OO-identity constraints (see §3.3.1.2).

(ii) Secondly, McCarthy stipulates that the ★-constraint must be an IO-faithfulness constraint. This assumption plays a vital rôle both in constraining the generative power of sympathy and securing its conceptual underpinnings. It has nonetheless come under fire (Itô & Mester 1997b, de Lacy 1998), and my analysis of West Germanic Gemination in §3.5. will show it to be untenable. The implications are discussed in §3.4.2.2.

High ranking ☉O-faithfulness constraints produce misapplication in a similar way to OO-identity constraints. In the case of Tiberian Hebrew /deʃʔ/, the epenthetic vowel of the output form [deʃe] incurs a violation of DEP_{IO}^V which is not required to satisfy any higher-ranking markedness or IO-faithfulness constraint; indeed, the markedness and IO-faithfulness violations of the unepenthesized transparent candidate *[deʃ] are a proper subset of those of the opaque output [deʃe] (Itô & Mester 1997b: 2). The sympathy analysis suggests that [deʃe] copies the epenthetic vowel from the ☉-candidate [deʃeʔ], which is the most harmonic candidate retaining underlying ʔ, i.e. $\aleph_{\text{MAX}_{\text{IO-C}}}$,⁶² the transfer is effected by high-ranking $\text{MAX}_{\text{☉O}}^V$. In [deʃeʔ], the epenthetic vowel is motivated by the need to fulfil high-ranking *COMPLEX. However, the ☉-candidate cannot itself be the most

⁶¹ Kiparsky (1998a: 74) speculates (but does not prove) that countenancing multiple ☉-candidates is likely to confer upon Sympathy Theory the ability to replicate rule-ordering analyses of unlimited depth, leading to overgeneration of opacity; cf. §3.3.3.1.

⁶² McCarthy (1998: 23) acknowledges that, in actual fact, the choice of ☉-candidate in (3,46) is indeterminate between [de.ʃeʔ] and [deʃ.ʔe]: $\aleph_{\text{MAX}_{\text{IO-C}}}$ is [de.ʃeʔ] under the ranking $\text{ANCHOR}_{\text{IO}}(\text{Root}, \sigma, \text{Final}) \gg \text{*CODA} \uparrow$, and [deʃ.ʔe] under the opposite ranking. The relative position of these two constraints in the hierarchy of Tiberian Hebrew is not independently known. Either ☉-candidate gives the right result, because both contain an epenthetic vowel which $\text{MAX}_{\text{☉O}}^V$ transfers to the optimal output.

harmonic output form because *CODA/? outranks ★-MAX₁₀^C; witness /qara?/→{qara:} (3,43b).⁶³

(3,46) Tiberian Hebrew /def?/→{deje} in Sympathy Theory (McCarthy 1998: 23)

/def?/	MAX _{OO} ^V	*COMPLEX	ANCHOR ₁₀ (Root,σ,R)	*CODA/?	★-MAX ₁₀ ^C	DEP ₁₀ ^V
def?	*!	*!		*!	✓	
def.?e ⁶²			*!		✓	*
de.je? *				*!	✓	*
def ⁶⁴	*!				*	
de.je *					*	*

Sympathy Theory, as propounded by McCarthy (1998), suffers from a number of serious conceptual difficulties. To begin with, the selection of the *O-candidate requires that the following assumption should be made:

(3,47) Invisibility of *O-faithfulness constraints (McCarthy 1998: 19)

Selection of *O-candidates is done without reference to *O-faithfulness constraints (on *any* sympathetic correspondence relation).

The import of this statement can be clarified by reference to (3,45). Recall that the *O-candidate under hierarchy (3,45a) was defined as the optimal candidate selected by (3,45b). The invisibility clause in (3,47) states that this definition is correct only if all correspondence constraints taking the *O-candidate as an argument are removed from (3,45b). In (3,46), for example, the *O-candidate is the optimal representation selected by

⁶³ Similarly, the alternative *O-candidate [def.?e] (see previous footnote) is ungrammatical as an output form because ANCHOR₁₀(Root, σ, Final) dominates ★-MAX₁₀^C.

⁶⁴ Following McCarthy (1998), transparent losing candidates are indicated in tableaux by a hand pointing away.

the hierarchy ★-MAX₁₀^C » { *COMPLEX, ANCHOR₁₀(Root, σ, Final), *CODA/? } » DEP₁₀^V; the *O-faithfulness constraint MAX_{OO}^V is invisible for the purposes of *O-selection.

According to McCarthy (1998: 19), the invisibility stipulation is necessary to prevent 'infinite loops': "the choice of N_{Fi} can't depend on performance on a constraint that needs to know what N_{Fi} is in order to be evaluated." The argument seems to be, in other words, that *O-faithfulness constraints cannot participate in the selection of the *O-candidate because evaluation of *O-faithfulness presupposes knowing what the *O-candidate is. This line of reasoning strikes me as incorrect. Under Sympathy Theory, EVAL chooses among pairs of representations, where one member aspires to *O-status and the other aspires to *F-status, with the proviso that the member aspiring to *O-status must best-satisfy the ★-constraint. Those pairs of representations whose members fail to be identical in respect of property *P* will fail FAITH_{OO}^F; and that is all. The situation is analogous in all relevant respects with that of symmetrical TCT, where, to the best of my knowledge, nobody has ever suggested that visibility of the base to OO-identity constraints leads to infinite loops; see e.g. tableau (3,31).

McCarthy's discussion of infinite loops diverts our attention from a more fundamental rôle which (3,47) is designed to play. Like OO-identity, and unlike BR-identity, the correspondence relationship between the *O-candidate and the output is *asymmetrical*. *O-faithfulness constraints are designed to induce misapplication in the output, not in the *O-candidate; in this sense, the invisibility stipulation in (3,47) guarantees the 'priority of the *O-candidate' (see §3.4.1.2). It is obviously undesirable that identity between the *O-candidate and the output should be attained through misapplication in the former; this would be empirically equivalent to normal application, for the *O-candidate is a virtual form. Yet, as in the case of OO-identity constraints, the priority of the *O-candidate cannot be directly read off the formulation of *O-faithfulness constraints. Consider, for example, MAX_{OO}^V:

(3,48) MAX_{OO}^V

Every vowel in the *O-candidate has a correspondent in the output.

If under normal application a vowel *V* is present in the *O-candidate but absent in the output, compliance with MAX_{OO}^V could be brought about in either of two ways, viz. by

deleting V in the ☉-candidate or by inserting V in the output; see (3,34). The invisibility clause evades this danger by guaranteeing that the content of the ☉-candidate is not affected by the sympathetic relationship.

In fact, Sympathy Theory needs the invisibility clause to be able to generate underapplication. If the relationship between the ☉-candidate and the ☞-candidate were symmetrical, then overapplication in the ☉-candidate would always be preferable to underapplication in the ☞-candidate. This is the same problem that affected symmetrical TCT; see diagram (3,32) above. To illustrate this undesirable prediction, I shall use one of McCarthy's own examples. According to McCarthy, Bedouin Arabic has a process of glide vocalization which counterfeeds the raising of /a/ in open syllables; see (3,49b).

(3,49) /a/-raising and glide vocalization in Bedouin Arabic (McCarthy 1998: 3)

	(a) 'he wrote'	(b) 'Bedouin'
UR	/katab/	/badw/
Raising	ki.tab	—
Vocalization	—	ba.du

McCarthy (1998: 25) proposes the following analysis:

$$(3,50) \{ * \text{COMPLEX}, \text{IDENT}_{\text{☉O}}^{[\text{high}]} \} \gg *a_{\text{c}} \gg \{ \text{IDENT}_{\text{IO}}^{[\text{high}]}, \star\text{-IDENT}_{\text{IO}}^{\text{Nuc}} \}^{65}$$

Assuming ☉-candidates to have priority over ☞-candidates, this hierarchy yields ☉[badw]~☞[ba.du], with the output opaquely matching the vowel height of the ☉-candidate. But it is easy to show that, without recourse to the invisibility clause, (3,50) would predict overapplication of raising in the ☉-candidate, rather than underapplication in the ☞-candidate. In fact, no ranking of $\text{IDENT}_{\text{☉O}}^{[\text{high}]}$ can trigger underapplication under conditions of symmetry. Demotion of $\text{IDENT}_{\text{☉O}}^{[\text{high}]}$ relative to $*a_{\text{c}}$ (yielding $*a_{\text{c}} \gg$

⁶⁵ In McCarthy's analysis, the \star -constraint is in fact $\text{DEP}_{\text{IO}}^{\text{M}}$. Its function is to block the vocalization of underlyingly non-moraic vocoids by preventing them from projecting a mora. I have substituted $\text{IDENT}_{\text{IO}}^{\text{Nuc}}$ because my own version of $\text{DEP}_{\text{IO}}^{\text{M}}$ is not designed to block glide vocalization; see footnote 23 in chapter 2. $\text{IDENT}_{\text{IO}}^{\text{Nuc}}$ states that an output segment cannot occupy the syllable nucleus if its input correspondent is non-nuclear. Crucially, I assume that the input contains partially specified syllable structure, including σ nodes; see chapter 2, particularly §2.2.2.1 and §2.2.2.3. For a related proposal, see Roca (1997).

$\text{IDENT}_{\text{☉O}}^{[\text{high}]} \gg \text{IDENT}_{\text{IO}}^{[\text{high}]}$ still selects ☉[bidw]~☞[bi.du]. Demotion of $\text{IDENT}_{\text{☉O}}^{[\text{high}]}$ relative to $\text{IDENT}_{\text{IO}}^{[\text{high}]}$ (yielding $*a_{\text{c}} \gg \text{IDENT}_{\text{IO}}^{[\text{high}]} \gg \text{IDENT}_{\text{☉O}}^{[\text{high}]}$) would in turn simply lead to unfaithfulness between the output and the ☉-candidate, with normal application of raising in both: i.e. ☉[badw]~☞[bi.du].

(3,51) Overapplication in the ☉-candidate under symmetrical sympathy

/badw/		*COMPLEX	$\text{IDENT}_{\text{☉O}}^{[\text{high}]}$	$*a_{\text{c}}$	$\text{IDENT}_{\text{IO}}^{[\text{high}]}$	$\star\text{IDENT}_{\text{IO}}^{\text{Nuc}}$
☉	☞					
bidw ☉*	bi.du ☉*	(*)			*,*	*
badw ☉	ba.du ☉	(*)		i*!		*
badw ☞	bi.du ☞	(*)	*!		*	*
bidw	ba.du	(*)	*!	*	*	*
bidw	bidw	(*), *!			*,*	
bidw	badw	(*), *!	*!		*	
badw	bidw	(*), *!	*!		*	
badw	badw	(*), *!				
bi.du	bi.du				**,*	**,*
bi.du	ba.du		*	*	*	**,*
bi.du	bidw	*			**,*	*
bi.du	badw	*	*		*	*
ba.du	bi.du		*	*	*	**,*
ba.du	ba.du			**		**,*
ba.du	bidw	*	*		*	*
ba.du	badw	*				*

The notation of (3,51) may require some clarification. In this tableau, shading signals the disqualification of all the candidate pairs in which the aspirant to \otimes -status violates the \star -constraint. This procedure implements the defining principle of Sympathy Theory: Confinement to $C_{\langle F \rangle}$.

(3,52) Confinement to $C_{\langle F \rangle}$ (McCarthy 1998: 19)

Selection of the \otimes -candidate N_F is confined to $C_{\langle F \rangle}$, the set of candidates that obey the IO faithfulness constraint F .

In all the remaining candidate pairs, the aspirant to \otimes -status violates \star COMPLEX in order to satisfy \star -IDENT_{IO}^{Nuc}. These violations all cancel each other out; their cancellation is signalled by enclosing the relevant marks between parentheses. By the Cancellation Lemma (Prince & Smolensky 1993: 139), cancelled marks are disregarded. Where a cell contains two marks separated by a comma, one is incurred by the aspirant to \otimes -status, and the other by the aspirant to \equiv -status. The pair which counterfactually wins is indicated by the \odot symbol; the pair which ought to have won (i.e. the one which wins under the invisibility clause) is marked by the \ominus symbol; the transparent pair is marked by the hands pointing away. The deciding mark appears as $\dagger!$.

We are thus back in familiar territory. Once more, the problem of asymmetry arises as strongly parallel OT attempts to substitute (quasi-)output-output correspondence for an input-output relationship. In §3.4.1.2 we saw how TCT had to resort to *ad hoc* stipulations to establish the priority of *o*-bases —which, unlike *i*-bases, could not lay claim to the inherent priority of inputs over outputs in production. In a similar vein, the \otimes -candidate performs the same rôle in Sympathy Theory as an intermediate representation in SPE, or as the output of a non-final level/cycle in interleaved OT —both of which have inherent priority over the surface representation. The \otimes -candidate, however, is a (virtual) output form, selected in parallel with the \equiv -candidate by reference to a common underlying representation; it therefore has no inherent priority when it comes to determining the outcome of misapplication. Rather, the asymmetry of the \otimes O-relationship has to be stipulated by means of (3,47). The fact that an asymmetry clause of this nature is unnecessary, indeed undesirable, in BR-identity highlights the anomalous nature of \otimes O-faithfulness.

McCarthy (1998: 20) attempts to forestall this criticism. The tenor of his comments is revealing enough to justify a somewhat lengthy quotation:

It is sometimes suggested that sympathy covertly reintroduces a kind of serialism. According to this view, selection of the \otimes -candidates must take place prior to selection of the \equiv -candidate, because the latter depends on the former. (The Invisibility property [3,47] would be seen as a necessary consequence of this ordering of events.) But it is simple wrong to insist that “A depends on properties of B” necessarily implies that “there is a serial derivation in which B is constructed earlier than A”. Dependencies of one form on another can also be understood in terms of satisfaction of constraints in parallel rather than serially. For example, reduplication may involve copying the base as it has been altered by phonological processes, but this does not entail that the base undergo phonology prior to reduplication. Rather, the effects of phonology on the base and reduplicant can be determined together, in parallel [...]. (Similar remarks apply to the analysis of “cyclic” effects in Benua (1997).) In both reduplication and sympathy, correspondence provides a way to express dependencies that does not depend on serial derivation.

This pre-emptive response fails to dispose of the main objection. It is certainly not in question that dependencies between the SR and representations other than the UR can be modelled using strongly parallel means. The crux of the matter is, however, that their asymmetry follows naturally from the architecture of so-called ‘serial’ models;⁶⁶ strongly parallel OT, in contrast, can only establish asymmetry by means of *ad hoc* stipulations, such as \mathcal{M} -recursion and the \otimes -faithfulness invisibility clause. Insofar as this sort of provision is out of place in BR-identity, both TCT and Sympathy Theory become suspect of deploying the resources of Correspondence Theory *ultra vires*. More generally, a research programme becomes suspect of having entered a ‘degenerating problemshift’ (Lakatos 1970) when it needs *ad hoc* stipulations to accommodate previously well-known facts, particularly if a competing research programme (in this case, interleaved OT) deduces the same facts from first principles.

Sympathy Theory faces further conceptual embarrassment over the arbitrariness of \otimes -candidates. As I have repeatedly stated in the course of this chapter, all the phonological representations postulated by interleaved OT are independently motivated in morphological or syntactic terms; see the discussion of Spanish *desdén* above, as well as §3.3.1.3 and Orgun (1996b: §1.4.4). In a derivational interpretation of interleaving, the output of each

⁶⁶ See §3.3.1.3 again for the issue of derivational vs declarative interpretations of interleaving.

level or cycle corresponds to the application of phonology over a domain defined by a certain set of morphological or syntactic operations; in Orgun's declarative interpretation, each non-surface phonological representation is associated with some node in the grammatical constituent structure of the relevant linguistic expression. In Sympathy Theory, in contrast, \otimes -candidates have no morphosyntactic status; they lack correlates outside the phonology. Very often, the properties of the \otimes -candidate coincide with those of a non-surface representation in interleaved OT;⁶⁷ nonetheless, Sympathy Theory cannot establish a formal link with the relevant grammatical domain, as Strong Parallelism acknowledges no domain for the application of phonology other than the entire underlying representation submitted to GEN. In this sense, \otimes -candidates are rather like the intermediate representations of SPE; they are *ad hoc* entities with no *raison d'être* other than to help strongly parallel OT deal with recalcitrant opacity effects. In comparison with interleaved OT, therefore, Sympathy Theory takes an alarmingly retrogressive step: it rejects level segregation and cyclicity (which provide a consistent model of the phonology-morphosyntax interface), and adopts instead an arbitrary device analogous to rule ordering.

McCarthy (1998: 11) nevertheless describes the status of \otimes -candidates in OT as 'legitimate' because they are independently required as members of G_i , where G_i designates the set of output candidates constructed by GEN for an input i , i.e. $\text{GEN}(i) \rightarrow G_i$. In the current state of knowledge, such a theory-internal argument amounts to very little, for G_i is usually characterized as being of indefinite size, and indeed potentially infinite. It is not clear, therefore, that G_i is smaller than the set of all possible phonological representations under a given theory of representations. By the same token, if all members of G_i can lay claim to theoretical status, then absolutely *any* representation can in principle be so legitimized: the trick lies in concocting a selection device that picks out the desired representation. In other words, any theory which relies on recruiting members of G_i as transderivational correspondents is no more constrained than the selection device on which it relies. McCarthy (1998) takes decisive steps to constrain his selection procedure. The \otimes -candidate is chosen by a special constraint hierarchy (in effect, a cophonology) which

⁶⁷ In footnote 54, for example, the \otimes -candidate [tʁɛŋ.a.θul] is identical in all relevant respects with the output of the word level in an interleaved analysis (i.e. the 'lexical representation' of Mohanan 1982, 1986: ch. 7). Note, in particular, the application of default velarization and the absence of postlexical resyllabification. McCarthy (1998: 30) observes that, "Though convergence between the \otimes -candidate and the serialist's intermediate stage is a consistent finding with simple opaque interactions [i.e. those involving a single \otimes -candidate], it is neither expected nor observed in situations of multiple opacity."

differs from the main constraint hierarchy in just two respects: (i) the promotion of a low-ranking IO-faithfulness constraint to absolute superordinate position, and (ii) the exclusion of all \otimes -faithfulness constraints. In the next section we shall see, however, that the confinement of \star -constraints to the IO-faithfulness set is untenable. This has dire consequences for Sympathy Theory, and for strongly parallel OT in general.

3.4.2.2 \star -Confinement and extended sympathy

McCarthy's (1998) version of Sympathy Theory incorporates the following principle:

(3,53) \star -Confinement

\star -C \in FAITH_{IO}

The constraints which select \otimes -candidates must be drawn from the set of IO-faithfulness constraints.

By dramatically restricting the choice of \otimes -candidates in any given grammar, (3,53) goes a long way towards curbing the generative power of sympathetic correspondence; but, as I shall presently show, this hypothesis also underpins the theory in other significant ways. First, \star -Confinement provides the otherwise unmotivated notion of sympathy with functional grounding in performance factors, specifically lexical access. Secondly, the exclusion of markedness constraints from \otimes -candidate selection safeguards the ban which strongly parallel OT is credited with imposing upon synchronic markedness reversal (see §3.3.3.3).

Under \star -Confinement, the \otimes -candidate is selected by a low-ranking, crucially dominated IO-faithfulness constraint.⁶⁸ As McCarthy notes, this implies that the \otimes -candidate will always be more faithful to the input than the \otimes -candidate in respect of the phonological property assessed by the \star -constraint. In turn, active \otimes O-faithfulness constraints demand that the output should be faithful to the \otimes -candidate along some dimension, thereby prompting otherwise unmotivated violations of markedness and/or IO-faithfulness constraints. These unexpected violations can be regarded as warning flags,

⁶⁸ As McCarthy (1998: 36) observes, if the \star -constraint is superordinate, then the \otimes -candidate is vacuously identical with the actual output form, for the \otimes -candidate will be the most harmonic member of the set of candidates satisfying the \star -constraint (which, in effect, is the definition of the \otimes -candidate). In other words, \otimes O-correspondence triggers misapplication only if the \star -constraint is crucially dominated.

drawing attention to the fact that an infraction of the \star -constraint (involving unfaithfulness to the input) has occurred in the output. From this, McCarthy (1998: 13) concludes that sympathy facilitates lexical access in performance:

The recoverability of underlying forms is supported when faithfulness constraints are obeyed. Sympathy provides another channel of faithfulness alongside the standard one — the \otimes -candidate is chosen because it obeys a specified faithfulness constraint, and the output is compelled to resemble (i.e., be faithful to) the \otimes -candidate. In this way, a sympathetic effect on the input \rightarrow output mapping indirectly improves the recoverability of the input from the output.

McCarthy traces this idea back to Kaye's (1974, 1975) suggestion that certain opaque rule orderings aid parsing.

As an example McCarthy gives the underapplication of nasal harmony in Sea Dayak (Kenstowicz & Kisseberth 1979; Scott 1957). In this language, nasality spreads from left to right, unless blocked by oral consonants. There is, however, an optional process of oral consonant deletion which counterfeeds nasal harmony.

(3,54) Deletion counterfeeds nasal harmony in Sea Dayak (McCarthy 1998: 4)

	'straighten'	'set up ladder'
UR	/naŋa/	/naŋga/
Harmony	nãŋã	nãŋga
Deletion (optional)	—	nãŋga ~ nãŋa
SR	[nãŋã?]	[nãŋga?] ~ [nãŋa?], *[nãŋã?]

In this case, the absence of nasality in the second vowel of [nãŋã?] 'set up a ladder' signals the deletion of an underlying oral consonant. As a result, /naŋga/ escapes merger with /naŋa/; the phonemic contrast between /g/ and /Ø/ surfaces as one between normal application and underapplication of nasal harmony. In a sympathy-theoretic analysis, ungrammatical *[nãŋã?] 'set up a ladder' is ruled out by IDENT_{*O}^[nasal], which demands faithfulness to the nasality specifications of the \otimes -candidate [nãŋga?], selected by \star -MAX_{IO}^C (McCarthy 1998: 15).

In other cases, however, this functional account of sympathetic effects is much less persuasive, particularly when \otimes O-identity directly conflicts with IO-faithfulness. Return, for example, to the sympathy analysis of Tiberian Hebrew /deʃʔ/ \rightarrow [deʃe] in (3,46). Here, the superiority of \otimes [de.ʃe] over \otimes [deʃ] depends on the ranking of MAX_{*O}^V above DEP_{IO}^V: the former requires that the epenthetic vowel present in \otimes [de.ʃe?] should be copied, even at the expense of unfaithfulness to the input /deʃʔ/. The opaque winner and the transparent loser incur otherwise identical sets of marks. In this case, sympathy does not diminish but actually increases IO-unfaithfulness: vowel epenthesis is opaquely added to ʔ-deletion. It could be argued that the unmotivated presence of an epenthetic vowel flags the deletion of /ʔ/. For the listener, however, perceiving the vowel as a warning flag (associated with an unmotivated violation of DEP_{IO}^V) implies recognising its epenthetic status, which in turn implies already having access to the target input representation. In this sense, the overapplication of epenthesis in Tiberian Hebrew can hardly be claimed to facilitate lexical access in the same way as the underapplication of nasal harmony in Sea Dayak does; at best, the epenthetic vowel could be regarded as a redundant cue for input-output matching, providing a somewhat circuitous means to confirm that /ʔ/-deletion has taken place.

Be that as it may, McCarthy's claims are representative of a consistent attempt within strongly parallel OT to provide a functional explanation for opacity. In the case of paradigm effects, Kenstowicz (1996: 370) suggested in a similar vein that, "Observance of Base-Identity serves to improve the transparency of morphological relationships between words and thus may enhance lexical access." Whilst markedness constraints are usually regarded as grounded on the physics or the physiology of speech production or perception, transderivational correspondence would therefore be grounded on considerations of lexical recognition. In this respect, strongly parallel OT emerges as a markedly functionalist or adaptationist research programme.⁶⁹

⁶⁹ For the concept of 'grounding', see Archangeli & Pulleyblank (1994); see also Hayes (1996) and Myers (1997) for non-innatist views of CON. In response to the argument that numerous phonological phenomena have no obvious functional motivation (Dinnsen 1980, Anderson 1981), Myers (1997: 147) suggests that this situation merely reflects our current ignorance, and that a functionalist research strategy remains advisable. I would suggest, however, that this methodological prescription is misleading and encourages the 'Panglossian fallacy' (Gould & Lewontin 1979). Research in evolutionary biology indicates that the evolved properties of organisms need not always be adaptations (i.e. properties selected for on fitness grounds), but may be trade-offs for adaptive properties, or may simply reflect basic engineering constraints. (Susceptibility to disease provides an obvious illustration of these possibilities; see Nesse & Williams 1998.) The same idea should in principle apply to the properties of the language faculty (see e.g. Lightfoot 1999: ch. 9). In interleaved OT, the viability of opaque phonological generalizations is treated as a (possibly maladaptive) epiphenomenon of the design of the phonology-morphosyntax interface.

★-Confinement also plays a vital rôle in relation to the issue of synchronic markedness reversal. As I pointed out in §3.3.3.3, Benua (1997) claims that phonological analyses relying on stratification need never resort to the reranking of markedness constraints between levels; the markedness hierarchy remains allegedly invariant across cophonologies. Benua argues that interleaved OT can accommodate this restriction only in a stipulative manner, whilst strongly parallel OT deduces it from first principles. In TCT, (morphological) levels are replaced by correspondence classes, associated with specific sets of OO-identity constraints; classes differ merely in the ranking of their respective OO-identity constraints. There is, in contrast, a single set of markedness constraints, whose relative ordering suffers no alteration. As a result, all output forms comply with the same markedness hierarchy, although they may be subject to different transderivational identity requirements.

The advent of Sympathy Theory alters this picture significantly. An output form and its *o*-base are different because they are associated with different input representations. In contrast, the \otimes -candidate is by definition derived from the same input representation as the \ominus -candidate; it is different from the latter only because it is selected by a different constraint hierarchy, more specifically one in which the ★-constraint has been raised from its lowly station —see again (3,45). In effect, Sympathy Theory institutes a form of synchronic constraint reranking. In this context, it is no longer clear that the ban on synchronic *markedness* reversal is independently motivated, for it requires a stipulation such as (3,53) —which McCarthy writes into (3,52). The question would be whether or not (3,53) does some other unrelated work in the grammar.⁷⁰

In this connection, ★-Confinement is representative of a theoretical outlook where it is regarded as incumbent upon UG to legislate against cophonology divergence (and, more generally, against abstractness). Following Orgun (1996b: 122), I have argued, in contrast, that the empirical limitations of cophonology divergence are an epiphenomenon

⁷⁰ McCarthy (1998: §5) claims that, under ★-Confinement, Sympathy Theory cannot model the effects of a Duke-of-York gambit (Pullum 1976). He argues, moreover, that this is a desirable result, suggesting that the instances of Duke-of-York derivation described in the literature are not convincing. Kiparsky's (1998a: §3.6.1) response to these claims is twofold. First, he demonstrates that, contrary to McCarthy's claims, Sympathy Theory can indeed mimic Duke-of-York derivations, even under ★-Confinement. Secondly, he argues that UG should not formally ban Duke-of-York derivations as a class; there is no principled reason why robustly learnable Duke-of-York gambits should be excluded. McCarthy (to appear) deals with the first objection; unfortunately, I was unable to obtain a copy of this manuscript in time to incorporate its contents into this discussion.

of acquisition and history: by the 'principle of free reranking', the constraint hierarchies of different levels are formally independent (§3.3.3.2).

As the preceding paragraphs indicate, ★-Confinement is endowed with the outmost theoretical importance: not only does it significantly constrain the generative power of Sympathy Theory, but it also vitally underpins key objectives of the strongly parallel programme. Nonetheless, research within strongly parallel OT has quickly stumbled upon putative counterexamples to ★-Confinement, leading to the appearance of 'Extended Sympathy Theory':

(3,55) **Extended sympathy (adapted from Itô & Mester 1997b)**

Other types of constraints, beside IO-faithfulness, can serve as ★-c.

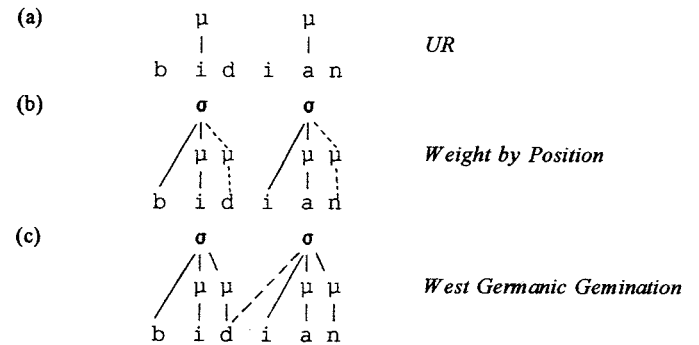
(i) German has a productive process of hypocoristic formation involving truncation of the base followed by suffixation of *-i*: e.g. *Gabriele* ~ *Gabi* (first name), *Gorbatschow* ~ *Gorbi* (surname), *Ostdeutscher* ~ *Ossi* 'East German' (Féry 1997). According to Itô & Mester (1997b), who follow Neef (1996: 282-283), the bare truncatum (i.e. the shortened form prior to *-i* suffixation) consists of the maximal syllable extractable from the base: hence **Gabr-i* (**Gabr-* is an ill-formed monosyllable) and **Gor-i* (*Gorb-* accommodates more base material). In Itô & Mester's strongly parallel analysis, the \otimes -candidate assumes the rôle of the bare truncatum; the actual output copies its segmental content through \otimes O-faithfulness. Crucially, the \otimes -candidate is itself defined as $\mathfrak{N}_{\text{ALIGN}(\sigma, L; \omega, L)}$, i.e. as the candidate which accommodates the maximum number of base segments within a single left-aligned syllable. To the extent that ★-ALIGN($\sigma, L; \omega, L$) is a markedness constraint, Itô & Mester's proposal is incompatible with ★-Confinement.

(ii) Certain stress systems (dubbed 'count systems' in van der Hulst 1994: §4) appear to require exhaustive footing in order to determine the location of primary stress, even though surface forms bear no secondary stress: e.g. Cairene Classical Arabic, Creek. In serial analyses, iterative footing is followed by a 'conflation' operation deleting the grid line which encodes secondary stress (Halle & Vergnaud 1987b: 50-55; Halle & Kenstowicz 1991: 462; see also footnote 29 above). In de Lacy's (1998) sympathy-theoretic analysis of Cairene Classical Arabic, IDENT \otimes ^{StressPeak} (see (3,23) above) determines the location of primary stress by reference to an exhaustively footed \otimes -candidate selected by ★-PARSE^o (see (2,5c) above). Again, this contradicts ★-Confinement.

Unfortunately, neither Itô & Mester's nor de Lacy's evidence establishes the untenability of \star -Confinement with absolute certainty. First, Féry (1999) denies that the bare truncatum in German *-i* hypocoristics must always be the maximal monosyllable extractable from the base: this would incorrectly predict *Ostdeutscher* ~ **Osti*, as *Ost* 'East' is evidently a well-formed monosyllable (cf. also *Manfred* ~ *Manni*, not **Manfi*). Instead, Féry adopts an ordinary OO-identity approach to truncation; she suggests that the composition of the consonant clusters preceding the suffix *-i* manifests Emergence-of-the-Unmarked effects (McCarthy & Prince 1994). Secondly, the opacity of 'count systems' is not universally acknowledged. In the case of Creek, it has been argued that weak feet are indeed present in the surface representation (Martin 1992, cited by Hayes 1995: 119). In other cases, representational alternatives such as headless or phonetically uninterpreted weak feet are difficult to discard (see de Lacy 1998: §4 for references).

In the following section, I shall present new empirical evidence against \star -Confinement. In West Germanic, underlyingly short consonants geminated before [j] and before the liquids /r l/ under conditions to be specified below: e.g. /bid-I-an/ → [bid.djan] 'ask' (OS *bidden*, OE *biddan*, OHG *bitten*). The historical record proves conclusively that this lengthening was an opaque form of mora preservation. In Proto-Germanic, the consonant clusters in question were heterosyllabic; accordingly, the first consonant was syllabified in the coda, where it became positionally μ -licensed (see §2.3): e.g. [bid^d.jan] (Gothic *bidjan*). In West Germanic, syllable contact was optimized by incorporating such coda consonants into the following onset; violations of CONTACT (2,29) were thereby avoided. Crucially, the erstwhile short coda consonants retained their moraic attachment, despite no longer fulfilling the conditions for positional μ -licensing: e.g. [bid.jan] > [bid.djan]. In more conventional terms, West Germanic Gemination counterbleeds Weight by Position (2,24a):

(3,56) West Germanic Gemination counterbleeds Weight by Position



In a sympathy-theoretic analysis, high-ranking IDENT_{IO}^μ forces \star [bid.djan] (=3,56c) to preserve the mora count of \star [bid.jan] (=3,56b). This results in opaque violations of DEP_{IO}^μ and IDENT_{IO}^μ, for the mora associated with the geminate [d] has no correspondent in the input (=3,56a), nor does it fulfil the criteria for positional μ -licensing; in particular, it violates clause (2,40ii). In contrast, the transparent loser \star [bi.djan], rejected by IDENT_{IO}^μ, fulfils both CONTACT and moraic IO-faithfulness. The crucial fact is that \star [bid.jan] is not selected by an IO-faithfulness constraint, but by \star -*[_oCj], which is the low-ranking markedness constraint banning Cj onset clusters. This constraint had been superordinate in Proto-Germanic, but was now dominated by antagonistic CONTACT. Thus, West Germanic Gemination confirms the untenability of \star -Confinement.

Curiously enough, Itô & Mester (1997b) and de Lacy (1998) are rather sanguine about the prospects of 'extended sympathy', in stark contrast with McCarthy's (1998: §6) misgivings. Both Itô & Mester and de Lacy regard \star -Confinement as an *ad hoc* stipulation (which, in formal terms, it most certainly is); by the same token, discarding (3,53) leads to greater theoretical simplicity:

In the original version of the theory, it was stipulated that [... \star -c] must be a faithfulness constraint. The suggestion here is simply to remove this stipulation. [...W]hereas Faithfulness might be dominant in sympathy candidate selection in many systems, the logic of OT itself compels us to expect other constraints in this role as well.

Itô & Mester (1997b: 10-11)

[...E]xtending Sympathy Theory to allow markedness constraints as selectors is desirable from an empirical point of view. Conceptually speaking, the proposed extension is also desirable since it removes the stipulation that selectors may only be faithfulness constraints.

De Lacy (1998: 14)

Nonetheless, the *undesirable* consequences of this move should now be clear:

(i) The number of potentially significant \otimes -candidates for any given input in a system expands enormously. In Extended Sympathy Theory, it equals the number of crucially dominated constraints in the grammar.

(ii) Discarding \star -Confinement severs the putative link (however tenuous) between sympathy and lexical access, for it implies that the \otimes -candidate need not be more faithful to the input than the \star -candidate in any particular respect. In this sense, the opaque constraint violations induced by \otimes O-faithfulness constraints no longer necessarily flag the infringement of an IO-faithfulness constraint in the output. Yet this means that Sympathy Theory (and, by implication, strongly parallel OT) can *describe*, but not *explain*, the existence of opacity: sympathy-theoretic mechanisms cannot be deduced as theorems from fundamental axioms of strongly parallel OT, nor can they be provided with functional grounding. In other words, sympathy proves neither an adaptation, nor a basic design constraint on the language faculty, nor a trade-off for some more fundamental property of phonological grammars (see footnote 69)! In contrast, interleaved OT deduces the possibility (as well as the formal limitations) of opacity from the architecture of the phonology-morphosyntax interface.

(iii) Extended Sympathy Theory implies an admission that synchronic markedness reversal does occur and, by implication, that interleaved OT misses no major typological generalization after all (*pace* Benua 1997). More generally, the jettisoning of \star -Confinement betrays the failure of a research programme which aims at curbing cophology divergence through UG, rather than through learnability and history. Not surprisingly, the proponents of strongly parallel OT have not yet devised an acquisition model comparable to that outlined in §3.3.3.2; recall that, as I suggested in that section, recursive input optimization within a stratified structure underpins the diachronic processes of morphologization and lexicalization that keep opacity in check despite cumulative innovations.

These problems compound the conceptual difficulties described in the §3.4.2.1: (i) the asymmetry of \otimes O-faithfulness is established in a purely stipulative manner, and (ii) \otimes -candidates are morphosyntactically unmotivated.

3.5 OPAQUE MORA PRESERVATION

3.5.1 Opacity in Hayes's canon of mora-preserving lengthening effects

The moraic analysis of West Germanic Gemination outlined in (3,56) possesses great theoretical significance. Not only does it conclusively refute \star -Confinement (with all that this result entails), but it does so in an area where there was little *prima facie* evidence to doubt the validity of the principle. McCarthy (1998: 46) confidently announced that: "The problem of compensatory lengthening in OT [...] can be resolved through sympathy to the moraic structure of the \otimes -candidate $\aleph_{\text{MAX-C}}$, which preserves input consonantism." As I shall show in this section, this promissory note would indeed have been redeemed had Hayes's (1989) typology of mora-preserving lengthening processes proved exhaustive, for all the opaque forms of compensatory lengthening listed by Hayes are triggered by segment deletion.⁷¹ The evidence of West Germanic Gemination, whilst strikingly confirming the superiority of Mora Theory over X-position Theory, nonetheless shows that mora-preserving lengthening need not be caused by segment deletion; it may be fuelled by resyllabification processes contingent on the reranking of markedness constraints. Hayes's article, on which recent studies of mora preservation seem to rely almost completely (e.g. Lee 1996), does not allow this fact to be fully appreciated.

Working within a rule-based framework, Hayes showed that all varieties of compensatory lengthening obey a unifying principle:

(3,57) **Moraic Conservation (Hayes 1989: 285)**

Compensatory lengthening processes conserve mora count.

⁷¹ Compensatory lengthening for prenasalization may be an exception to this statement, but the representation of prenasalized segments is too poorly understood to establish the point conclusively; see below.

Translated into optimality-theoretic terms, this assertion implies that compensatory lengthening processes are faithful to the moraic content of the input. More specifically, segments lengthen without mora epenthesis: i.e. no violation of DEP^μ is incurred.

(3,58) **Moraic Conservation (optimality-theoretic version)**

Compensatory lengthening processes respect DEP^μ.

Of course, the constraint-based formulation of Moraic Conservation is fully equivalent to the rule-based version only if morae inserted by Weight by Position during the initial round of syllabification are not counted as epenthetic; see §2.3. In chapter 2, I suggested that this effect can be achieved by reformulating DEP^μ and IDENT^μ so as not to be infringed by positional μ -licensors. The relevant statements are repeated here for ease of reference. (Recall that \wedge symbolizes Smolensky-conjunction.)

(3,59) (=2,40) **Positional μ -Licensing**

A segment α is positionally μ -licensed by a mora μ_α if, and only if,

- (i) α does not have an input correspondent linked to a mora,
- and (ii) α is immediately dominated by μ_α , and by μ_α only.

(3,60) (=2,43) **DEP^μ**

Let μ be a mora in the output.

DEP^μ = (a) \wedge (b)

- (a) μ has a correspondent in the input.
- (b) μ is a positional μ -licenser.

(3,61) (=2,44) **IDENT^μ**

Let α be a segment in the input.

Let β be a correspondent of α in the output.

Let α be linked to n morae.

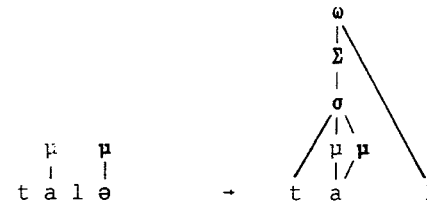
IDENT^μ = (a) \wedge (b)

- (a) β is linked to n morae.
- (b) β is positionally μ -licensed.

Lengthening is transparently mora-preserving when the trigger (i.e. the segment which donates the transferred mora) is underlyingly moraic. Hayes (1989) lists three types of transparent mora-preserving lengthening:

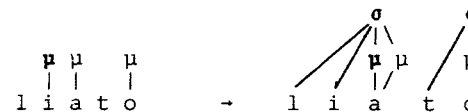
(3,62) **Vowel Loss (Hayes 1989: §4.2, §5.1.6)**

ME /talə/ → [ta:l] 'tale' (Minkova 1982; see also ch. 4 below)



(3,63) **Glide Formation (Hayes 1989: §4.3, §5.1.3)**

LuGanda /li-ato/ → [lja:to] 'boat' (Clements 1985b: 47)



(3,64) **Inverse Compensatory Lengthening (Hayes 1989: §5.1.7)**

Pāli *thūla-* / *thūlla-* 'big' < Sanskrit *sthūla-* (Hock 1986: 441)



In all three cases, the first conjunct of DEP^μ is fulfilled: the transferred mora has an input correspondent. Hence, the macro-constraint itself is not violated.

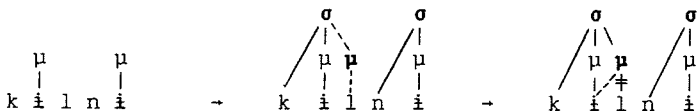
In contrast, lengthening is opaque when the mora subject to the preservation effect is not present in the underlying representation, but is rather inserted by Weight by Position during an initial round of syllabification in order to license an underlyingly non-moraic coda segment. On the surface, Weight by Position appears to misapply either because the mora in question has subsequently been attached to a different segment, or because the relevant coda consonant has acquired a new prosodic licensor. In optimality-theoretic terms,

DEP^μ seems to be violated because the mora neither has a correspondent in the underlying representation nor qualifies as a positional μ-licenser on the surface. Hayes (1989) lists a variety of processes conforming with this description.

(3,65) **Classical Compensatory Lengthening** (Hayes 1989: §3.2, §5.1.1)

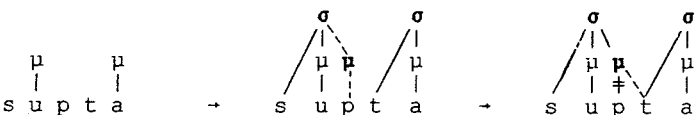
Komi /kil-ni/ → [ki:ni] 'hear' inf.; cf. /kil-i/ → [kili] 1sg.past

(Kenstowicz 1994: 295, data from Harms 1968)



(3,66) **Anticipatory Total Assimilation** (Hayes 1989: §5.1.2)

Pāli /sup-ta/ → [sut:a] 'sleep' past.ptc.; cf. /sup-ati/ → [sup:ti] (Zec 1995a: 174)



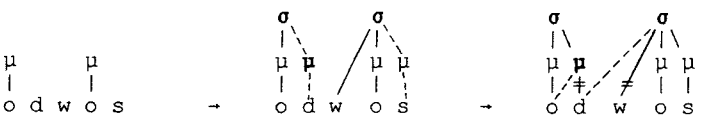
(3,67) **Perseverative Total Assimilation** (Hayes 1989: §5.1.2)

Pāli /lag-na/ → [lag:a] 'attach' past.ptc.; cf. /lag-ati/ → [lagati] (Zec 1995a: 174)



(3,68) **Double Flop** (Hayes 1989: §4.1, §5.1.5)

Ancient Greek (East Ionic) /odwos/ → [o:dos] 'threshold' (Wetzels 1985: 310)



In all these examples, the relevant mora (highlighted in boldface) ends up dominating a segment that has some other prosodic licenser; hence, the mora fails to meet criterion (ii)

of positional μ-licensing. Since it has no underlying correspondent either, it must be treated as epenthetic.

In interleaved OT, the solution to this problem is obvious. The intermediate representations in derivations (3,65) to (3,68) must correspond to the output of a non-final phonological level; in each of these representations, the relevant mora meets the criteria for positional μ-licensing, and DEP^μ is therefore satisfied via conjunct (b). The lengthening operation takes place at a later level; but here the mora in question is present in the input, and therefore fulfils DEP^μ through conjunct (a). In Sympathy Theory, in contrast, the intermediate representations in (3,65) to (3,68) are treated as ⚡-candidates. In the ⚡-candidate, the relevant mora is a positional μ-licenser and hence fulfils DEP_{IO}^μ through conjunct (b). In the ⚡-candidate, in contrast, DEP_{IO}^μ is not satisfied, but superordinate MAX_{IO}^μ demands that the mora count of the ⚡-candidate should be preserved.

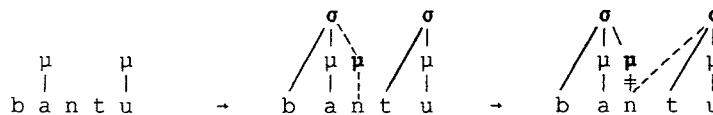
The crucial point is that, in all four cases, the ⚡-candidate characteristically retains a segment which is deleted in the actual output. In (3,65) and (3,66), the ⚡-candidate lacks the positionally μ-licensed consonant present in the ⚡-candidate. In (3,67) and (3,68), in contrast, it is a following onset consonant that is missing, prompting the resyllabification of the positionally μ-licensed segment. Under such circumstances, McCarthy's claim that the ⚡-constraint is MAX_{IO}^C holds up; ⚡-Confinement can only be challenged if the ⚡-candidate differs from the ⚡-candidate merely in terms of prosodification, rather than segmental content.

Tantalizingly, Hayes lists one form of compensatory lengthening that may just fit the bill:

(3,69) **Prenasalization** (Hayes 1989: §5.1.4)

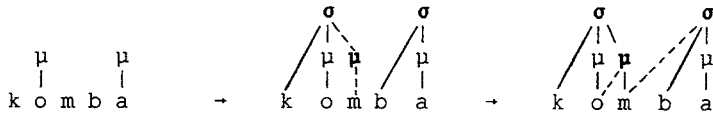
(a) 'Full-mora' compensatory lengthening

LuGanda /ba-ntu/ → [ba:ntu] 'people' (Clements 1985b: 52)



(b) 'Split-mora' compensatory lengthening

Runyambo /komba/ → [ko'mba] 'lick' (Hubbard 1995b: 250-1)



If these analyses are correct, then the σ -candidate is no more faithful to the input than the actual output: the only difference between the two lies in the syllabification of the underlying nasal consonant. The σ -candidate fulfils ONSETFORM (see §2.3.3) at the expense of * μ /SON (2,9), whilst the μ -candidate does exactly the opposite; see also tableau (2,38). In this view, the selector constraint in cases of prenasalization would be \star -ONSETFORM, a markedness constraint.

Unfortunately, the evidence of prenasalization is unreliable owing to the uncertainty of the phonological representation of prenasalized consonants. Although scholars such as Padgett (1995: 72-80) argue that prenasalized segments are onset clusters, this view is by no means universally shared; see Clements & Hume (1995: §1.3) for discussion and references. If prenasalized sequences are single-root contour consonants, then the output forms in (3,69) violate $\text{MAX}_{\text{IO}}^{\text{C}}$, as the root node of the underlying nasal is deleted and its [+nasal] autosegment spreads to the following consonant.

Accordingly, West Germanic Gemination provides absolutely crucial evidence that mora-preserving lengthening need not be triggered by segment deletion, but may involve a simple resyllabification operation. In fact, (3,56c) differs from (3,56b) solely in the addition of an association line, and so West Germanic Gemination is purely structure-building. This finding is important not only because it fills a gap in Hayes's (1989) typology, but because it conclusively refutes \star -Confinement. Additionally, West Germanic Gemination holds considerable interest in relation to the diachronic ontogenesis of opacity because its historical origins can be elucidated in fine detail.

3.5.2 West Germanic Gemination

3.5.2.1 Description

West Germanic Gemination is one of the crucial 'synapomorphies' (Lass 1997: §3.4) which define the eponymous branch of the Germanic family; see Lass (1994: §2.8). Its operation can be determined very precisely by comparing OS, OE and OHG wordforms with their cognates in Gothic (East Germanic) and ON (North Germanic), which attest to the Common Germanic forms in the relevant respects. Here I follow Simmler (1974), who examines the evidence of OS and OHG texts with remarkable thoroughness.

In all branches of West Germanic, [j] causes gemination of any immediately preceding consonant other than /r/. In OS, OE and the majority of OHG texts, [j]-triggered gemination is only attested after a short stressed vowel. After a long stressed vowel or a stressed vowel plus consonant, pre-[j] gemination is only found in Oldest Upper German (henceforth OUpG), by which I mean essentially the oldest Alemannic and Bavarian monuments; see Braune/Eggers (1975: §96, *Anm.* 1).⁷² The triggering [j] is well preserved in OS, spelt <i>; it is also occasionally attested in the oldest High German texts.

(3,70) West Germanic Gemination before [j]

(a) After 'V

Gothic		OS	OE	OHG
<i>saljan</i>	'offer' inf.	<i>sellien</i>	<i>sellan</i>	<i>zellen</i>
<i>kunjis</i>	'race' gen.sg.	<i>kunnies</i>	<i>cynnes</i>	<i>chunnes</i>
<i>hafjan</i>	'lift' inf.	(<i>af</i>) <i>hebbien</i>	<i>hebban</i>	<i>heffen</i>
<i>bidjan</i>	'ask' inf.	<i>biddien</i>	<i>biddan</i>	<i>bitten</i>

⁷² Nach Langvokal und Diphthong ergeben sich vollkommen andere Verhältnisse als bei der Gemination vor j nach Kurzvokal. Im Altenglischen und Altostniederfränkischen können keine Geminaten belegt werden, und auch im Altfrisischen und Altsächsischen sind nur Spuren vorhanden; verhältnismäßig stark ist die Gemination vor j aber im Oberdeutschen durchgeführt. (Simmler 1974: 375)

We shall see below that this divergence between OUpG and the rest of West Germanic dialects may be imputed to a difference in the distribution of [j] in their respective ancestors. This difference is in turn connected to the history of Sievers' Law. See Kiparsky (1998b), as well as §3.5.2.4 and §3.5.2.5 below.

(b) After \bar{V} or \bar{VC}

Gothic	OS	OE	OHG	OUppG
<i>dāiljan</i> 'divide' inf.	<i>dēlien</i>	<i>dāēlan</i>	<i>teilen</i>	(<i>ar</i>) <i>teillan</i> ⁷³
<i>wēnjan</i> 'expect' inf.	<i>wānian</i>	<i>wēnan</i>	<i>wānen</i>	(<i>far</i>) <i>vānan</i> ⁷⁴
<i>lāusjan</i> 'release' inf.	<i>lōsian</i>	<i>lȳsan</i>	<i>lösen</i>	<i>lōssan</i> ⁷⁵
<i>fōdjan</i> 'feed' inf.	<i>fōdian</i>	<i>fēdan</i>	<i>fuoten</i>	<i>vuottan</i> ⁷⁶

When preceded by a short vowel, the voiceless stops /p t k/ also lengthen before the liquids /r l/. This type of gemination is found in all West Germanic dialects.⁷⁷

(3,71) West Germanic Gemination before liquids

ON	OS	OE	OHG
<i>snotr</i> 'wise' nom.sg.	<i>snottar</i>	<i>snot(t)or</i>	<i>snottar</i>
<i>bitr</i> 'bitter' nom.sg.	<i>bittar</i>	<i>bit(t)er</i>	<i>bittar</i>

Note that, in these examples, the nom.sg. is usually regarded as reflecting the restructuring of the underlying representation on the basis of oblique forms such as OE *snottres* str.gen.masc.sg., where the stop is immediately followed by the liquid (Campbell 1959: §408; Murray & Vennemann 1983: fn. 7). Interestingly, pre-liquid gemination remained active but variable during the historical period in OE (Campbell 1959: §453; Hogg 1992a: §7.78), hence competing variants such as str.nom.masc.sg. *biter* ~ *bitter*, wk.nom.masc.sg. *bitra* ~ *bittra* ~ *bitera* (with epenthesis instead of gemination).⁷⁸

⁷³ Simmler (1974: 215).

⁷⁴ Simmler (1974: 221).

⁷⁵ Simmler (1974: 231).

⁷⁶ Simmler (1974: 238).

⁷⁷ Occasionally, /t/ triggers gemination after etymologically long vowels concomitantly with vowel shortening. This phenomenon is rather frequent in OE (Hogg 1992a: §7.79).

⁷⁸ Die Gemination vor *r* ist ebenfalls in allen westgermanischen Einzelsprachen vorhanden, es läßt sich aber ein Nebeneinander von geminierten und ungeminierten Beispielen feststellen. Dieses ist besonders ausgeprägt im Altenglischen, während im deutschen Sprachgebiet die geminierten Kasus und Flexionsformen die ungeminierten deutlich überwiegen. (Simmler 1974: 374)

3.5.2.2 Common Germanic syllabification

Thanks to a long-accumulated wealth of philological and linguistic research into early Germanic syllable structure, the operation of West Germanic Gemination is now easy to understand. The crux of the matter is this: all the consonant sequences in which West Germanic Gemination is attested constituted heterosyllabic coda-onset clusters in the proto-language. Various evidence, both internal and external, underpins this generalization.

First, Cj onset clusters were absolutely prohibited in Common Germanic, as in Latin; in other words, the constraint *[_σCj] was surface-true.⁷⁹ This state of affairs is profusely attested to in Gothic. Thus, words beginning with a consonant followed by yod are entirely absent from Gothic texts.⁸⁰ Calabrese (1994: 164-166) observes that the infinitive /fi-an/ 'hate' is realized as [fi.an] <fian> rather than *[fjan] <fjan>,⁸¹ even though elsewhere Gothic avoids the hiatus sequence [i.a]: e.g. /food-I-an/ → [food.jan] <foodjan>, not *[foo.di.an] <foodian> (see §3.5.2.4 and §3.5.2.5). Ungrammatical *[fjan] is ruled out by superordinate *[_σCj]. Furthermore, word-division in Gothic manuscripts falls consistently between [j] and a preceding consonant, rather than before the latter (Hechtenberg Collitz 1906; Schulze 1908; Hermann 1923; Braune/Ebbinghaus 1981: §1, *Anm.* 9; Murray & Vennemann 1983: 515; Vennemann 1987):

⁷⁹ *[_σCC] penalizes all complex onsets; see (2,28). In this sense, *[_σCj] is more specific and will be active (in the technical sense of Prince & Smolensky 1993: 82) only if *[_σCj] » *[_σCC]. For justification of *[_σCj], see Calabrese & Romani (1991) and Calabrese (1994: 163-165).

⁸⁰ The Gothic script (see e.g. Braune/Ebbinghaus 1981: §1; Marchand 1973: ch. 1) has different characters for [i] and [j]. This might be regarded as surprising: to the extent that the distribution of [i] and [j] is controlled by Sievers' Law (see §3.5.2.5), the status of these two sounds would appear to be purely allophonic, inhibiting an orthographic distinction. However, Sievers' Law was partly morphologized in Gothic (see e.g. Riad 1992: §2.3.2), and there appears to have been a marginal phonemic contrast word-initially: e.g. *iupr* 'above' ~ *juka* 'yoke' nom.acc.pl. (Marchand 1973: 70; Riad 1992: §2.3.2.3).

⁸¹ The alternative [fi.jan] <fjan> is a variant of [fi.an] with optional low-level hiatus resolution; see Braune/Ebbinghaus (1981: §10, *Anm.* 5) and §3.5.2.5.

- (3,72) (a) *ma* | *jan* 'to eat'
al | *ja* 'than, except'
sun | *ja* 'truth'
- (b) *band* | *jan* 'prisoner' acc.
sōk | *jandans* 'seeking' acc.pl.
hāus | *jan* 'to hear'

Braune/Ebbinghaus (1981: §44, *Ann.* 1) conclude, "Inlautendes *j* bildet [...] stets den Silbenanfang."

In contrast, *muta cum liquida* clusters were permissible onsets in Common Germanic and occurred fairly frequently in word-initial position; in other words, * $[_o]CC$ was not surface-true. This does not mean, however, that the constraint was totally inactive. Word-medially CR clusters were heterosyllabic after a short stressed vowel, but tautosyllabic after a long stressed vowel or a stressed vowel followed by a consonant: i.e. ${}^1\check{V}C.RV$ contrasted with ${}^1\check{V}.CRV$ and ${}^1VC.CRV$.⁸² Thus, Common Germanic *muta cum liquida* clusters turn out to have been heterosyllabic precisely in those environments where West Germanic Gemination was later to apply.

Word-division in Old English manuscripts has long been known to comply with this generalization (Wetzels 1981, Lutz 1986), but Kiparsky (1998b) finds additional confirmation in other old Germanic texts:⁸³

(i) According to Kiparsky (1998b: §8), in the *Abrogans* (Bavarian, circa 765) and the Monsee fragments (Bavarian *Umschrift* of a Rhenish Franconian original?, early 9th c.) consonant clusters are generally divided before the last consonant, which implies that CR sequences are split. Nonetheless, Kiparsky cites Schulze (1908: 408, fn.) as reporting that the Monsee fragments contain several instances of word-division *before* a CR cluster. In

⁸² Murray (1986) suggests that, in fact, Proto-Germanic liquids behaved exactly like high vocoids: i.e. when preceded by a heavy sequence, the sonorant in a CR cluster would become syllabic following Sievers' Law (see §3.5.2.5). If so, we would have ${}^1\check{V}C.RV$ versus ${}^1\check{V}.CRV$ - ${}^1\check{V}.CRLRV$, at least at an early stage. For the extension of Sievers' Law from high vocoids to sonorants, see e.g. Collinge (1985: 160-161).

⁸³ Metre provides another interesting source of evidence on the syllabification of CR clusters. Murray & Vennemann (1983: 517) quote Hermann's (1923[1978]: 280) pronouncement that, "In der altgermanischen Poesie bilden alle zweiteiligen Konsonantengruppen Position." If so, CR clusters must be heterosyllabic after a short stressed vowel.

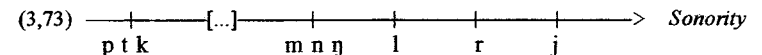
all such cases, the CR cluster is crucially preceded by a heavy syllable or two light syllables: e.g. *for* | *drono*, *hun* | *grita*, *tem* | *ple*.⁸⁴

(ii) The Gothic manuscripts differ from each other in their treatment of CR clusters. In the Codex Argenteus and in the Ambrosianus A such clusters are always kept together. Nonetheless, it has been noted that the words divided just before CR sequences in these two codices tend to be either borrowed from Greek (and hence possibly influenced by Greek orthographic conventions) or morphologically complex (and hence possibly subject to *etymologische Trennung*),⁸⁵ see Murray & Vennemann (1983: 516) and references therein. In the Codices Ambrosiani E (the *Skeireins*) and B, in contrast, Kiparsky (1998b: §7) reports that 17 out of 20 relevant instances of word-division comply with the generalization ${}^1\check{V}C.RV$ vs ${}^1\check{V}.CRV$, ${}^1VC.CRV$.⁸⁶

3.5.2.3 The rôle of CONTACT

By causing Cj and OR clusters to be heterosyllabified, superordinate * $[_o]Cj$ and high-ranking * $[_o]CC$ caused gross violations of CONTACT in Common Germanic. Recall that CONTACT requires that an onset-initial consonant should be less sonorous than an immediately preceding rhyme-final segment. In this light, it becomes clear that West Germanic Gemination optimized syllable contact by attracting the less sonorous coda consonant into the onset (Murray & Vennemann 1983: §4). In optimality-theoretic terms, one must assume that West Germanic demoted * $[_o]Cj$ and * $[_o]CC$ relative to CONTACT.

For additional support, consider the following sonority hierarchy:



⁸⁴ Related to the Monsee fragments in the Paris manuscript of the OHG translation of Isidor of Seville's *De fide catholica* (South Rhenish Franconian?, circa 800). Vennemann (1987: *Ahnang*) examines word-division practice in this manuscript, but the evidence for *muta cum liquida* clusters is unfortunately almost insignificant.

⁸⁵ "Komposita werden in der Regel nach der Etymologie getrennt, d. h. in der Kompositionsfuge (*ga* | *swalt*, *ufar* | *ist*); ebenso wird die Reduplikation behandelt (*skai* | *skaid*)" (Braune/Ebbinghaus 1981: §1, *Ann.* 9).

⁸⁶ Like Kiparsky (1998b), Vennemann (1987) takes the word-division practice of the *Skeireins* and of the Ambrosianus B to attest to the original syllabification algorithm of Gothic, closely reflecting that of Proto-Germanic. However, Vennemann (1987) also ventures the suggestion that the Codex Argenteus and the Ambrosianus A belong to a relatively innovative dialect of Gothic, where CR sequences are consistently tautosyllabic. This putative innovation he dubs the *correptio argentea*, on the analogy of the *correptio attica* (Allen 1973: 210ff.); on the latter see §3.5.2.4.

If violations of CONTACT are gradient, then this analysis of West Germanic Gemination makes two predictions:

(i) the more sonorous the onset consonant, the greater the number of preceding coda consonants subject to gemination;

(ii) if a coda consonant C_1 geminates before an onset consonant C_2 , then all coda consonants whose sonority value is lower than that of C_1 will also geminate before C_2 .

As Murray & Vennemann (1983: §4) point out, both predictions are confirmed. First, yod triggers gemination of a larger set of consonants than the liquids (and, moreover, it does so with less variability). Secondly, the only consonant not subject to pre-[j] gemination is highly sonorous /r/; conversely, the liquids only trigger lengthening of the least sonorous set of segments, i.e. the voiceless stops.

3.5.2.4 Opacity

At this point, Murray & Vennemann (1983: 525) pose the obvious question: "If PGmc. syllabication represented an undesirable structure, why didn't a simple syllable-boundary shift such as the following [sc. $^1sat.ja > ^1sa.tja$] take place?" An optimality-theoretic formulation drives home the urgency of the problem: in tableau (3,74), the marks incurred by the resyllabified candidate (c) are a proper subset of those incurred by the geminated candidate (b).

(3,74)

/akra/ ⁸⁷	CONTACT	*[_c CC	DEP ^u
(a) ak.ra	*!		
(b) ak.kra ☹		*	*!
(c) a.kra ☹*		*	

According to Murray & Vennemann, (3,74b) proved more harmonic than (3,74c) in West Germanic because the stressed syllable is heavy in the former but not the latter (see also Vennemann 1988: 45); bimoric stressed syllables, Vennemann suggests, are favoured

⁸⁷ Cf. Gothic *akrs* nom.sg. ~ *akra* dat.sg. 'field', OS *akkar*.

by the Weight Law (Vennemann 1988: 30).⁸⁸ For our purposes, the Weight Law can be replaced by the optimality-theoretic constraint PKPROM (Prince & Smolensky 1993: 39).

(3,75) PKPROM

If a syllable bears primary stress, then it is bimoric.

However, this answer is unsatisfactory for two reasons:

(i) It fails to account for the geminated OUppG forms in (3,70b): i.e. *arteillan*, *farvvānnan*, *lössan*, *vuottan*. In these examples, the long vowel or diphthong suffices to render the stressed syllable heavy; for the purposes of PKPROM, the geminate is superfluous. Moreover, assuming the corresponding protoforms to have been syllabified along the lines of their Gothic cognates [dail.jan], [ween.jan], [laus.jan] and [food.jan] (see below), resyllabification instead of gemination would have offered the opportunity of getting rid of marked superheavy syllables whilst still fulfilling PKPROM.

(ii) In order for PKPROM to trigger unfaithfulness to input mora specifications, it must dominate DEP^u and IDENT^u; but this ranking incorrectly predicts the disappearance of all light tonic syllables.

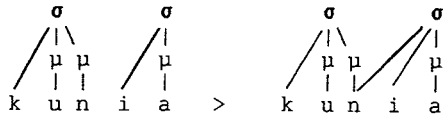
⁸⁸ In a similar vein, Lass (1994: §3.2) also suggests that the incidence of West Germanic Gemination is primarily determined by considerations of syllable weight.

(3,76)

		CONTACT	PKPROM	*[_σ Cj]	DEP ^u
kunja ⁸⁹	kun.ja	*!			
	ku.nja		*!	*	
	kun.nja ⁹⁰			*	*
dailjan	dail.jan	*!			
	dail.ljan ⁹⁰			*	
	dai.ljan ⁹¹			*	*!
trumida ⁹¹	tru.mi.da ⁹¹		*!		
	trum.mida ⁹¹				*

From the viewpoint of Mora Theory, in contrast, the answer to Murray & Vennemann's question is rather simple: West Germanic Gemination enforced CONTACT with the minimum disruption of existing structure; more specifically, the mora count of Germanic protoforms was preserved.

(3,77)



In the light of the cases of Mora Conservation discussed in §3.5.1, this development is not particularly remarkable. From a theoretical viewpoint, the interest lies, of course, in the productive alternations it gave rise to in West Germanic: e.g. ¹[trum.mjan] 'strengthen' inf.

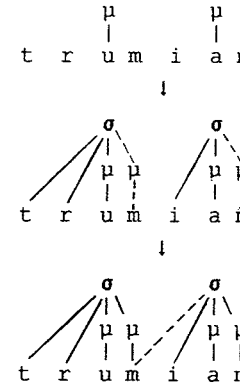
⁸⁹ Cf. Gothic *kunja* 'race' dat.sg.; OE *cynne*; OHG *kunniu* - *kunnu*, -o.

⁹⁰ Cf. OUpG *teillan*; see (3,70b).

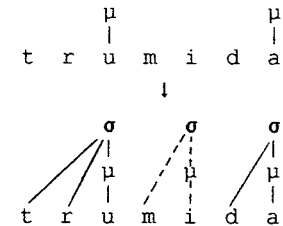
⁹¹ Whence OE *trymede* 'strengthen' 1sg.pret.ind.; cf. ¹[trum.jan] > OE *trymman* inf. (see Hogg 1992b: 157-159).

vs ¹[tru.mi.da] 1sg.pret.ind. (see footnote 91). These demonstrate that the system became synchronically opaque: the regular alternation of ¹[tru.mi.da] with ¹[trum.mjan] implies that the final consonant of the root /trum/ remained non-moraic in the underlying representation, and so gemination came to counterbleed Weight by Position in the grammar.

(3,78) (a)



(b)



In this analysis, West Germanic Gemination has nothing to do with PKPROM, or with constraints on rhyme structure in general: consonant lengthening is regarded as a by-product of the rise of CONTACT under conditions of moraic stability. Thus, liquids failed to cause gemination of voiceless stops after ¹V̄ or ¹VC sequences simply because OR clusters were already tautosyllabic in such environments. More interestingly, pre-yod gemination is predicted to occur in all environments, for Cj clusters were always heterosyllabic in Common Germanic. As Kiparsky (1998b) suggests, the puzzling behaviour of OUpG in comparison with other West Germanic dialects can be imputed to a difference in the distribution of [j] prior to gemination. As is well known, the realization of an underlying high vocoid /I/ as either a vowel [i] or a glide [j] was predictable in Proto-Germanic: roughly, Sievers' Law dictated that an underlying /C-I-V/ sequence would surface as [C.jV] when preceded by ¹V̄, but would undergo vocalization to [CiV] when

preceded by \bar{V} or \bar{VC} .⁹² This implies that Cj clusters failed to occur after bimoric sequences, which correctly predicts the incidence of pre-yod gemination in OE, OS and most OHG dialects:

(3,79)	(a)	(b)
UR	/bid-I-an/	/food-I-an/
<i>Sievers' Law</i>	bid.jan	foo.di.an ⁹³
<i>Gemination</i>	bid.djan	—

In Gothic, however, a phonological innovation is attested whereby the operation of Sievers' Law becomes restricted: vocalization is no longer allowed except in underlying /C-I-/ clusters, where a regular [C.ji]~[Cii] alternation is observed; elsewhere, the medial high vocoid becomes a glide, i.e. /C-I-V/ → [C.jV]. Kiparsky (1998b) suggests that the innovative, restricted version of Sievers' Law found in Gothic also came to be operative in OUppG. This yields the right results:

(3,80)	(a)	(b)
UR	/bid-I-an/	/food-I-an/
<i>Restricted Sievers' Law</i>	bid.jan	food.jan
<i>Gemination</i>	bid.djan	food.djan ⁹⁴

As Kiparsky (1998b: §8) puts it, "We can thus say that West Germanic gemination in Old [Upper] German testifies to a prosodic system that resembled the Gothic system [...] in having *j* after both heavy and light stems."

Kiparsky's proposal concerning the contrast between OUppG and the other West Germanic dialects has a reassuringly close counterpart in Ancient Greek. From a formal viewpoint, compensatory lengthening for the loss of /w/ in East Ionic (see (3,68) above)

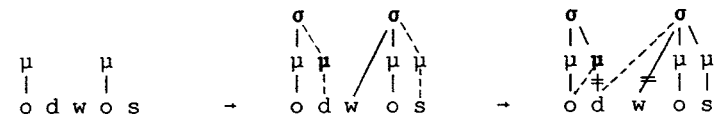
⁹² Proto-Germanic inherits Sievers' Law from Indo-European. The law was originally formulated in Sievers (1878). For the vicissitudes of its statement over more than a century of Indo-European studies, see e.g. Collinge (1985) and references therein. For the early Runic evidence attesting to the operation of Sievers' Law in Germanic, see Springer (1975).

⁹³ Whence ungeminated reflexes such as OS *fōdian*, OE *fēdan*, OHG *fuoten*; see (3,70b).

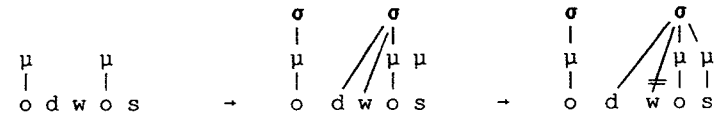
⁹⁴ Whence geminated reflexes such as OUppG *vuottan*; see (3,70b).

is entirely analogous with West Germanic Gemination. In both cases, an underlying sequence of consonant plus high vocoid is initially heterosyllabified, with the effect that the consonant acquires a mora by Weight by Position; lengthening occurs as the coda consonant subsequently spreads to the following onset. There are only two differences. First, the spreading of the coda consonant was triggered by the deletion of the following onset glide in East Ionic; in West Germanic, it was caused by CONTACT. Secondly, in East Ionic the spreading consonant vacates its positional μ -licenser as it occupies the onset of the following syllable; as a result, the floating mora redocks onto the preceding vowel. Significantly, however, the absence of compensatory lengthening in Attic is due to the fact that "Cw clusters were tautosyllabic in this dialect at the time of the loss of *w*" (Steriade 1982: 220); see also Allen (1968), Lejeune (1972), Wetzels (1985), Rialland (1993).

(3,81) (a) *East Ionic*



(b) *Attic*



As Rialland explains, the tautosyllabicity of Cw clusters was an Attic peculiarity, known as the *correptio attica*:

La «correptio attica» correspond à une restructuration syllabique. Alors que les autres dialectes répartissaient les groupes intervocaliques tels que *kn, tr* entre deux syllabes (ex.: *pat-ri, tek-non*), l'attique les a regroupés dans l'attaque syllabique (ex.: *pa-tri, te-knon*). Cette restructuration est largement attestée dans la métrique (où la première syllabe des mots *patri* et *teknon* est normalement lourde chez Homère mais légère chez Aristophane). Cette «correptio attica» concerne les groupes dont la première consonne est une occlusive sourde et la seconde une sonante (*r, l, m, n*) et devait aussi s'appliquer aux groupes avec *w* en deuxième position.

Rialland (1993: 75-76)

3.5.2.5 Sievers' Law

As the previous section indicates, to analyse West Germanic Gemination one must have an account of Sievers' Law (both in its pristine Proto-Germanic version and in the innovative restricted version attested in Gothic), for Sievers' Law determines the occurrence of Cj clusters. In this matter I shall follow Kiparsky's (1998b) proposals with only minor elaboration.⁹⁵

At the outset, it is important to emphasize the analytical constraints imposed by Richness of the Base (see §2.2.2.1). The goal of the analysis is to describe the predictable alternation of [i] and [j] in all environments other than word-initially (see footnote 80). It is therefore not possible to analyse this alternation as crucially depending on the underlying moraic specifications of non-word-initial high vocoids (assuming them to be either moraic or non-moraic), for OT treats non-contrastive input properties as emerging from the output through Lexicon Optimization. If one assumed *a priori* that underlying /I/ is, say, non-moraic, and one fashioned the analysis on this assumption, one would be incurring the Duplication Problem (Clayton 1976, Kenstowicz & Kisseberth 1977), as the prohibition of moraic non-word-initial vocoids in underlying representations would then remain to be accounted for. By the same token, constraints on moraic faithfulness such as IDENT^u and DEP^u will not be active in Sievers' Law (nor, for that matter, will IDENT^{NUC}; see footnote 65).

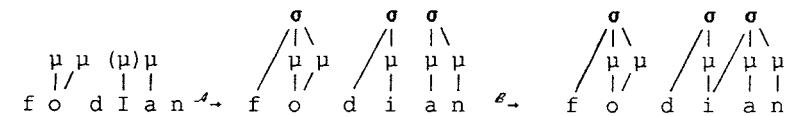
Not surprisingly, underlying /I/ vocalized whenever it constituted a sonority peak: e.g. Gothic /kun-I/ → [ku.ni] 'race' nom.acc.sg., /sal-I-da/ → [sa.li.da] 'offer' 1sg.pret.ind. For our purposes, however, the interesting alternations are those directly addressed by Sievers' Law, i.e. those arising from underlying /C-I-V/ sequences.

Let us first see how these sequences are treated under the original or full version of Sievers' Law—which operated in Proto-Germanic, and which is responsible for the distribution of yod in the input to West Germanic Gemination as attested in OE, OS, and OHG dialects other than OUppG. Here, the overriding constraint is foot well-formedness: feet must consist of exactly two morae, which excludes both [ɛ̃ʊ̃] and monosyllabic feet containing a superheavy syllable.⁹⁶ Since superordinate *[_vCj] dominates CONTACT and

crucially prohibits consonant-plus-yod onsets, a non-syllabic realization of /I/ in a /C-I-V/ sequence pushes the initial consonant into the coda of the preceding syllable: e.g. /bid-I-an/ → [bid.jan] 'ask' inf., /bid-I-Is/ → [bid.jis] 2sg.pres.ind. However, when putting /C/ in the coda would result in an ill-formed foot, high ranking FTFORM compels /I/ to become syllabic: e.g. /food-I-an/ → [foo.di.an] 'feed' inf., /food-I-Is/ → [foo.diis] 2sg.pres.ind.

Observe that, in the case of /food-I-an/ → [foo.di.an], the cost of avoiding a superheavy stressed syllable is hiatus; this indicates that ONSET is ranked below FTFORM. This hiatus can nonetheless be resolved by spreading [i] to the following vacant onset position: i.e. /food-I-an/ → [foo.di.jan]. Indeed, this realization predominates in the early Runic evidence: e.g. *ramijaR* 'trier, challenger' (Øvre Stabu spearhead, 3rd c.), *holtijaR* 'man from/ son of Holt' (Gallehus horn, circa 400); but cf. *mākia* (*mākija?*) 'sword' (Vi sword-clamp, 250-300); see Springer (1975). Other sources, geographically as well as chronologically disparate, suggest that hiatus resolution was in fact an optional low-level process throughout the history of old Germanic: e.g. one encounters free variants such as Gothic [fi.an] <fian> ~ [fi.jan] <fijan> 'hate', OE [lu.vi.an] <lufian> ~ [lu.vi.jan] <lufigean> 'love'.⁹⁷ For this reason, I am inclined to believe, with Kiparsky (1998b), that the output of Sievers' Law was [foo.di.an]; I assume, moreover, that Sievers' Law applied in the lexicon, whilst hiatus resolution could optionally apply postlexically. This assumption is reinforced by the observation that, thus construed, hiatus resolution turns out to be mora preserving, and to that extent analogous with West Germanic Gemination.

(3,82)



A: Sievers' Law (lexical)

B: Hiatus resolution (postlexical)

Nonetheless, this approach raises a tricky question: what constraint prevented the mapping /food-I-an/ → [foo.di.jan] from being effected in one fell swoop in the lexicon? Again, if

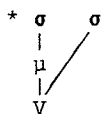
⁹⁵ I am also grateful to John McCarthy and René Kager for helping me to refine my thinking, particularly in relation to Richness of the Base and [foo.di.an]–[foo.di.jan].

⁹⁶ Kiparsky conflates both requirements under a single FTFORM constraint. A plausible alternative is to impute the rejection of [ɛ̃ʊ̃] to WSP (Prince 1990: 358), and the rejection of superheavy syllables to *[_vμμμ] (Prince & Smolensky 1993: 210-211; see also chapter 4 below). This is immaterial to the analysis of West Germanic Gemination. For evidence that VVC syllables are trimoraic in Germanic, see Hayes (1989: §7).

⁹⁷ Cf. OE [ner.jan] <nerian> 'save' (see Hogg 1992b: 161-162). Of course, *nerian* is a class 1 weak verb, whilst *lufian* is class 2, with a Proto-Germanic stem-formative /-o:l-/ rather than /-I-/.

one takes Richness of the Base seriously, then the answer cannot be DEP^u or IDENT^u, for —recall— we cannot rely on a stipulation that underlying non-word-initial high vocoids are non-moraic. In the absence of a clear answer, I will provisionally adopt a markedness constraint *i.j penalizing vowels that straddle the boundary between nucleus and onset:

(3,83) *i.j (tentative)



It is clear that this type of hiatus resolution deserves further research. Present-day English may be a useful source: hiatus glides (e.g. [fii:jə] *freer*) are found in many, though reportedly not all, dialects.

Given the above, the constraint hierarchy responsible for full Sievers' Law would be as follows:

(3,84) Full Sievers' Law

*[_σCj], *i.j, FtFORM » ONSET » CONTACT

The following tableau shows how it works. To facilitate the assessment of FtFORM, I indicate the boundaries of the strong foot.

(3,85)

		*[_σ Cj]	FtFORM	*i.j	ONS	CONTACT
bid-I-an	[_σ bi.dja]n	*!				
	[_σ bi.di.]jan			*!		
	[_σ bi.di.]jan				*!	
	[_σ bid.]jan ^{wa}					*
food-I-an	[_σ foo.]djan	*!				
	[_σ food.]jan		*!			*
	[_σ foo.]di.jan			*!		
	[_σ foo.]di.an ^{wa}				*	
bid-I-Is	[_σ bi.dji]s	*!				
	[_σ bi.dii]s		*!			
	[_σ bi.di.]jis			*!		
	[_σ bi.di.]is				*!	
	[_σ bid.]jis ^{wa}					*
food-I-Is	[_σ foo.]djis	*!				
	[_σ food.]jis		*!			*
	[_σ foo.]di.jis			*!		
	[_σ foo.]di.is				*!	
	[_σ foo.]diis ^{wa}					

As Kiparsky shows, the restricted version of Sievers' Law found to operate in Gothic (which also controlled the distribution of [j] in the West Germanic dialects from

which OUpPG descends) can be obtained by means of a minimal reranking of constraints in (3,84): FTFORM is simply demoted relative to ONSET.

(3,86) **Restricted Sievers' Law**

*[_σCj, *i.j » ONSET » FTFORM » CONTACT

In essence, this reranking renders [foo.di.an] less harmonic than [food.jan], as hiatus avoidance takes priority over quantitative considerations. Note, however, that all other ranking relations must be preserved; more specifically, something like *i.j must remain highly ranked in the lexicon, for the rival candidate [foo.di.jan], unattested in Gothic, manages to satisfy all of *[_σCj, ONSET, FTFORM and CONTACT. In this sense, the entire analysis is underpinned by the assumption that hiatus resolution is not available until the postlexical level.⁹⁸

(3,87)

		*[_σ Cj	*i.j	ONS	FTFORM	CONTACT
bid-I-an	[_ε bi.dja]n	*!				
	[_ε bi.di.]jan		*!			
	[_ε bi.di.]jan			*!		
	[_ε bid.]jan					*
food-I-an	[_ε foo.]djan	*!				
	[_ε food.]jan				*	*
	[_ε foo.]di.jan		*!			
	[_ε foo.]di.an			*!		
bid-I-Is	[_ε bi.dji]s	*!				
	[_ε bi.dii]s				*!	
	[_ε bi.di.]jis		*!			
	[_ε bi.di.]jis			*!		
	[_ε bid.]jis					*
food-I-Is	[_ε foo.]djis	*!				
	[_ε food.]jis				*!	*
	[_ε foo.]di.jis		*!			
	[_ε foo.]di.is			*!		
	[_ε foo.]diis					

The syllabification of OR sequences (Kiparsky 1998b: §7) is a comparatively simple affair. Essentially, heterosyllabicity is preferred (i.e. *[_σCC » CONTACT) unless placing the

⁹⁸ Additionally, Gothic [fi.an] » *[fjan] 'hate' shows that *[_σCj » ONSET; see §3.5.2.2.

obstruent in the coda disrupts foot well-formedness (i.e. FTFORM » *_lCC). Tableau (3,88) illustrates this pattern by means of two Gothic words which are found divided at the end of the line in the Codices Ambrosiani B and E: *ak* | *ran* 'fruit' and *af* | *tra* 'again' (see §3.5.2.2).⁹⁹

(3,88)

		FTFORM	* _l CC	CONTACT
akran	[₂ a.kra]n		*!	
	[₂ ak.]ran			*
aftra	[₂ aft.]ra	*!		*
	[₂ af.]tra		*	

3.5.2.6 Summary

In Germanic there occurred a number of coda-onset consonant clusters with a rising sonority profile. These clusters were generated by hierarchy (3,89a) in dialects ancestral to OE, OS and OHG, and by hierarchy (3,89b) in dialects ancestral to OUppG.

⁹⁹ As I pointed out in footnote 82, Murray (1986) defends a more purely Sieversian pattern for earliest Proto-Germanic, viz. ^hV̇C.RV vs ^hV̇.CR.V, with the liquid occupying a nuclear position when the OR cluster is preceded by a bimoric sequence. This is generated by the ranking FTFORM » *_lCC » NUC » CONTACT, where NUC requires that a syllable should be headed by a vowel.

		FTFORM	* _l CC	NUC	CONTACT
V̇ORV	V̇.ORV		*!		
	V̇. OR.V			*!	
	V̇O.RV				*
V̇ORV	V̇O.RV	*!			*
	V̇. ORV		*!		
	V̇. OR.V			*	

The later Common Germanic pattern of tableau (3,88) is obtained by promoting NUC over *_lCC.

- (3,89) (a) *_lCj, *i.j, FTFORM » ONSET, *_lCC » CONTACT
 (b) *_lCj, *i.j » ONSET » FTFORM » *_lCC » CONTACT

West Germanic Gemination repaired the violations of CONTACT incurred by a subset of such clusters: more specifically, C.j clusters where C is any consonant other than /r/, and O.R clusters where O is a voiceless stop. Their sonority profile was normalized by linking the coda consonant to the following onset position. This implies the demotion of *_lCj and *_lCC relative to CONTACT.

West Germanic Gemination involved opaque mora preservation. The initial consonant of the cluster remained attached to its positional μ -licenser whilst spreading to the following syllable. As a result, the preserved mora neither has a correspondent in the underlying representation nor fulfils the criteria for positional μ -licensing on the surface. In a strongly parallel analysis, the markedness and IO-faithfulness constraint violations incurred by the losing resyllabified candidate ^h[bi.djan] are a proper subset of those incurred by the winning geminated candidate ^h[bid.djan]; see tableau (3,74). Thus, West Germanic Gemination proves to be intractable in an optimality-theoretic framework combining Strong Parallelism and Strict Surface Orientation; see (3,5). Since the process is not amenable to a treatment relying on TCT or Generalized Alignment Theory, strongly parallel OT must resort to a sympathy-theoretic analysis. As I have already suggested, however, West Germanic Gemination will demonstrate the untenability of \star -Confinement, with all its deleterious consequences for the strongly parallel programme; see §3.4.2.2.

3.5.3 Peripeteia: free reranking upheld, \star -Confinement refuted

In interleaved OT, the analysis of West Germanic Gemination proves remarkably simple. The old Germanic syllabification patterns, embodied in (3,89), continue to hold within the lexicon. At the postlexical level, however, *_lCj and *_lCC are demoted relative to CONTACT. Whilst rising-sonority coda-onset clusters are thus repaired postlexically, the moraic content of the lexical output is preserved by high-ranking faithfulness constraints; accordingly, gemination is preferred to resyllabification.

A few pieces of evidence confirm our assumption that Sievers' Law applied in the lexicon, whilst gemination initially took place at the postlexical level:

(i) First of all, we have showed that hiatus resolution was a variable postlexical process, and that it was fed by Sievers' Law: e.g. Runic /holt-I-az/ → [hol.ti.az] → [hol.ti.jaz]; see (3,82). Hence, Sievers' Law must have been lexical.

(ii) In Gothic, the alternation between [ji] and [ii] showed signs of morphologization. Notably, neuter long stems resisted vocalization: e.g. /riik-I-Is/ → [riik.jis] <reikjis>, rather than the expected *[rii.kiis] *<reikeis> 'kingdom' gen.sg.; see Braune/Ebbinghaus (1981: §95) and footnote 80. This is an unequivocal indication that at that stage Sievers' Law was not postlexical.

(iii) On the other hand, there is reason to believe that gemination must have started its life as an optional postlexical phenomenon: recall that the application of gemination to OR clusters remained variable in OE for a very long period, yielding variants such as *bitra* ~ *bitra* ~ *bitera*.

Therefore, the output of tableaux (3,85), (3,87) and (3,88) corresponds to the input to the postlexical phonology. Postlexically, the selection of the optimal surface representations is straightforward. Tableau (3,90) illustrates the application of gemination in the West Germanic dialects from which OS, OE and OHG descend. Observe that the ranking *i,j » ONSET blocks hiatus resolution: i.e. [foo.di.an] > [foo.di.jan]. To obtain the opposite result, *i,j and ONSET must simply swap places in the hierarchy. No further reranking is needed; notably, IDENT⁺ can remain superordinate, as hiatus resolution is mora-preserving (see 3,82).

(3,90) West Germanic Gemination; dialects ancestral to OS, OHG and OE

Postlexical level		IDENT ⁺	CONTACT	FTFRM	*i,j	ONS	*[Cj]	*[CC]
bid.jan	bi.di.jan	**!			*!			
	bi.di.an	*!				*		
	bi.djan	*!					*	*
	bid.jan		*!					
	bid.djan						*	*
foo.di.an	food.djan	**!		*!			*	*
	foo.djan	*!					*	*
	food.jan	*!	*!	*!				
	foo.di.jan				*!			
	foo.di.an					*		
ak.ran	ak.ran		*!					
	a.kran	*!						*
	ak.kran							*
af.tra	aft.ra		*!					
	aft.tra	*!		*!				*
	af.tra							*

In the West Germanic ancestor of OUppG, the postlexical hierarchy differs only in terms of the ranking of FTFORM, which is crucially dominated by IDENT⁺ and CONTACT: hence the trimoraic syllable in [food.djan], whence OUppG *vuottan*.

(3,91) West Germanic Gemination; dialect ancestral to OUpG

Postlexical level		IDENT*	CONTACT	*i,j	ONS	FtFRM	*[_o Cj	*[_o CC
bid.jan	bi.di.jan	**!		*!				
	bi.di.an	*!			*!			
	bi.djan	*!					*	*
	bid.jan		*!					
	bid.djan						*	*
food.jan	foo.di.jan	**!		*!				
	foo.di.an	*!			*!			
	foo.djan	*!					*	*
	food.jan		*!			*		
	food.djan					*	*	*
ak.ran	ak.ran		*!					
	a.kran	*!						*
	ak.kran							*
af.tra	aft.ra		*!					
	aft.tra	*!				*		*
	af.tra							*

An interesting question concerns the learnability of these derivations. Why does the West Germanic child not posit [bid.djan] as the lexical-level representation of surface [bid.djan] —or, in OUpG, [food.djan] as the lexical-level representation of surface [food.djan]? What triggers [bid.jan] as the input to the postlexical phonology? The question may appear particularly pressing in the light of input optimization (§2.2.2.1, §3.3.3.2), which would appear to encourage the projection of SRs onto the lexical*output.

The answer lies in the fact that the acquisition routine must incorporate some device to avoid superset grammars. Note that the West Germanic child's trigger experience contained VC.CjV sequences, but no *V.CjV.¹⁰⁰ Crucially, in an ordinary Richness-of-the-Base scenario there is no hierarchization of CON generating a language *L* such that *L* has VC.CjV but lacks *V.CjV. This fact unambiguously indicates that West Germanic surface VC.CjV must be the opaque realization of something else. I suggest, in sum, that the acquisition algorithm provided by UG operates in such a way as to eschew superset grammars; in West Germanic, the absence of *V.CjV triggers this anti-superset feature,¹⁰¹ leading the child to posit VC.jV as the lexical representation of postlexical VC.CjV.¹⁰² This is consistent with the assumption that in an interleaved grammar Richness of the Base only holds at UR (see §3.3.1.1).

Recall now Benua's (1997: 90) challenge to interleaved OT: "Promotion of faithfulness is the key to serial OT analyses of cyclic effects—in fact it is only faithfulness, and never markedness, that changes its ranking position between levels"; see (§3.3.3.3). Benua suggests that, from the viewpoint of interleaved OT, this phenomenon is entirely unexpected and can only be accommodated in a stipulative fashion. It should now be clear that West Germanic Gemination refutes Benua's claim. Tableaux (3,85) and (3,87) show that, at the lexical level, the ranking *[_oCj] » CONTACT crucially selects [bid.jan] at the expense of [bi.djan]. Yet, equally crucially, the opposite ranking, i.e. CONTACT » *[_oCj], rules out [bid.jan] in favour of [bid.djan] at the postlexical level, as shown in tableaux (3,90) and (3,91). There is no doubt that both CONTACT and *[_oCj] are markedness constraints. Therefore, West Germanic Gemination provides a robust instance of synchronic markedness reversal. The principle of free reranking is thereby upheld; Benua's claim does not stand up.

¹⁰⁰ Recall also that no Cj clusters occurred word-initially; see §3.5.2.2.

¹⁰¹ This anti-superset feature is a logical necessity, but its nature remains to be specified. Of course, the key problem is how to build a superset-avoiding property into UG without recourse to negative data (on which see Marcus 1993). Pulleyblank & Turkel (1997) offer some specifically optimality-theoretic suggestions based on the notion of 'gradient retreat', but they are primarily intended for a strongly parallel model and do not directly cast light on the problem at hand.

¹⁰² There is another factor which could conceivably play a rôle. Recall that OE displays variation between *bitra*, *bittra* and *bitera*. A child exposed to such variants has direct access to VO.RV parses: e.g. [bit.ra]. It is perhaps significant in this respect that pre-liquid gemination remained productive in OE for centuries since its West Germanic inception.

I shall now turn to the sympathy-theoretic analysis of West Germanic Gemination. I shall begin with the gemination pattern attested in OS, OE, and in OHG dialects other than OUppG. Representative examples are listed under (3,92) for ease of reference.

- (3,92) /bid-I-an/ → [bid.djan]
 /food-I-an/ → [foo.di.(j)an]
 /akra/ → [ak.kra]
 /aftra/ → [af.tra]

The analysis is summarized in tableau (3,93).

Consider first pre-yod gemination. Given input /bid-I-an/, a geminated output form such as \mathfrak{M} [bid.djan] violates IDENT_{IO}^μ, as the lengthened consonant is neither underlyingly moraic nor positionally μ-licensed. This constraint violation is opaque to the extent that gemination is not required to satisfy a higher-ranked markedness or IO-faithfulness constraint: recall that the markedness and IO-faithfulness infractions incurred by the resyllabified candidate \mathfrak{M} [bi.djan] form a proper subset of those of the opaque winner. Accordingly, identity with the mora count of the \mathfrak{C} -candidate [bid.jan] must outrank IO-faithfulness.

Now comes the main question. Which is the \star -constraint responsible for selecting \mathfrak{C} [bid.jan]? As we know, for there to be opacity the output form must violate the \star -constraint (which is therefore crucially dominated). There are only *three* constraints simultaneously fulfilled by \mathfrak{C} [bid.jan] and broken by \mathfrak{M} [bid.djan]: viz. IDENT_{IO}^μ, *[_CCj] and *[_CCC. Obviously, IDENT_{IO}^μ cannot be the \star -constraint: as a matter of pure logic, the constraint opaquely violated by the \mathfrak{M} -candidate selects the transparent loser as \mathfrak{C} -candidate, in this case \mathfrak{M} [bi.djan]. Therefore, the selector constraint must be either *[_CCj] or *[_CCC. Yet, beyond the shadow of a doubt, both are markedness constraints. \star -Confinement is thereby refuted, *pace* McCarthy (1998).

In fact, the analysis is simplified if we choose *[_CCC as the \star -constraint. In this way, the sympathetic effect extends to pre-liquid gemination. Given input /akra/, \star -*[_CCC selects [ak.ra] as the \mathfrak{C} -candidate. Faithfulness to the mora count of the latter prompts the opaque violation of IDENT_{IO}^μ in \mathfrak{M} [ak.kra]; cf. transparent \mathfrak{M} [a.kra]. Observe, however, that

\star -*[_CCC also designates [aft.ra] as the \mathfrak{C} -candidate for input /aftra/. In this case, however, IDENT_{IO}^μ is unable to trigger gemination: *[aft.tra] is ungrammatical. This implies that FTFORM must dominate IDENT_{IO}^μ.

Another point of interest concerns input /food-I-an/. Under the ranking { \star i.j, ONSET} » {*[_CCj, *[_CCC}, which emerges in the analysis of /bid-I-an/,¹⁰³ the transparent candidate is \mathfrak{M} [foo.djan], without vocalization. Nonetheless, $\mathfrak{N}_{[C]CC}$ turns out to be [foo.di.an], which wins opaquely through sympathy to itself!

¹⁰³ This happens by transitivity. First, it is essential to the selection of the \mathfrak{C} -candidate [bid.jan] that CONTACT should be outranked by \star i.j and ONSET; these two constraints discard the other candidates which fulfil \star -*[_CCC, viz [bi.di.jan] and [bi.di.an] respectively. Secondly, for [bid.jan] not to be the \mathfrak{M} -candidate CONTACT must dominate both *[_CCj and *[_CCC. It follows that { \star i.j, ONSET} » {*[_CCj, *[_CCC}.

	FTFORM	IDENT _σ ^h	*i,j	ONSET	CONTACT	IDENT _{IO} ^h	*[_σ Cj	★-*[_σ CC
bid-I-an	bi.di.jan	**	*			*		✓
	bi.di.an	*		*				✓
	bid.jan				*			✓
	bi.djan	*				*		*
	bid.djan					*		*
food-I-an	food.djan	**				*	*	*
	foo.djan	*				*	*	*
	food.jan				*			✓
	foo.di.jan	*	*			*		✓
	foo.di.an			*				✓
akran	ak.ran				*			✓
	a.kran							*
	ak.kran	*				*		*
aftra	aft.ra				*			✓
	aft.tra	*				*		*
	af.tra		*					*

N. B. For hiatus resolution, i.e. /food-I-an/ → [foo.di.jan], reverse the relative ranking of *i,j and ONSET.

Tableau (3,93) achieves the main goal of section §3.5: it conclusively demonstrates that ★-Confinement is untenable. The implications of this finding for Sympathy Theory, and for strongly parallel OT in general, were discussed in §3.4.2.2 and need not be repeated here. Suffice it to say that none of the instances of opacity associated with West Germanic Gemination is amenable to a functional explanation in terms of lexical recognition. Quite the reverse: the opaque winners violate IDENT_{IO}^h solely in order that they may fulfil an antagonistic O -faithfulness constraint.

Before proceeding to the conclusions of this chapter, however, let us consider the gemination pattern attested in OUppG. In (3,94) I give representative examples for ease of reference:

- (3,94) /bid-I-an/ → [bid.djan]
 /food-I-an/ → [food.djan]
 /akra/ → [ak.kra]
 /aftra/ → [af.tra]

The crucial difference with respect to OS, OE and OHG is the the fact that long stems display pre-yod gemination: e.g. [food.djan], OUppG *vuottan*; cf. (3,92). Pre-liquid gemination, in contrast, remains circumscribed to short stems: e.g. *[aft.tra].

Here the point of interest is that not one but *two* distinct sympathetic relationships are required; see §3.4.2.1. Both in the case of pre-yod gemination and in that of pre-liquid gemination, the O -candidate contains a rising-sonority heterosyllabic cluster: e.g. O [bid.jan], O [food.jan], O [ak.ra], O [aft.ra]; the selector constraints ★-*[_σCj and ★-*[_σCC reject alternative candidates with complex onsets. Recall, however, that liquids do not trigger gemination in long stems. This implies that the O -constraint responsible for gemination in OR clusters must be crucially dominated by FTFORM. In contrast, the O -constraint that triggers gemination before [j] outranks FTFORM, given that [food.djan] > [foo.djan].¹⁰⁴

¹⁰⁴ In tableau (3,95), the name of each separate O -identity constraint is prefixed with the name of the ★-constraint that selects the particular O -candidate which the O -constraint in question demands faithfulness to: e.g. *[_σCj-MAX_σ^h requires that all the morae in N_{σ,Cj} should have a correspondent in the output.

	*i,j	ONS	*[Cj]- MAX _{oo} ^μ	CONTACT	FTFORM	*[LCC]- IDENT _{oo} ^μ	IDENT _{io} ^μ	★-[LcJ]	★-[L _o CC]
bid-I-an	bi.di.jan	*				**	*	✓	✓
	bi.di.an	*				*		✓	✓
	bid.jan N _{r[cj]} N _{r[cc]}			*				✓	✓
	bi.djan		*			*		*	*
	bid.djan						*	*	*
food-I-an	foo.di.jan	*				**	*	✓	✓
	foo.di.an	*				*		✓	✓
	food.jan N _{r[cj]} N _{r[cc]}			*	*			✓	✓
	foo.djan		*			*		*	*
	food.djan						*	*	*
akra	ak.ra N _{r[cc]}			*				✓	✓
	a.kra N _{r[cj]}					*		✓	*
	ak.kra						*	✓	*
aftra	aft.ra N _{r[cc]}			*	*			✓	✓
	af.tra N _{r[cj]}					*		✓	*
	aft.tra				*		*	✓	*

3.6 EPILOGUE: WHERE THE MORAL OF THE TALE IS DRAWN

Interleaved OT, as instantiated in LPM-OT (Kiparsky 1998a) and SBPM (Orgun 1996b), innovatively combines two of the most important ideas ever to have arisen in the history of generative phonology: on the one hand, the LPM model of the phonology-morphosyntax interface, based on level segregation and cyclicity; on the other, the optimality-theoretic apparatus of constraint ranking and harmonic evaluation. In this chapter I have argued, in agreement with Kiparsky (1998a), that this fusion of insights leads to a major theoretical breakthrough in relation to opacity. In the context of the problematic descriptive gaps left open by Prince & Smolensky (1993), interleaved OT succeeds in dispensing with an *ad hoc* theory of misapplication; the possibility of opacity is deduced from the design of the phonology-morphosyntax interface. At the same time, stringent formal limitations on the complexity of misapplication effects emerge from the fact that input-output mappings are bound by Strict Surface Orientation, and are therefore transparent within each level or cycle. Indeed, as Orgun has suggested, it should be easy to implement the theory non-derivationally, for it postulates a relatively pared-down array of phonological representations; no entities devoid of morphosyntactic status, such as SPE-style intermediate representations or \otimes -candidates, are countenanced.

I have been at pains to emphasize, moreover, that interleaved OT successfully integrates the synchronic and diachronic aspects of the opacity problem. Innovations arising through stochastic processes of hypocorrection are the main source of misapplication phenomena: they typically produce the effect of ‘adding a rule at the end of the grammar’ (Halle 1962: 68; Kiparsky 1970: 309). In interleaved OT such surface-true innovations can create opacity unimpeded by undesirable global side-effects: by being encoded at the postlexical level, the new phonological generalizations remain transparent and morphologically insensitive, whilst they are free to counterfeed or counterbleed phonological processes located in the lexicon. In this sense, misapplication is a ‘Rube Goldberg property’ of phonological grammars (Lightfoot 1999: ch. 5): opacity-causing processes infect grammars through contingent misacquisition phenomena, and thrive parasitically on the levelled architectures designed by UG.

On the other hand, the recursive operation of Lexicon Optimization during the acquisition process keeps misapplication effects from accumulating indiscriminately. In the absence of robust cues (such as recurrent alternations) in the trigger experience, the child posits input representations for a particular phonological level which are identical with the

corresponding outputs. Thus, if any contingent circumstance attenuates the relevant cues, an innovative postlexical ranking will percolate upwards to the word level, and thence eventually to the stem level. Ostensibly, Lexicon Optimization underpins OT's answer to the Duplication Problem; but we can now see that, when applied within an interleaved model, it also predicts the life cycle of sound change, whereby innovative changes tend to become morphologically sensitive until they eventually undergo morphologization or lexicalization.

In contrast, strongly parallel OT shows clear signs of having entered a 'degenerating problemshift', in the technical sense of Lakatos (1970). This research programme is defined by the idea that misapplication effects are caused by transderivational correspondence relationships. This hypothesis has—as it should—fostered novel, previously unanticipated empirical predictions; but corroboration has not been forthcoming. Among the predictions that have already been withdrawn or are currently under fire, the following stand out:

(i) that OO-correspondence is a symmetrical relationship, capable of inducing misapplication in base forms;

(ii) that *o*-bases can be identified by some general criterion of compositionality;

(iii) that non-vacuous Duke-of-York gambits are impossible in Sympathy Theory under ★-Confinement;

(iv) that ★-constraints must be drawn from the IO-faithfulness set and that, therefore, ✱O-correspondence functions as a marker of IO-unfaithfulness;

(v) that there is no synchronic markedness reversal.

Proponents of strongly parallel OT have typically reacted to falsifying empirical evidence in a regressive, content-decreasing fashion. Two theoretical moves are representative of this trend:

(i) Benua (1997) desists from the attempt to formulate a general compositionality criterion; she suggests that in an OO-correspondence relationship the *o*-base is simply stipulated in the subcategorization frame of the affix;

(ii) Itô & Mester (1997b) simply give up ★-Confinement.

In my view, however, the most serious indictment of strongly parallel OT, in its current state, is that it renders the very existence of misapplication phenomena profoundly baffling. The theory offers no internal explanation: e.g. the principles governing the selection of ✱-candidates are axiomatic; they cannot be deduced from more fundamental postulates. Yet, at the same time, Extended Sympathy Theory cuts itself off from hopes of

finding some external, functional justification. More generally, OO- and ✱O-correspondence cannot even bask in the reflected glory of BR-identity, for, as I have emphasized, the latter is not a transderivational but a syntagmatic relationship, and, like IO-faithfulness, it does not require stipulative anti-symmetry devices. Nor can I glimpse a meaningful relationship between the synchronic structure of strongly parallel grammars and the diachronic processes fuelling the rise and fall of misapplication effects. I feel, in sum, that under the present circumstances only a major theoretical counterblow, a veritable *coup de main*, could restore confidence in Strong Parallelism.

Indeed, a question to ponder is why Strong Parallelism has generated such enthusiasm. Notably, it is not immediately clear why 'noncyclicity' should have been lumped together with 'nonsequentiality' (i.e. the rejection of rule ordering) as a corollary of output orientation; see Itô & Mester (1997a: 419). It is true that, as Orgun (1996a: §2.2) suggests, any form of recursive mapping in phonology tends to be regarded as undesirably complex in computational terms; but the founding text of OT contains salutary reminders of the feebleness of arguments from computational complexity, given our state of knowledge (Prince & Smolensky 1993: §10.1.2.Q1). In fact, the current opacity debate is in certain respects strikingly reminiscent of the abstractness debate of the late 60's and 70's. In particular, two elements recur: the attempt to legislate whole classes of opacity phenomena out of existence, and the pursuit of a thorough-going functionalism.

Phonologists working in the aftermath of SPE were quick to realize that SPE-style derivations posed massive problems of learnability and overgeneration. Often enough, the chosen strategy was to legislate the worst-looking types of abstractness out of existence. The austerity measures suggested ranged from the mild, e.g. (3,96), to the draconian, e.g. (3,97).

(3,96) Alternation Condition (Kiparsky 1968)

Obligatory neutralization rules cannot apply to all occurrences of a morpheme.

(3,97) True Generalization Condition (Hooper 1976: 13)

All rules should express transparent surface generalizations, generalizations that are true for all surface forms and that, furthermore, express the relation between surface forms in the most direct manner possible.

Yet, with the benefit of hindsight, the anti-abstractness furore now seems misdirected. In phonology, where 'poverty of the stimulus' arguments are hard to construct, one is justified in attributing some property to UG when one observes a consistent mismatch between the facts of phonology and those of phonetics.¹⁰⁵ In this sense, it is not for UG to legislate against unlearnably opaque phonological systems: these rule themselves out of existence by their very nature. Yet one seems to feel the allure of anti-abstractness regimentation in two strongly parallel proposals: McCarthy (1998) attempts to devise a theory of misapplication where non-vacuous Duke-of-York gambits are not just unlearnable, but strictly unimplementable with UG means (see footnote 70); similarly, Benua (1997) takes the impossibility of synchronic markedness reversal to be a virtue of TCT—counterfactually, as it turns out.

In this sense, strongly parallel OT appears to place a high premium on UG models whose ability to encode opaque phonological generalizations is stringently curtailed. This, I suspect, reflects a gap in the overall research programme. Strongly parallel OT has hitherto fostered little systematic reflection upon the ontogeny of opaque phonological systems. This becomes apparent in Benua's handling of the cophonology divergence problem. When Benua notes that in interleaved OT ranking differences across levels are neither random nor unlimited, she holds grammar theory accountable to the observation, failing to broach the possibility of a historical explanation. Yet important typological generalizations about natural language, as it is known to us, may be 'path-dependent', i.e. contingent upon the dynamics of language change. Typically, when one hypostatizes such path-dependent properties of language as UG principles, empirical anomalies begin to crop up.¹⁰⁶

There is another feature of the strongly parallel programme that is reminiscent of the abstractness-concreteness debate. Proponents of Strong Parallelism insist on dealing with misapplication phenomena through CON, typically by devising tranderivational identity constraints. Since, in turn, the popular view is that constraints are 'grounded', strongly parallel OT holds up the tantalizing hope that functional explanations for opacity will be ultimately forthcoming. McCarthy (1998: 13) spurs this hope when he suggests that sympathetic effects "indirectly improve the recoverability of the input from the output";

¹⁰⁵ See e.g. Hayes (1996: §6), where such mismatches are analysed in terms of gradience vs categorialness and asymmetry vs symmetry.

¹⁰⁶ The Affix Ordering Generalization is probably a case in point; see §3.3.1.2.

along similar lines, Kenstowicz (1996: 370) states that OO-correspondence "serves to improve the transparency of morphological relationships between words and thus may enhance lexical access."

In comparison, interleaved OT offers a decidedly unglamorous explanation for the fact that human beings can efficiently learn and compute opaque phonological generalizations. It is not suggested that this ability confers adaptive advantages in terms of language use; rather, opacity is regarded as a 'Rube Goldberg feature of grammars', like ergative splits (Lightfoot 1999: §5.4). The interaction between phonology, morphology, syntax and the lexicon throws up the possibility of misapplication. This possibility becomes actualized, in turn, because the stochastic hypocorrection processes which lead to phonological innovation are to a large extent blind; top-down constraints on phonological innovation (Kiparsky 1995: §2) are not powerful enough to maintain transparency.

...Yet, mystifyingly, some global diachronic effects are in fact observed in the history of languages. To these I turn in the next chapter.

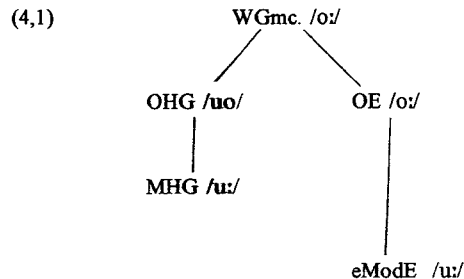
Chapter 4

Globality

4.1 GLOBAL EFFECTS IN PHONOLOGICAL CHANGE

4.1.1 Is sound change not so blind after all?

Misapplication defeats our untutored intuitions, but there are phenomena in phonological change which upon reflection prove no less baffling for all their appeal to our sense of pattern. One such phenomenon is convergent drift (Sapir 1921: ch. 8; Jakobson 1936; R. Lakoff 1972; Lass 1987: §8, 1997: §3.5, §6.3.5). Convergent drift is said to occur when genetically related but non-contacting languages undergo identical or near-identical innovations. Lass (1997: 120-121) gives the example of the raising of West Germanic /o:/ in English and High German. In both languages, the present-day reflex of /o:/ is /u:/, hence the cognates *cool* ~ *kuhl*, *stool* ~ *Stuhl*. Nonetheless, the two sisters underwent the change independently from each other: in High German the raising took place during the middle period via diphthongization in OHG; in English the date of the raising is somewhat controversial, but it was complete in the early Modern period and did not go through an intermediate stage with falling diphthongs. Needless to say, High German and English were not in close contact at this stage of their evolution.



Convergent drift is representative of a class of diachronic phenomena to which I shall refer as 'global', on the analogy of Kisseberth's (1973) notion of global rule

application in synchronic grammar (for other putative instances of diachronic globality, see Lass 1987). Global diachronic phenomena bear two hallmarks:

(i) Languages subject to global diachronic effects describe non-random trajectories in the finite space of possible grammars defined by UG. In this sense, their evolution seems not to comply with a random-walk model, but appears to require some (partially) deterministic dynamics (e.g. a Markov chain). This is clearly illustrated in the case of convergent drift. Take a set of grammar populations occupying a small subregion of linguistic phase space;¹ given enough time to evolve randomly, one would expect these populations to disperse, a bit like the molecules of an ideal gas filling a chamber. If instead all the members of the set move into some other confined subregion within the space of possible grammars, then the dynamics of change cannot be completely random.

(ii) Very often, the outcome of a global diachronic effect can be characterized in terms of some constraint or set of constraints. In other words, successive innovations either fail to violate or actively implement a particular constraint. As an example, Green (1997: 21-22) mentions the Open Syllable Law of Proto-Slavic. Proto-Slavic underwent, both before and after its dialectal breakup, a series of changes whose cumulative effect was to remove closed syllables; these changes included loss of word-final obstruents, cluster simplification, vowel nasalization with deletion of the trigger, metathesis, and epenthesis (Bräuer 1961, Carlton 1990). As Green notes, each innovation targets some set of structures where NOCODA is violated.²

For Jakobson (1929), the evidence of globality refutes the Neogrammarian assumption that innovations arise through the accidental phonologization of automatic phonetic effects. Unfortunately, Jakobson had nothing better to substitute for this assumption than teleological mystification: "Ce n'est pas en renonçant à la notion de 'loi phonique' qu'on dépasserait la tradition des néo-grammairiens, mais bien en abandonnant la conception mécanique de cette notion et en l'interprétant téléologiquement" (Jakobson 1929[1962]: 21-22).³ Kiparsky (1995: §2.1) agrees that the Neogrammarian doctrine of

¹ For attempts to model regions of linguistic phase space as lattices, see Greenberg (1978, 1995) or Anttila & Cho (1997).

² In the moraic framework outlined in chapter 2, the effects of NOCODA are modelled by the ranking *BRANCH#, *μ/C » PARSE^{2#}; see (2,11).

³ Gallons of ink have been spilt on the problem of teleology. Lass has grappled with the evidence, as well as with the metaphysics, in a large number of works (Lass 1974, 1980, 1985, 1987, 1992, 1997). For a skeptical view which nevertheless acknowledges the empirical issues raised by global effects, see Vincent (1978).

blind sound change cannot be entirely correct, but propounds instead what he calls a 'variation/selection model' of phonological innovation. According to Kiparsky, phonetic variation is rife in the primary data to which the child is exposed, but not all automatic phonetic effects have an equal chance of becoming phonologized; notably, there must be top-down selection in phonological acquisition, whereby the child's emergent grammar constrains the admission of new variants.

On our view, the learner [...] selectively intervenes in the data, favoring those variants which best conform to the language's system. Variants which contravene language-specific structural principles will be hard to learn, and so will have less of a chance of being incorporated into the system.

Kiparsky(1995:655)

Kiparsky's proposal can be formally modelled without incurring teleological perplexities. Dresher's cue-based approach to phonological acquisition gives a few pointers (Dresher & Kaye 1990; Dresher 1992, 1993, 1994, 1996; in particular, see Dresher 1996: 252):

- Phonological acquisition is not instantaneous, but follows a 'learning path' (Lightfoot 1989). Constraint rankings⁴ are established according to a particular sequence, which may be at least partially determined by UG.
- The learner does not attempt to match the primary linguistic data, but establishes a relative constraint ranking when exposed to a specific configuration in the trigger experience, i.e. a 'local cue'.
- The acquisition strategy is deterministic (Marcus 1980, Berwick 1985). Once a constraint ranking is established, it may not be undone.

A learner modelled on these assumptions can produce the sort of top-down effect that Kiparsky seems to have in mind. Notably, the assumption of determinism entails that, once a constraint ranking is established on the basis of some local cue, no phonetic effect requiring an opposite ranking may become phonologized, regardless of its salience. This is reminiscent of Lightfoot's (1989) claim that, once a child sets a syntactic parameter in a particular way, all evidence contradicting this parameter setting drops from the triggering experience; the child is literally unable to 'attend to' the counterevidence.

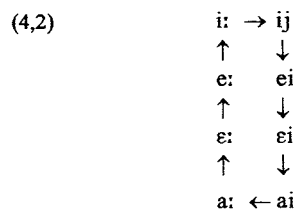
⁴ Dresher's acquisition model is parameter-based. Here I reformulate his postulates in terms of constraints. Of course, I am not denying the possibility that the translation from parameters to constraints may turn out to raise non-trivial obstacles.

For our purposes, Kiparsky's (1995: §2.1) suggestions raise two problems. First, are all global diachronic effects to be imputed to top-down pressures from the child's emergent grammar? Secondly, and more importantly, if phonological innovation is indeed subject to such top-down pressures, is our approach to the ontogeny of opaque systems compromised? Recall that, in chapter 3, we based our account of opacity on the assumption that innovation at the postlexical level is not constrained by rankings obtaining at higher levels. In the next two sections I shall answer both questions in the negative: some —though crucially not all— global diachronic effects can be explained on purely phonetic grounds; and top-down effects in the acquisition of the constraint hierarchy of a given phonological level need not involve selective pressures from higher levels.

4.1.2 Globality without top-down selection

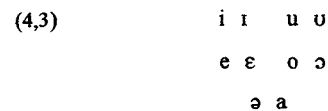
The likelihood that a child will misinterpret some coarticulation or acoustic distortion phenomenon as a phonological alternation controlled by the target grammar varies according to a number of factors. Crucially, some of this factors are purely physical or physiological; there normally is, for example, an ascertainable correlation between acoustic dissimilarity and auditory discrimination. The rôle of such extragrammatical factors in determining the inherent probability of hypocorrection processes can be quantified in the laboratory using experimental techniques. Winitz, Scheib & Reeds's (1972) confusion matrix for stop+vowel syllables indicates, for example, that the probability of /ku/ being misheard as /pu/ is far smaller than the probability of /ki/ being misheard as /ti/, quantified as 0.24 against 0.47 (Ohalá 1989: 185); this asymmetry can be related to purely physical factors such as the similarity of formant transitions. Interestingly, the asymmetry accounts for the fact that velar coronalization (e.g. pre-OE ¹/kinn-/ > OE /tʃinn-/ 'chin') is recorded far more frequently than velar labialization (e.g. Proto-Bantu ¹/-kumu/ > West Teke /pfumu/ 'chief'). The explanation is quite simple: the greater the incidence of a particular hypocorrection phenomenon, the more likely the child is to acquire the corresponding grammatical variant; the more frequently a grammatical variant is acquired, the more likely it is to propagate across a population. Given a numerically quantified misperception rate, a dynamical systems model can derive precise predictions concerning the rate of propagation of the corresponding grammatical variant across a population (Pulleyblank & Turkel 1996; see also Nigoyi & Berwick 1997).

Note that this proposal does not incur the type of self-contradiction which besets so-called 'typological' approaches to syntactic change (McMahon 1994: §6.2; Lightfoot 1999: §8.2). The proponents of these approaches often suggest that a certain type of syntactic system is inherently unstable and prone to change because the sentences it generates are hard to process; the problem then becomes how to account for the fact that the very same processing difficulties did not block the birth of the relevant system in the first place. The probabilistic model of phonological hypocorrection avoids this perplexity because the incidence of a hypocorrection process $\beta \rightarrow \gamma$ is logically and empirically independent from the incidence of a complementary process $\alpha \rightarrow \beta$, moreover, the probability of $\alpha \rightarrow \gamma$ may be very low. In topological terms, therefore, the dynamics of hypocorrection need not only define 'sources' and 'sinks', but may also generate other attractors such as 'limit cycles' (I. Stewart 1989: 86-93). Boersma (1997: §5) gives an example of an eternal hypocorrection loop, all of whose steps are attested in the history of the West Germanic languages:



Greenberg (1978: 75-76; 1995: 151-152) expresses the same idea when he argues that the frequency and stability of linguistic types are independent magnitudes: by 'frequency', he means the probability that languages of whatever type will enter a particular state; by 'stability', the probability that a language, once in that state, will not exit (see also Croft 1990: 206ff.).

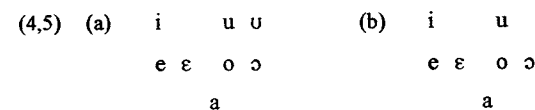
To illustrate the applicability of these ideas to problems of diachronic globality, let us consider a case of convergent drift reported by Archangeli & Pulleyblank (1994). It is a robust typological observation that perfectly symmetrical 5+5 ATR-harmony systems, such as (4,3), are very rare (Lindau 1978: 553):⁵



In most cases, the ATR low vowel is missing, yielding (4,4):



Additionally, many languages also have gaps among the high vowels, lacking either [ɪ] or both RTR vowels [ɪ ʊ]. As a result, in many ATR-harmony systems only the mid vowels harmonize.



These gaps are grounded in articulation: the tongue being an object of relatively constant volume, root advancement tends to trigger body raising, whilst root retraction tends to trigger body lowering (Lindau 1975; Archangeli & Pulleyblank 1994: §3.1.2). In grammatical terms, a gap at [ə] requires that the constraint LO→RTR be highly ranked; similarly, a gap at [ɪ ʊ] indicates that HI→ATR is dominant. Interestingly, it is normally the case that, if a language lacks [ɪ ʊ], it also lacks [ə]; i.e. there is an implicational tendency $*[ɪ ʊ] \supset * [ə]$ (J. M. Stewart 1971: 199). Crucially, this is not an absolute universal, for Lokəə (Benue-Congo, Upper Cross; SE Nigeria) has the following inventory (Williamson 1983: §6; Winston 1962):



⁵ I use the notation favoured in the literature, rather than IPA subscripts, for the sake of typographical distinctness. In diagram (4,3), for each pair of vowels of equal height and backness, the leftmost is ATR and the rightmost RTR.

These typological facts turn out to be reflected in the history of the Edoid family (Niger-Kordofanian, Benue-Congo; SW Nigeria —see Elugbe 1989). Proto-Edoid had a perfectly symmetrical 5+5 vowel system like (4,3). Phonological change, however, has removed [ə] in approximately 95% of its present-day daughters, according to a 20-language sample collected by Elugbe (1982). Additionally, around 60% of Edoid languages have lost [ɪ] and [u]. No language in Elugbe's sample has a system like that of Lokɔ̀.

(4,7) **Elugbe's (1982) 20-language sample of Edoid**

system	N	%	e.g.
(4,3)	1	5	Degema (Delta Edoid)
(4,4)	5	25	Isoko (SW Edoid)
(4,5a)	2	10	Epie (Delta Edoid)
(4,5b)	12	60	Edo (NCentral Edoid)
(4,6)	0	0	—

Archangeli & Pulleyblank (1994) note that Edoid displays the effects of convergent drift: promotion of LO→RTR has occurred in most languages, and more frequently than promotion of Hi→ATR. We know that the promotion of LO→RTR occurred separately in different members of the family, and that it is not a shared inherited trait, because Proto-Edoid [ə] shifted in different directions in different languages, merging sometimes with [a], sometimes with [ɔ], [o], [ɛ], or [e] (Elugbe 1982: 109).

Is a probabilistic explanation based on purely physical factors possible? The question can be decided experimentally. All one needs to do is generate a confusion matrix for the vowels of Degema, which retains the Proto-Edoid system. If it proves to be the case that the rate of misperception of [ə] exceeds that of [ɪ u] by an appropriate margin, then top-down grammatical selection is not necessary to account for convergence in the Edoid family. It is not every day that a problem of causation in historical linguistics can be settled in the laboratory!

4.1.3 Top-down selection of innovations in interleaved OT

Nonetheless, the probability of hypocorrection processes is not exclusively determined by physical and physiological factors. If that were the case, then a given sound structure would

have the same chance of undergoing a particular form of hypocorrection in every language. As Kiparsky (1995: 655-656) insightfully notes, however, this prediction is contradicted by the evidence of tonogenesis: a voice opposition between consonants is normally reinterpreted as a tone or register distinction only in languages which already possess a tone system, or which are subject to strong areal pressure from tone languages. In other words, the pitch variations associated with contrastive voicing resist phonologization in the absence of 'priming' from existing tone contrasts. In terms of the acquisition scenario outlined in §4.1.1, a child whose emergent grammar forbids lexical tone contrasts is unlikely to attend to the pitch variations associated with consonantal voicing distinctions.

Nonetheless, priming effects of this nature have proved extremely difficult to pin down. A notorious case is that of compensatory lengthening. De Chene & Anderson (1979) claimed that consonant deletion can only trigger 'classical' compensatory lengthening —see (3,65)— in languages where a vowel length contrast already exists; in this view, the reassociation of the floating mora to a short vowel has to be primed by the prior existence of bimoraic vowels in the language. However, Hayes (1989: 290) adduced evidence from Ilokano reduplication to show that classical compensatory lengthening can indeed occur in languages without phonemic vowel length; Kiparsky (1995: 668) mentions Morin (1992) as making similar claims for Occitan. Hayes hypothesized that the priming factor enabling compensatory lengthening was not a segmental length contrast, but a syllable weight distinction (e.g. light CV vs heavy CVC). Nonetheless, it appears that even this weakened priming requirement can be overridden. According to Lin (1997), Piro (Arawakan, Eastern Peru) has *neither* phonemic vowel length *nor* heavy CVC syllables; yet, in this language, the deletion of extrasyllabic moraic consonants (see footnote 6 in chapter 2) triggers compensatory vowel lengthening. Lin suggests that the contrast between non-moraic onset consonants and moraic extrasyllabic consonants suffices to prime compensatory lengthening.

The spectacular climb-down from de Chene & Anderson's (1979) to Lin's (1997) claims should have a sobering effect. At this stage, we simply know too little about the child's 'learning path' to determine with any precision what elements of phonological grammar are acquired early; recall that, according to the hypothesis advanced in §4.1.1, it is these bits of grammar that later have the ability to prime or inhibit the phonologization of other phonetic phenomena. It is perfectly possible that several elements of phonological grammar may exert top-down pressures on a single aspect of the primary linguistic data. Moreover, it is to be expected that top-down effects become increasingly complex and

system-specific as the grammar grows, rendering cross-linguistic generalization difficult.⁶ My hunch is that, for the time being, we should be content with observing top-down effects after the event. This is in itself no mean task, for, if properly done, recognizing a top-down effect requires that one should control for purely physical or physiological factors, as described in §4.1.2. We should look out for scenarios where a constraint ranking which has already been active for some time in the language determines the outcome of some later innovation. If we are lucky enough to find recurrent patterns, then perhaps (and only perhaps) there may be grounds to claim that the older ranking is attained relatively early in the acquisition process and 'primes' further developments.

For my purposes, however, the crucial question raised by the possibility of top-down selection in language acquisition is the following: is it the case that innovations at level L may be primed or inhibited by constraint rankings established at levels higher than L ? In chapter 3, I suggested that interleaved OT's account of the ontogeny of opaque systems is superior to that of strongly parallel OT precisely because, in a stratified grammar, innovation at low levels is untrammelled by global interactions with generalizations holding at higher levels. If, in order to account for global diachronic effects, one were to admit that high levels exert selective pressures upon the growth of the constraint hierarchy at low levels, then one would be forfeiting this theoretical advantage.

Indeed, if the acquisition of level-specific constraint hierarchies proceeds as suggested in §3.3.3.2, then it is strictly impossible that high strata should exert selective pressures upon low strata. Recall that, in our model, constraint ranking at level L_n is triggered by the input to level L_{n+1} . Crucially, the hierarchy of L_{n+1} must at this stage be fully established, for the input to L_{n+1} becomes fixed in relation to L_{n+1} rankings by Lexicon Optimization. If, as suggested by Dresher, the acquisition process is deterministic, then ranking relations at L_{n+1} cannot now be reversed. Fortunately, this does not imply that all top-down selection in interleaved OT is impossible; clearly, priming is still possible within levels. For example, once the child has established the ranking $C_1 \gg C_2$ postlexically, the phonologization of phonetic effects requiring $C_2 \gg C_1$ will be inhibited. The effect is level-internal.

⁶ Dresher (1996: 252) warns that "Cues become increasingly abstract and grammar internal the further along the learning path they are".

4.1.4 Preview of the chapter

In the following sections I shall develop the guidelines set in §4.1.3 with an analysis of a classic case of diachronic globality: the ME weight adjustment. Traditionally, the sequence of changes affecting vowel length in IOE and eME has been interpreted as a step-wise conspiracy to implement a system of prosodic constraints on syllable quantity. The prevalent approach, even on the part of Neogrammarian titans such as Karl Luick, has been brazenly teleological (Luick 1898). Modern work (e.g. Lass 1974, Ritt 1994) largely echoes Luick's ideas, with Lass (1974) putting the notion of conspiracy firmly back on the agenda.

Paradoxically, the teleological interpretation of the ME quantity adjustment depends on a conservatively Neogrammarian approach to reconstruction. The linchpin of the debate is ME Open Syllable Lengthening (henceforth MEOSL), the lengthening process which yields OE *nāma* > PDE *name* or OE *nōsu* > PDE *nose*. Traditionally, MEOSL has been thought to involve mora epenthesis, mirroring similar changes in continental West Germanic. This formulation, however, bears an uncomfortably high number of exceptions; accordingly, Neogrammarian handbooks resort to postulating extensive analogical levelling within paradigms, thereby preserving the tenet of regularity. Minkova (1982), however, has challenged the received wisdom by suggesting that MEOSL was in fact a compensatory process, triggered by the loss of stem-final schwa. Minkova's revision of MEOSL has major implications: if she is correct, then the whole teleological edifice erected by Luick (1898) collapses.

This chapter provides new empirical arguments in favour of Minkova's position. In theoretical terms, I will show that, under the compensatory reformulation of MEOSL, the development of syllable quantity in ME retains elements of diachronic globality. However, these prove innocuous, as they can be straightforwardly interpreted in non-teleological terms. Notably, I shall argue that the outcome of the loss of thematic schwa (viz. compensatory lengthening) was determined by constraint rankings involved in processes already active in the language. This result accords well with the suggestions made above in respect of top-down effects. It will be shown, moreover, that lexical diffusion propagated underlying specifications which minimized the infringement of high-ranking markedness constraints at the stem level. The argument will proceed as follows:

- section 4.2 deals primarily with issues of reconstruction;

• section 4.3 addresses the element of diachronic globality involved in the operation of MEOSL in its revised compensatory formulation;⁷

• section 4.4 deals with diffusing change in underlying representations, presenting an optimality-theoretic reinterpretation of Kiparsky's (1995) proposals.

4.2 MIDDLE ENGLISH QUANTITY: EMPIRICAL ISSUES

4.2.1 MEOSL: mora epenthesis or mora conservation?

4.2.1.1 The traditional account

Luick's (1964) classic account of the evolution of vowel length in IOE and ME is replicated with remarkable uniformity in most other twentieth-century handbooks (e.g. Brunner, 1948; Fisiak, 1968; Jordan, 1934; Mossé, 1949; Wright & Wright, 1923). He posits three *Lautgesetze*:

(4,8)

(a) Shortening before Consonant Clusters (SHOCC)

Long vowels shorten before geminates and before consonants clusters not forming a complex onset:

e.g. *sōfte* > *sōfte* 'soft'
cēpte > *cēpte* 'kept'
wræþþu > *wrāþpe* 'wrath'

(b) Trisyllabic Shortening (TRISH)

Long vowels shorten in stressed antepenultimate syllables:

e.g. *ārende* > *ērende* 'errand'
hāligdæg > *hōlidai* 'holiday'
sūðerne > *sūðerne* 'southern'

⁷ Sections 4.2 and 4.3 reproduce the content of Bermúdez-Otero (1998a) with some significant additions and revisions.

(c) Middle English Open Syllable Lengthening (MEOSL)

Short non-high vowels lengthen in stressed open antepenultimate syllables:

e.g. *nāma* > *nāme* 'name'
nōsu > *nōse* 'nose'
mēte > *mēte* 'meat'

In a vast number of cases, these sound laws predict the wrong PDE reflex for an OE vowel. To attain observational adequacy, therefore, it is necessary to postulate certain patterns of analogical levelling:

- MEOSL may fail in disyllabic stems on the analogy of trisyllabic inflected forms, where TRISH requires a short vowel (Luick, 1964: §392.2):
e.g. *hēofon* > *hēven* 'heaven', without MEOSL, on the analogy of the plural *hēofonas* > *hēvenes*; also *gānet* 'gannet', *bōdi* 'body', etc.
- Etymologically long vowels may shorten in stressed open penults by analogy with the short root-vowels created by TRISH in the trisyllabic forms of the paradigm (Luick, 1964: §387):
e.g. *hāring* > *hēring* 'herring' through levelling from the plural *hāringas* > *hēringes*; also *wāpn* > *wēpen* 'weapon', *bōsm* > *bōsem* 'bosom', etc.
- Monosyllables with VC rhymes may be affected by lengthening on the analogy of disyllabic inflected forms, where MEOSL introduces a long root-vowel (Luick, 1964: §392.1):
e.g. *hwǣl* > *whāl* 'whale' by analogy with the plural *hwālas* > *whāles*; also *blǣd* > *blād* 'blade', etc.

Levelling plays as crucial a rôle in the overall analysis as the sound laws themselves. The 'exceptions' of MEOSL (i.e. PDE forms of the *heaven*, *body* and *gannet* type) are extremely numerous (§4.2.1.2). Moreover, direct attestations of TRISH such as *ārende* > *ērende* 'errand' are virtually non-existent (§4.2.2); the exceptions of MEOSL, together with analogically restructured stems such as *hāring* > *hēring* 'herring', are the main reason for positing TRISH. Finally, analogy also takes care of the sporadic lengthening of -VC monosyllables (e.g. *hwǣl* > *whāl* 'whale').

Since these patterns of levelling work from derivative to base, they must involve restructuring of the underlying representation, rather than reflecting a synchronic misapplication effect in the grammar of ME ('priority of the base'; see §3.4.1.2). Elan

Dresher and Aditi Lahiri, in research reported in Dresher (1998), have recently put forward a hypothesis about the cause of the levelling of alternations and the restructuring of stems. They point out that the operation of MEOSL and TRISH was rendered opaque by the loss of schwa, both as a thematic and as an inflectional vowel:

(4,9) **Schwa Loss renders MEOSL and TRISH opaque**

	'name' sg. ~ pl.	'whale' sg. ~ pl.	'weapon' sg. ~ pl.
<i>UR</i>	/namə/ ~ /naməz/	/hwal/ ~ /hwaləz/	/we:pən/ ~ /we:pənəz/
<i>MEOSL, TRISH</i>	na:mə ~ na:məz	hwal ~ hwa:ləz	wɛ:pən ~ wɛ:pənəz
<i>Schwa Loss</i>	na:m ~ na:mz	hwal ~ hwa:lz	wɛ:pən ~ wɛ:pənz

Dresher (1998: abstract) claims that this opaque interaction was neither learnable nor amenable to a transparent reanalysis; as a result, children failed to acquire the alternations and generalized one stem-allomorph throughout the paradigm:

In our view, the levelling of length alternations is the result of a failure by language learners to acquire a rule, whether phonological or morphological, that could generate the alternation. This failure came about because the interaction of various Middle English processes had rendered the length alternation not only opaque, but incoherent.

Dresher and Lahiri's argument is not implausible. It is not clear, however, that it solves a fundamental problem: why was the basis for the restructured stem chosen at random, sometimes from the singular, sometimes from the plural forms? In a normal restructuring scenario, one expects an asymmetry between the alternants, with one member of the paradigm consistently prevailing over the others as the model of the restructured stem; indeed, restructuring usually favours the singular over the plural, at least in Germanic (Lahiri 1982). In this respect, the behaviour of ME appears strange.

4.2.1.2 Minkova's revision

It was largely out of dissatisfaction with this reliance on random levelling that Minkova (1982) undertook a thorough reappraisal of the empirical evidence. She gathered a corpus of 326 OE words fulfilling the traditional structural description of MEOSL, which she

supplemented with a further 107 potential inputs borrowed from Anglo-Norman before 1400.⁸ From this material, she obtained startling results:

(i) Lengthening was regular in items affected by Schwa Apocope: e.g. *nāma* > *name*. Only sixteen such items in the corpus, constituting just 4.9% of the OE material, fail to display lengthening: e.g. *drōpa* > *drop*. Of these sixteen items, five contain potentially interfering -sC- clusters: e.g. *fǣstan* > *fast* (verb) (Minkova, 1982: n. 5).⁹

(ii) Most of the items which remain disyllabic in PDE did **not** undergo lengthening: e.g. *bōdig* > *body*, *gǣnot* > *gannet*, *hēofon* > *heaven*. Of 225 unapocoped words in the entire corpus, including reflexes of both OE and Anglo-Norman material, only 36 (i.e. 16%) show lengthening: e.g. *hrǣfn* > *raven* (Minkova, 1982: 41).

Reviving an old proposal of Sarrazin's (1898), Minkova concluded that MEOSL was triggered by the apocope of schwa. Hayes (1989: 266-9) formalized this analysis within Mora Theory as a straightforward case of compensatory lengthening; see (3,62). However, both Riad (1992: 335) and Ritt (1994: 32) take Minkova to task over the 36 lengthened unapocoped forms in the corpus, for which she made no special provision; Riad dubs this the 'data problem' of Minkova's analysis. Below I will show that these items comply with a significant, and hitherto overlooked, phonological generalization (§4.2.4).

Minkova's compensatory reformulation of MEOSL has a knock-on effect on the status of TRISH, as analogy with shortened trisyllabic forms is no longer required to explain the absence of lengthening in disyllabic stems such as *heaven* or *body*. Moreover, there is only a vanishingly small number of cases where TRISH can be claimed to have applied to uninflected forms, as Ritt (1994: 103-5) observed. This leads Minkova & Stockwell (1996: 19) to state categorically that "TRISH never existed in the history of English". Crucially, alternations such as *sane* ~ *sanity*, which characterize the Romance wordstock in PDE, are not the result of historical TRISH application. Developing suggestions of Marchand's (1960), Minkova and Stockwell show that such words were not borrowed as derivationally related pairs; often, the allegedly derived item entered the language centuries before the alleged base (e.g. *procession* 1150 ~ *proceed* 1350, *sanity* 1432 ~ *sane* 1694). Thus, proparoxytones such as *sanity* and *vanity* were directly borrowed into English with a short vowel: "In this sense, it is certain that [sæ:ntɪ] never existed in

⁸ Minkova's selection of Anglo-Norman items (1982: n. 6) reproduces a list drawn up by Bliss (1952/3: §56), with four exclusions. It contains only words which have remained disyllabic in PDE.

⁹ Kim (1993: 272-3) provides a somewhat larger list of exceptions. Of course, these forms are troublesome for everyone.

English (if at all) and that it is not the historical ancestor of [sænrɪʃ]" (Minkova & Stockwell, 1996: 1).

It is important to understand that, between Luick's classic description of the changes and Minkova and Stockwell's revisionary scholarship, there has intervened a major shift in the theoretical outlook of historical linguists. Luick, together with most other handbook writers, was wholeheartedly committed to the Neogrammarian principle that all sound change is lexically regular. In the light of this principle, analogical levelling was left (once interdialectal borrowing had been discarded) as the only possible explanation of irregularities such as the 'exceptions' of MEOSL, the lengthening of *whale*-type monosyllables, and the shortening of *weapon*-type disyllables. Neogrammarian orthodoxy has since been qualified by the work of the lexical diffusionists, who have shown that some sound changes are implemented in a lexically selective fashion. As a result, analogy is no longer the only explanation for irregular change: lexical diffusion is now acknowledged as a significant source of 'residue' (Wang, 1969; Chen, 1972). In the current theoretical climate, therefore, the pattern of random analogical restructuring traditionally postulated for ME can no longer be accepted without independent corroboration. In the following sections, I will concentrate on testing Luick's analogical account of the 'exceptions' to MEOSL. Far from yielding independent support for it, this investigation will throw up important counterevidence, which is better accommodated within Minkova's compensatory analysis.

4.2.2 The circularity of TRISH

Without independent proof of the existence of TRISH, Luick's analogical account of the 'exceptions' to MEOSL becomes viciously circular. Such proof, however, is not forthcoming. The manuscript record does not directly attest to the hypothesized alternations. PDE alternations such as *sane* ~ *sanity* are, as Minkova and Stockwell point out, irrelevant, for they arise through borrowing rather than change. In turn, the evidence of shortening in disyllabic stems of the *weapon* type enlarges, but does not break, the vicious circle; it can only be attributed to TRISH on the otherwise unproven premise that lexical irregularities in the development of ME quantity arise from levelling, and not from any other plausible source such as diffusing change.

Ideally, then, we would need to find shortened PDE reflexes of OE trisyllabic stems; but these, as I pointed out above, are extremely scarce. The examples usually cited include *ǣrende* > *errand*, *sūðerne* > *southern*, and *ǣmētte* > *emmet*. Minkova & Stockwell (1996:

14) go as far as claiming that "the number of genuine (uninflected) native forms showing TRISH in OE is zero"; any OE form with the structure $\bar{\sigma}\sigma\sigma$ would have become disyllabic under syncope. Indeed, *ǣrende* may be included under this heading, as it shows up as disyllabic <errnde> in the *Ormulum*, and therefore subject to SHOCC. Syncope did not necessarily eliminate *all* uninflected trisyllables: certain words containing a sonorant consonant in the medial rhyme retain three sonority peaks under vowel syncope (e.g. *sūðerne*; see Luick, 1964: §353 Anm. 3).¹⁰ But this is a minor quibble: the fact remains that, in practice, there is no evidence of TRISH in uninflected forms.

The situation hardly alters if we take into account compound words. In cases such as *Hlāfmæsse* > *Lanmas*, *Mōnendæg* > *Monday*, and *Punresdæg* > *Pūresdæg* > *Thursday*, shortening is attributable to SHOCC. Only in a few compounds has the stressed syllable remained open: e.g. *hāligdæg* > *holiday*, *Mīcheles mæsse* > *Michaelmas*. Again, a few compounds do not warrant the setting up of a sound law with the critical attribute of Neogrammarian regularity. These forms, as well as those of the *weapon* type, can be handled just as adequately by a lexically diffusing process of Trochaic Shortening (Prince, 1990; Hayes, 1995).

4.2.3 Middle English lexical phonology: SHOCC in the *Ormulum*

The handbooks assume not only that TRISH existed once as a regular process, but also that it was still capable of triggering alternations in nominal paradigms when MEOSL kicked into action. In other words, it is presumed that TRISH remained insensitive to word morphology since its alleged inception in the eleventh century until well into the thirteenth century.¹¹ In Drescher and Lahiri's account, this presumption of morphological insensitivity extends for an even longer period. Recall that Drescher and Lahiri assume that alternations were lost, and stems restructured, when schwa dropped from inflectional endings such as plural *-es*. This happened rather late. According to Mossé (1949: §37.1), the loss of schwa in *-es* plurals only reached completion in the North around 1400. Lass (1992: 102) provides

¹⁰ With regard to regularity, the evidence of *sūðerne* > *southern* is cancelled out by that of *ēasterne* > *eastern*.

¹¹ Luick (1964: §353 Anm. 4) describes TRISH as definitely active before the end of the eleventh century. He mentions the appearance of forms such as *hālidai* alongside *hōli* in Southumbrian dialects as evidence that TRISH preceded the raising of OE /a:/ to ME /ɔ:/ (see also Jordan, 1934: §24).

metre-based evidence from the General Prologue to the *Canterbury Tales* (late 1380's): in 500-odd lines, the *-es* inflection retains its vowel in 48 out of 68 instances.

Since we have no direct attestation of the existence of TRISH, all arguments concerning its mode of application (and whether or not it was sensitive to word morphology) must *a fortiori* be inferential. However, if one assumed that TRISH did exist, there is evidence one could resort to in order to make an informed guess. Notably, the traditional assumption, energetically encouraged by Luick, is that TRISH and SHOCC were coeval and evolved in parallel; this view gathers support from the fact that the PDE alternations alleged to reflect ME TRISH and SHOCC live together at the stem level. It would therefore not be unwarranted to make a tentative extrapolation from the behaviour of SHOCC to that of TRISH.

For the precise location of SHOCC in the grammar of a South Lincolnshire dialect *circa* 1180, we have the precious evidence of Orm's holograph manuscript of the *Ormulum* (Holt 1878; for palaeographical evidence of date and provenance, see Parkes 1983). As is well known, Orm indicates vowel length in closed syllables by doubling the graph corresponding to a coda consonant when the preceding vowel is short (see Anderson & Britton 1997 for details). On this basis, I have carried out a detailed examination of the application of SHOCC in Orm's dialect; the results are tabulated in Appendix A. The evidence shows that SHOCC had already reached the stem level, and was counterfered by word-level phonology. Thus, underived stems are consistent in their compliance with SHOCC: e.g. <blosstme> 'blossom' (OE *blōstma*), <lihht> 'light' (OE *lēoht*). In contrast, consider the noun <dækenn> 'deacon' (OE *dēacon*). Addition of the plural suffix <-ess> at the word level triggers syncope of the medial vowel, but, crucially, the resulting superheavy syllable resists SHOCC: i.e. <dæcness>, not *<deccness>. Similarly, the plural of <tákenn> 'token' (OE *tācn*) is spelt <tacness>, rather than *<taccness>; cf. the spelling of <reccnenn> 'reckon' (OE *rēcenian*). The morphological conditioning of SHOCC in Orm's dialect is summarized in (4,10):

(4,10) **Morphological conditioning of SHOCC in the *Ormulum***

(a) *SHOCC-triggering constructions (stem level)*

- underived stems
- e.g. <blosstme> 'blossom' OE *blōstma*
<lihht> 'light' OE *lēoht*
- athematic past tense and past participle forms of weak verbs
- e.g. <demmdē> 'deem' pret.3sg. OE *dēman*
<hidd> 'hide' p.ptc OE *hýdan*
- *-pe/-te* suffixation
- e.g. <ma33þe> 'kin' OE *māg*
<seollþe> 'happiness' OE *sāel*

(b) *SHOCC-blocking constructions (word level)*

- nominal inflection
- e.g. sg. <dækenn> ~ pl. <dæcness>, not *<deccness>, 'deacon' OE *dēacon*
sg. <tákenn> ~ pl. <tacness>, not *<taccness>, 'token' OE *tācn*
- zero-derived weak verbs
- e.g. noun <tákenn> ~ vb. <tacnenn>, not *<taccnenn>, 'betoken' OE *tācn*
noun <wæpenn> ~ vb. <wæpnedd>, not *<weppnedd>, 'arm' OE *wāpen*
- derivational suffixation
- e.g. <god> 'good' ~ <godnesse> 'goodness', not *<goddnesse>, OE *gōd*
<soþ> 'true' ~ <soþlike> 'truly', not *<soþþlike>, OE *sōþ*
- compounding
- e.g. <boc> 'book' ~ <bocstaff> 'letter', not *<bocstaff>, OE *bōc*
<shep> 'sheep' ~ <shephirde> 'shepherd', not *<shepphirde>, OE *scēap*

The implications of these data are striking: the major lineaments of the stratal organization of PDE morphology were already in place in a South Lincolnshire dialect by the end of the twelfth century!

If, as Luick suggests, TRISH developed in parallel with SHOCC, then we are led to surmise that its application became circumscribed to the stem level rather quickly. But if that was so, then TRISH could not have triggered the alternations that are supposed to

account for the exceptions of MEOSL. MEOSL had not even started in Orm's dialect when SHOCC had already reached the stem level.¹²

4.2.4 Lengthening in unapocopated disyllables

The levelling hypothesis can also be tested by looking for unexpected bias in the distribution of the 'exceptions' to MEOSL. Dobson (1962: 126-7) seems one of the first to have realized that certain well-known distributional trends are anomalous in the context of an analogical account. First, MEOSL-candidates containing the ending *-ig* regularly escaped the change: e.g. *bōdig* > *body*.¹³ Secondly, unpredictability in the length of the root-vowel is particularly acute among MEOSL-candidates whose second syllable contains a nasal or liquid consonant in the rhyme: e.g. *hēofon* > *heaven* vs *hræfn* > *raven*. These tendencies appear to be phonological in nature; Sweet (1888: §629) spoke of 'back-shortening' before *-er*, *-el*, *-en*, *-ing* and *-i*. Dobson observes that, if the exceptions to the traditional statement of MEOSL had a purely paradigmatic cause, one would not expect them to display any significant phonological bias.

In other words, the handbooks suggest that a MEOSL-candidate is vulnerable to analogical shortening whenever it has an imparisyllabic paradigm, i.e. it has trisyllabic inflected forms subject to TRISH. Luick accounts for the regularity of lengthening in apocopated words of the *name* type through the fact that, prior to the loss of stem-final *-e*, these words were disyllabic throughout their inflection, and so lacked forms targeted by TRISH: e.g. sg. *nāme* ~ pl. *nānen/nānes*. The class of MEOSL-candidates vulnerable to analogical levelling thus coincides with the class of unapocopated disyllabic stems: e.g. sg. *hēven* ~ pl. *hēvenes*. If MEOSL-failure was caused by paradigm levelling, exceptions should be randomly scattered across this vulnerable group, regardless of the segmental

¹² That MEOSL had not yet begun in Orm's dialect is now a well-established fact, thanks to the work of Fulk (1996) and Anderson & Britton (1997). Incidentally, the absence of MEOSL in the *Ormulum* is fully compatible with Minkova's compensatory formulation of the process: although final *-e* was being lost as an oblique case ending, it remained fairly stable as a stem-vowel (Fulk, 1996: 492).

¹³ According to Morsbach (1896: 94), *-ig* bore secondary stress, which prevented lengthening. However, Luick (1964: §392 Ann. 1) has convincingly shown that the assumption of secondary stress in this case is groundless and *ad hoc* (see further Minkova, 1982: 30). From a phonological viewpoint, moreover, the structure *ōð* should favour, rather than inhibit, lengthening. In contemporary American English, for example, secondary stress on the last syllable of *Arab* has been observed to trigger lengthening of the tonic vowel, clearly in order to maintain foot binarity: /æɹəb/ ~ /eɹ,ɹæb/. This is the so-called 'Arab rule' (Fidelholtz, 1967; and see Pater, 1995: §1.1, for an OT discussion).

makeup of its individual members. How far the actual pattern of exceptions departs from this expected distribution can now be accurately assessed through Minkova's (1982) corpus.

Analysis yields striking results; see Appendix B, summarized below as table (4,11). Variation turns out to be strictly confined to those words whose post-tonic syllable contains a rhymal sonorant consonant, i.e. to words of the *heaven* and *raven* types. If an unapocopated disyllabic stem fails to meet this segmental condition, then lengthening is predictably absent. As was already known, items whose post-tonic rhyme contains an unchecked non-deleting vowel (the *body* type) have invariably escaped lengthening. However, items whose post-tonic rhyme consists of a vowel + obstruent sequence (the *gannet* type) also display short root-vowels. In the latter case, Minkova's corpus contains just one solitary counterexample: OE *nācod* > *naked*.¹⁴

(4,11) Incidence of MEOSL among unapocopated disyllabic stems

Total number of items: 231

(a)	The post-tonic rhyme contains a sonorant consonant.....	166 items		
		N	%	e.g.
(i)	Lengthened:	39	23.5	<i>raven</i>
(ii)	Unlengthened:	127	76.5	<i>heaven</i>
(b)	The post-tonic rhyme contains an unchecked vowel.....	24 items		
		N	%	e.g.
(i)	Lengthened:	0	0	—
(ii)	Unlengthened:	24	100	<i>body</i>
(c)	The post-tonic rhyme consists of vowel plus obstruent.....	41 items		
		N	%	e.g.
(i)	Lengthened:	1	2.4	<i>naked</i>
(ii)	Unlengthened:	40	97.6	<i>gannet</i>

¹⁴ Minkova's corpus also includes the noun *process*, but this item must be discarded since it has remained bipedal in PDE: /ˈprəʊ,ses/. According to Wells (1990: 559), "the *-is*, *-əs* forms seem to arise only as occasional weakenings of a basic *-es* pronunciation." Since we therefore have a *ōð* structure, the length of the tonic vowel can be put down to the 'Arab rule' (see previous footnote).

It is out of the question that this biased distribution should be a mere accident. A simple statistical test reveals that the positive correlation between lengthening and the presence of a sonorant consonant in the post-tonic rhyme attains a very high level of significance, with $\chi^2=12.973$ and $p<0.0005$:

(4,12) Correlation between lengthening and presence of a sonorant in the post-tonic rhyme¹⁵

		Lengthening			
		YES		NO	
		observed	expected	observed	expected
Sonorant C in post-tonic Rh	YES	35	25.5	123	132.5
	NO	1	10.5	64	54.5

In view of this correlation, we must conclude that the traditional account of the 'exceptions' to MEOSL is either simply wrong, or else in need of phonological qualification. Luick's proposals could be salvaged by retreating into an empirically weaker position, where the alleged allomorphic alternations between disyllabic and trisyllabic inflectional forms are regarded as merely creating the *possibility* of levelling; the actual incidence of analogical shortening would then be determined by other factors.¹⁶ However, this manoeuvre faces a serious obstacle. One would have to explain why levelling from inflected forms was particularly favoured (indeed compulsory) in two types of word which seem to have nothing in common: the *body* class and the *gannet* class. This avenue is likely to lead to a disjointed heap of *ad hoc* stipulations.

We have found, in sum, that MEOSL complies with the following generalizations:

¹⁵ The test excludes four etymological doublets.

¹⁶ I am grateful to an anonymous *ELL* reviewer for suggesting this approach: "An interesting issue is raised here as to whether analogical levelling could display a phonological bias. [...] One could argue that it is precisely in such conditions —i.e., where a selection has to be made between variants, and there is no rule for doing so— that all kinds of factors, including phonological ones, could become relevant, and we might expect to find local subregularities in an otherwise random set."

(4,13) MEOSL

- (a) Lengthening is regularly present in items affected by apocope of stem-final -ə.
- (b) Lengthening applies unpredictably among unapocopated disyllables containing a sonorant consonant in the post-tonic rhyme.
- (c) Lengthening is regularly blocked in unapocopated disyllables lacking a sonorant consonant in the post-tonic rhyme.

This summary indicates that lengthening was conditioned by the nature of the post-tonic syllable to an even greater extent than Minkova (1982) envisaged; more importantly, it suggests, as I shall argue presently, that the compensatory interpretation of MEOSL is correct in essence.

Let us assume, in accordance with the compensatory hypothesis, that ME grammar forbade lengthening through mora-insertion, but allowed lengthening through mora-transfer. This would imply that DEP^h enjoyed a high position in the constraint hierarchy, dominating antagonistic constraints such as PKPROM; see (3,58) and (3,75). Under such circumstances, lengthening in apocopated disyllables would be regular because stem-final schwas were deleted without exception; see §4.3. In contrast, lengthening in words such as *bōdi* and *gānet* would be prevented by the absence of a mora-donor. In these forms, deletion of the unstressed vowel is either ungrammatical (e.g. **bod*), or it yields a monosyllable closed by a SHOCC-triggering consonant cluster (e.g. OE *mūnuc* > ME *mūnec* > PDE *monk*).

When the post-tonic rhyme contains a nasal or a liquid, however, a different possibility arises. In PDE such rhymes allow more or less free variation in the nucleus between a syllabic sonorant and schwa. This state of affairs has been a persistent feature in the history of the language: so-called parasiting can be observed in the earliest OE texts (e.g. *hræfn* vs *hræfen*; Campbell 1959: §363) and is known to have had important repercussions on OE metre (see Fulk 1992: ch. 1). One can therefore safely assume that unstressed syllables allowed variation between nuclear sonorants and schwa also in ME.

In this light, consider a ME child in whose mental lexicon *rāven* is represented as /rævən/. Such a child will have no difficulty in recognizing both [rævən] and [ravn̩] as

ordinary allophonic variants of this lexical item. What happens, however, if, through an accident of performance, the same child is exposed to a stimulus [ra:vŋ]? Given the fact that compensatory lengthening is active elsewhere in the language, [ra:vŋ] may be parsed as an acceptable match of input /ravən/: provided that the tonic vowel is seen to have attracted the mora delinked from the underlying schwa, the match will satisfy DEP^u. In other words, this particular form of hypocorrection is facilitated by the fact that the child is primed to associate schwa loss with compensatory lengthening.

The effect was never likely to become systematic because the trigger was not robust, but highly variable. To the extent that purely physical factors may have been involved, however, one can make certain predictions concerning the incidence of lengthening (see §4.1.2): notably, the frequency with which vowels are misanalysed as lengthened will increase in direct proportion to their intrinsic duration. Since low vowels are intrinsically longer than high ones (Lehiste 1970), we expect that, within the class of ME disyllabic stems containing a sonorant in the post-tonic rhyme, /a/ will lengthen with greater frequency than either /e/ or /o/. The data in Appendix B (§B.1) bear this prediction out: /a/ lengthens in 34.9% of cases (30 out of 86), /o/ in 12.9% of cases (4 out of 31), and /e/ in 10.2% of cases (5 out of 49); see graph in Ritt (1994: 39). The positive correlation between /a/ and lengthening is very highly significant, with $\chi^2=11.598$ and $p<0.0005$.

4.2.5 Middle English in a West Germanic perspective

Lahiri (1995) puts forward an interesting argument in favour of the traditional approach to MEOSL. She suggests that Minkova's compensatory reformulation destroys the uniformity of West Germanic developments. According to Lahiri, OSL operated in exactly the same way in ME, MHG and Middle Dutch. Differences in the distribution of long vowels across the West Germanic languages arose only through interference from other changes such as the *hochdeutsche Lautverschiebung* and TRISH, and were amplified by variable patterns of analogical levelling: "Thus, although open syllable lengthening was found in all the West

Germanic languages in the middle period, due to language independent interactions and levelling, cognate words can easily have different vowel quantity" (Lahiri 1995: §47).¹⁷

This argument is not persuasive. First, it is factually wrong. OSL did not occur in all the West Germanic dialects: High Alemannic did not participate in the change. Thus, the typical paradigm of short *a*-stems in High Alemannic is represented by *grāb* ~ *grēber* 'grave' (OE *græf*); see e.g. Hotzenköcherle (1986), Keller (1961: ch. 2), Ritzert (1898: §1).

Secondly, English is not the only Germanic language where analogical accounts of irregular vowel length developments have come under attack. In Danish, monosyllables with VC rhymes underwent sporadic lengthening, in the same way as the English *whale* class. Since Danish has genuine OSL, rather than mere compensatory lengthening, the analogical explanation of this irregularity (Boberg 1896) would in this case appear to be cast-iron; but Hansen (1962: 389) presents counterevidence. As Riad (1992: 355) reports,

Hansen notes that several original CvC forms that have lengthened their vowel could not have become that way analogically, for lack of bisyllabic forms to model on. Typical cases are nouns with zero plurals, e.g. *mæd* 'food', *loof* 'permission', *taak* 'roof', *høer* 'flax', *øel* 'beer', *hool* 'hole'. The definite article [suffix; RBO], which would render these words bisyllabic, was only in severely restricted use at this time. Hence, analogy alone could not explain the situation.

4.2.6 Summary

In the previous sections, I have shown that Luick's analogical account of the 'exceptions' to MEOSL suffers from serious empirical inadequacies:

First, there is no independent evidence that TRISH created the required allomorphy. There are practically no instances of TRISH in uninflected forms to break the circularity of the argument. Moreover, the morphophonological behaviour of SHOCC, as manifested in the *Ormulum*, indicates that it is highly unlikely that TRISH would have remained morphologically-insensitive until the thirteenth (or, *a fortiori*, the late fourteenth) century.

¹⁷ Lahiri's idea that West Germanic underwent convergent drift is actually consistent with a long and venerable tradition inspired by so-called 'Prokosch's Law' (Prokosch 1938: 140). According to Lahiri (1995), this alleged West Germanic convergence was fuelled by the properties of the so-called 'Germanic foot' (Dresher & Lahiri 1991); but Hutton (1996: 9-10) has convincingly argued that the Germanic foot is an unnecessary addition to universal foot inventories (cf. e.g. Hayes 1995).

Minkova & Stockwell's (1996) skepticism regarding the very existence of TRISH as a sound change is justified.

Secondly, Luick's paradigm-based approach fails to make sense of the phonological conditioning of lengthening in unapocopated disyllables. An investigation of Minkova's (1982) corpus of MEOSL-candidates has revealed that unapocopated disyllabic stems could undergo lengthening only if their post-tonic rhyme contained a sonorant consonant, which alternated freely with schwa in the syllable nucleus. I have shown how this fact can be incorporated into a compensatory analysis of MEOSL: as the child was primed to associate schwa loss with compensatory lengthening, hypocorrected forms could contingently arise.

Finally, claims that OSL was a Pan-West-Germanic phenomenon are factually wrong. They are also vitiated by the difficulties incurred by levelling analyses in languages other than English.

I conclude that the compensatory formulation of MEOSL is correct. Alternative non-analogical accounts of anomalous *whale*-type and *weapon*-type words must be sought. For the former, I sketch some suggestions in section 4.4.

4.3 GLOBALITY IN MEOSL

4.3.1 Teleology

In their traditional formulations, SHOCC, TRISH and MEOSL dovetail in a striking way; they create a powerful impression of purposeful historical momentum. It is therefore not surprising that a linguistic tradition has emerged where they are analysed together as a complex of functionally interrelated changes. In this tradition, vowel length in stressed syllables is regarded as having remained largely faithful to its etymological origins until IOE. At that point, the prosody of the language was reorganized upon new principles whereby vowel length came to depend on certain syllabic and accentual preferences (see e.g. Mossé 1952: §§18-9). The *locus classicus* in this line of research is Luick (1898). This article, now over a century old, has cast a long shadow upon all subsequent approaches to IOE and ME quantity. Thus, Luick deserves the credit of having single-handedly manufactured the two most important 'objects' of English historical phonology: the Great Vowel Shift (see Lass, 1997: 32-40) and the ME length adjustment.

However, Minkova's (1982) compensatory reformulation of MEOSL and Minkova & Stockwell's (1996) dismissal of TRISH go against the very grain of this tradition (cf. Lahiri 1995, Lass 1992, Riad 1992, Ritt 1994). If one takes Minkova and Stockwell's proposals on board, the idea that ME quantity was heading in some particular direction loses much of its appeal. In §4.3.1.1 and §4.3.1.2, I show how the empirical results of section 4.2 put the traditional theoretical outlook in jeopardy. I shall show that global connections can indeed be drawn between the different phenomena involved in the ME quantity adjustment; however, these connections need not be understood in teleological terms, for they simply involve the parallel interaction of old and innovative constraint rankings in the lexical phonology.

4.3.1.1 Luick's *Normalmaße*

Luick (1898) classifies syllables into three quantitative grades (*Quantitätsstufen*): light or monomoric (*Stufe 1*), heavy or bimoric (*Stufe 2*), and superheavy or trimoric (*Stufe 3*). He suggests that, in the transition from OE to ME, stressed syllables were assigned a standard or optimal measure (*Normalmaß*). SHOCC, TRISH and MEOSL were triggered in order to eliminate deviations from this optimal quantity. The standard quantitative grade assigned to a stressed syllable was inversely proportional to the number of unstressed syllables immediately following it within the word; see (4,14). Hence, in the absence of secondary stress (i.e. "in phonetisch einfachen Wörtern"), monosyllables were optimally superheavy, disyllables were optimally headed by a heavy syllable, and trisyllables were optimally headed by a light syllable (Luick, 1898: 336).

(4,14) The inverse proportion rule in Luick's *Normalmaß* system

No. of syllables in the word:	1	2	3
	†	†	†
<i>Quantitätsstufe</i> of the head:	3	2	1

Luick (1898: 337) encounters serious empirical difficulties with stressed monosyllables. SHOCC repaired those words which exceeded the requirements of a grade-3 syllable: e.g. *fȳlð* > *fȳlð* 'filth'. There was, however, no concomitant sound law raising the quantity of -VC monosyllables to their alleged *Normalmaß*; Luick argues that the final consonant was geminated in ME, but there is no evidence to suggest that this was anything other than a purely orthographic phenomenon (e.g. OE *Gōd* spelt <Godd> in the

Ormmulum). Similarly, unchecked monosyllables could only have attained grade 3 by means of a paragogic consonant or an overlong vowel; here, Luick makes a half-hearted attempt to conjure up some *ursprünglich überdehnung* on the basis of present-day allophonic variation.

In polysyllabic words, Luick's hypothesis appears to work neatly if one adheres to the traditional formulations of MEOSL and TRISH; but these have now proved inadequate.

4.3.1.2 Lass: the Great Length Conspiracy

Lass has addressed the IOE and ME vowel length changes in numerous publications (1974, 1980, 1985, 1987, 1992, 1997). His *leitmotif* has been the idea that these changes constitute an episode within a larger development: the Great Length Conspiracy (Lass 1974: 326-334, 1987: 159-161, 1992: 72-76). Lass (1974) defined 'conspiracy' as a sequence of chronologically discrete and superficially unrelated changes whose cumulative effect is the enactment of a relevant synchronic generalization.¹⁸ According to Lass, the various changes incorporated into the Great Length Conspiracy can be grouped on the basis of two common effects: "reducing the number of environments in which vowel length was contrastive; and tending to stabilise certain syllable shapes as 'preferred' or 'optimal'" (Lass 1992: 70). In Lass (1974: 315, 333), it was suggested that these effects may be captured by means of higher-order instructions or 'metarules' (see also Lakoff 1972). Thus, at the most abstract level, the Great Length Conspiracy would obey a metarule of the type "Maximise the predictability of vowel length" (Lass 1974: 333); at a more concrete level, the changes tend towards the implementation of a Luickian *Normalmaß* system (see table 32 in Lass 1992: 75-76).

Lass is aware that postulating 'metarules' poses more questions than it answers, but my arguments against his approach will be empirical, rather than methodological. I shall argue that, in view of the results of section 4.2, Lass's metarules require violence to the facts. It is intriguing that Lass should take a sympathetic view of Minkova's reformulation of MEOSL: "This requires more research, but there is no doubt that Minkova's solution, rather than the traditional one, is essentially right" (Lass 1992: 74; see also Lass 1985). As I will show presently, the compensatory approach to MEOSL is incompatible with Lass's description of the Great Length Conspiracy.

¹⁸ The terms 'teleology', 'orthogenetic series', 'drift' and, most recently, 'chreod' (Lass, 1997: §6.3.5) all refer to the same empirical construct, although they evoke subtly different views of the underlying mechanics.

First, let us consider MEOSL from the viewpoint of the metarule "Maximise the predictability of vowel length". I have shown that the input to MEOSL is invariably a disyllabic word with a light stressed syllable whose post-tonic syllable contains a deletable schwa. In the vast majority of cases, where the input undergoes deletion of stem-final -ə, the output is a -VVC monosyllable: e.g. /namə/ ↔ [na:m]. Only in a few instances does the output remain disyllabic: e.g. /ravən/ ↔ [ra:vŋ]. Consequently, MEOSL did not neutralize vowel length distinctions anywhere at all. After MEOSL applied, there remained a plentiful stock of disyllabic words with light stressed syllables, including the *body* and *gannet* classes; these contrasted with items whose tonic syllable was heavy (e.g. words like PDE *token* < OE *tācn*). Similarly, the new -VVC monosyllables (e.g. *name*) merged with an already existing phonological class, which contrasted with -VC monosyllables (e.g. *god*). Length distinctions were retained even among unapocopated disyllables whose post-tonic syllable contained a sonorant consonant: here, the irregularity of the hypocorrection effect prevented neutralization (e.g. *heaven* vs *raven*). Lass's (1992: 74) claim that the strong syllable of a disyllabic foot was a neutralizing environment is thus factually wrong.

Let us now turn to Lass's espousal of Luick's *Normalmaß* idea. Did MEOSL tend to "stabilise certain syllable shapes as 'preferred' or 'optimal'"? In this connection, Lass (1992: 74) argues that Minkova's compensatory reformulation of MEOSL "makes very little difference". He contends that, "whatever 'open-syllable lengthening' was, its effects fit into the conspiracy" (Lass 1987: 173; italics in the original). This is because "the 'aim' of the change is to substitute heavy for light in the strong syllable of a foot"; thus, "the overall quantitative effect (light > heavy)" remains unaffected whether one follows the traditional version of MEOSL or its compensatory reformulation (Lass 1992: 74).

This escape manoeuvre is obviously fallacious, inasmuch as Lass resorts to a selective use of the empirical evidence. The implication is that the metarule "Eliminate light stressed syllables", or some such, can be invoked *a posteriori* when a phonological process happens to fit the bill. The fact that the metarule is profusely violated elsewhere, as in the *body*, *gannet* and *heaven* classes, is declared irrelevant. As in Vennemann (1983), the teleological interpretation sits atop the actual phonological processes, invulnerable to empirical challenge. Prince & Smolensky (1993: 198) comment acidly on such practice. It is disconcerting that Lass (1987, 1992) should use this ploy, since Lass (1980) clearly diagnoses its specious nature.

The following pages explore the connections between SHOCC and MEOSL within an optimality-theoretic framework. It will become apparent that both changes are structurally linked in a rather simple, non-teleological way, through the rôle of a constraint requiring final consonant extrasyllabicity. This constraint exempts -VVC monosyllables from SHOCC, and forces morae cast adrift by deleted stem-final schwas to redock onto the root-vowel. We will see that MEOSL had nothing to do with PKPROM.

4.3.2 SHOCC

SHOCC has received a great deal of indirect attention from synchronic theorists because some of the allomorphic alternations it created have survived into PDE. The consensus view holds that SHOCC is a form of closed syllable shortening (Myers 1987; Prince & Smolensky 1993; Rubach 1996; Stampe 1972). Closed syllable shortening is driven by $*[\mu\mu\mu]$, a prosodic maximality constraint prohibiting superheavy, or trimoric, syllables (Prince & Smolensky 1993: 210-1). $*[\mu\mu\mu]$ is activated by the ranking $*\text{BRANCH}^{\mu} \gg *_{\mu}/\text{C}$, which forbids weightless coda consonants; see §2.2.1, particularly (2,10) and (2,11).

(4,15)

	PARSE ^μ	PARS ^{3σ}	*BRNCH ^μ	MAX ^{3σ}	*[_σ μμμ]	*μ/C	MAX ^μ
$\begin{array}{c} \mu \mu \quad \mu \\ \quad \\ k \quad e \quad p \quad t \quad \emptyset \end{array}$							
$\begin{array}{c} \sigma \\ / \quad \quad \backslash \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ k \quad e \quad p \quad t \quad \emptyset \end{array}$	*!					*	
$\begin{array}{c} \sigma \\ / \quad \quad \backslash \\ \mu \quad \mu \quad \mu \\ \quad \quad \\ k \quad e \quad p \quad t \quad \emptyset \end{array}$		*!					
$\begin{array}{c} \sigma \quad \sigma \\ / \quad \backslash \quad / \quad \backslash \\ \mu \quad \mu \quad \mu \quad \mu \\ \quad \quad \quad \\ k \quad e \quad p \quad t \quad \emptyset \end{array}$			*!				
$\begin{array}{c} \sigma \quad \sigma \\ / \quad \backslash \quad / \quad \backslash \\ \mu \quad \mu \quad \mu \quad \mu \\ \quad \quad \quad \\ k \quad e \quad t \quad \emptyset \end{array}$				*!			
$\begin{array}{c} \sigma \quad \sigma \\ / \quad \backslash \quad / \quad \backslash \\ \mu \quad \mu \quad \mu \quad \mu \\ \quad \quad \quad \\ k \quad e \quad p \quad t \quad \emptyset \end{array}$					*!	*	
$\begin{array}{c} \sigma \quad \sigma \\ / \quad \backslash \quad / \quad \backslash \\ \mu \quad \mu \quad \mu \quad \mu \\ \quad \quad \quad \\ k \quad e \quad p \quad t \quad \emptyset \end{array}$						*	*

In ME, monosyllables with VVC rhymes escaped closed syllable shortening because the final consonant was extrasyllabic. This is independently confirmed by the facts of stress assignment, both in ME (Minkova 1997: §6.1) and PDE (Hayes 1982). Final consonant extrasyllabicity is enforced by the ranking WEAKC \gg PARSE^{3σ}; see §2.2.1.

(4,16)

lif	*[_σ μμμ]	WEAKC	PARSE ^{Seg}	MAX ^μ
[_σ [_σ lif]]	*!	*!		
[_σ [_σ lif]]		*!		*
[_σ [_σ li:]f]			*	

Unrestricted final consonant extrasyllabicity would have undesirable side-effects, for it would render -VC monosyllables monomoric and, hence, subminimal: e.g. *[_σ[_σva]t]. Such conflicts between extraprosodicity and word minima are instances of a widespread 'Unstressable Word Syndrome', whose resolution varies from language to language (Hayes 1995: 110-3). A number of members of the Germanic family opt for vowel lengthening, notably the Scandinavian languages and High Alemannic: e.g. Icelandic *mǎ<n>* → *mā<n>* 'remember' 1sg.pres. (Kiparsky 1984). In contrast, ME chose to cancel final consonant extrasyllabicity in VC monosyllables: e.g. [_σ[_σvat]]. In the context of standard parametric theory, however, extraprosodicity cancellation constituted an anomaly: parameters are, in principle, 'all-or-nothing affairs' (Mester, 1994: 16). In OT, in contrast, such 'non-uniformity' effects (Pater, 1995) follow naturally from run-of-the-mill constraint interaction: like any other constraint, WEAKC can be violated under the compulsion of higher-ranked constraints (Prince & Smolensky, 1993: 43-7).

The Unstressable Word Syndrome is triggered by a strict ban on monomoric lexical words, which arises when FTBIN and LX≈PR are undominated.¹⁹ The resolution of the syndrome depends on the ranking of DEP^μ. The Icelandic option is generated by the ranking WEAKC » DEP^μ, the vowel is lengthened at the expense of DEP^μ so that FTBIN can be fulfilled whilst maintaining final consonant extrasyllabicity. In ME, in contrast, DEP^μ outranks WEAKC, forcing the syllabification of the final consonant.

¹⁹ FTBIN requires that all feet should be binary either at the moric or at the syllabic level (Prince 1990: 360). In strongly parallel OT, LX≈PR(Word) is a constraint on the morphology/phonology interface requiring that every lexical word should correspond to a prosodic word (Prince & Smolensky, 1993: 43). Since, by Proper Headedness (2.4), every prosodic word must in turn contain at least one foot, LX≈PR(Word) causes the requirements of FTBIN to percolate to content words. Under interleaving, I assume that LX≈PR simply requires that the visible domain should be parsed by at least one *σ*, see §3.4.1.1 for reference to morphological categories by phonological constraints.

(4,17)

va ^μ t	FTBIN	DEP ^μ	WEAKC	PARSE ^{Seg}
[_σ [_σ va ^μ]t]	*!			*
[_σ [_σ va ^{μμ}]t]		*!		*
[_σ [_σ va ^μ t ^μ]]			*	

The lengthened candidate violates DEP^μ because the inserted mora neither has an input correspondent nor qualifies as positional *μ*-licenser. In the winning candidate, the final consonant is solely licensed by its moraic host, which accordingly qualifies as a positional *μ*-licenser.²⁰

Old VVCC monosyllables undergo SHOCC because WEAKC demands the extrasyllabicity of only one consonant. In tableau (4,18) I assume that PARSE^{Seg} plays the crucial rôle of rejecting candidates with two extrasyllabic segments. The job could also be done by the constraint WLWF (for WEAKLAYERINGWELL-FORMEDNESS), which states that two phonological units may not be adjacent if both fail to be licensed by the next higher category in the prosodic hierarchy: i.e. no two adjacent unfooted syllables, no two adjacent extrasyllabic morae (Lin 1997: 420).

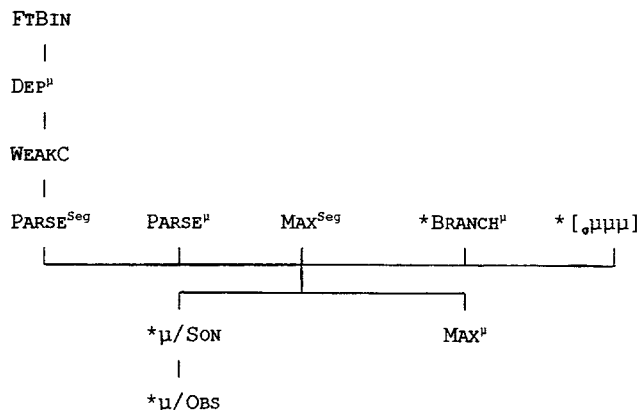
(4,18)

fy:lθ	*[_σ μμμ]	MAX ^{Seg}	WEAKC	PARSE ^{Seg}	MAX ^μ
[_σ [_σ fy:lθ]]	*!		*		
[_σ [_σ fy:l]θ]	*!			*	
[_σ [_σ fy:]θ]		*!		*	
[_σ [_σ fy:l]θ]				**!	
[_σ [_σ fy:l]θ]				*	*

²⁰ The relative ranking of FTBIN and DEP^μ is not crucial to the example at hand. However, FTBIN must dominate DEP^μ to capture alternations such as unstressed *bū* vs stressed *bū* 'thou'.

In sum, the application of SHOCC was regulated by the following hierarchy:

(4,19) SHOCC



In Orm's dialect, this ranking holds only at the stem level; at the word level, $*[\text{o}\mu\mu\mu]$ is demoted relative to MAX^μ (see §4.2.3). I assume that all other dominance relationships remain unaltered (at least in the lexical phonology), thereby maximizing inter-level faithfulness.

4.3.3 The global interaction of schwa loss and consonant extrasyllabicity

MEOSL was triggered by the presence of a deletable input schwa in the post-tonic rhyme. The irregular hypocorrection mechanism involved in cases such as *hræfn* > *raven* was discussed in §4.2.4; this section will focus on words affected by the loss of stem-final -ə: e.g. *nāma* > *name*. Schwa loss involved the promotion of a constraint against placeless vowels over MAX^{seg} . At the moraic level, however, I will show that the outcome of deletion, i.e. compensatory lengthening of the stem vowel, can be predicted on the basis of rankings independently established for SHOCC. Thus, the older rankings globally determine the outcome of the innovative process of schwa deletion. This global interaction is level-internal and causal rather than teleological; see §4.1.3.

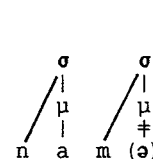
In IOE or eME schwa emerged as a placeless reduction vowel; I will therefore assume that schwa deletion was driven by a constraint $*\text{EMPTYV}$, which requires that

output representations should not contain vowels lacking oral features. Of course, $*\text{EMPTYV}$ can trigger the deletion of an input vowel position only if it dominates the faithfulness constraint MAX^{seg} .²¹ We know that deletion in stem-final position predated deletion in checked inflectional endings such as *-ed* and *-es*, and so we must infer that, during ME, $*\text{EMPTYV}$ underwent crucial reranking with respect to other constraints on syllable structure; but the details are not relevant to my purposes here.

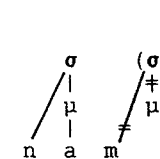
Diagram (4,20) offers a derivational representation of the mora-transfer process whereby the loss of stem-final schwa triggered lengthening of the root-vowel (for 'parasitic delinking', see Hayes 1989: 268):

(4,20)

Nucleus deletion



Parasitic delinking



Mora relinking



There seems to be a problem in this chain of operations: after parasitic delinking, why does the floating mora skip the final consonant and dock onto the tonic vowel? Concern over this step has been independently voiced by Bickmore (1995: 148) and Riad (1992: 335-6), the latter quoting Lahiri (1991); Riad calls this 'the analytical problem' of the compensatory formulation of MEOSL. Hayes invokes Itō's (1986) proposal that prosodic structure is erected maximally, but Riad responds that default mora-to-segment association is one-to-one (as per the Association Convention; see e.g. Goldsmith, 1990: 14). It is obvious, however, that the solution cannot be a matter of universal principle, since, cross-linguistically, vowel apocope sometimes triggers compensatory lengthening, and sometimes does not. As the analysis stands, the last step in (4,20) has to be stipulated.

In an optimality-theoretic analysis, however, this difficulty vanishes. The floating mora must dock onto the root-vowel because, through extrasyllabification, the final consonant becomes unavailable as a landing site. Crucially, the extrasyllabicity of the final

²¹ $*\text{EMPTYV}$ covers phenomena similar to those addressed by the Empty Category Principle in Government Phonology. The Empty Category Principle describes the licensing conditions under which the 'cold vowel' is phonetically unrealized (Kaye, 1990: 313-4). In her analysis of certain aspects of the historical behaviour of German schwa, Raffelsiefen (1996) posits a constraint $*\text{SCHWA}$.

consonant need not stipulated: it is required by already existing constraint rankings, established independently in the analysis of SHOCC. When one places *EMPTYV on top of MAX^{Seg} in (4,19), all the rest follows automatically. Here is the demonstration:

- Given the ranking *EMPTYV » MAX^{Seg}, stem-final -ə deletes, casting its own mora adrift.
- With WEAKC dominating PARSE^{Seg}, the preceding consonant, now word-final, occupies a preferably extrasyllabic position.
- Since superordinate LX≈PR and FTBIN ban monomoric lexical items, the floating mora is compelled to reattach to the only available intrasyllabic landing site: the root-vowel.
- The lengthening of the root-vowel respects DEP^M because the weak mora was already present in the input as the host of stem-final -ə.

The reassociation of the floating mora could be imputed to MAX^M. As it turns out, the satisfaction of MAX^M is not crucial. As tableau (4,21) demonstrates, the superordinate constraints FTBIN, *[_σμμμ] and *EMPTYV narrow the pool of viable parses down to two: a [_σ[_σCV^MC^M]] and a [_σ[_σCV^M]C] monosyllable. Since both candidates fulfil DEP^M and tie on MAX^{Seg}, the final decision is down to WEAKC.

In this light, the discussion of diachronic globality in §4.1.3 leads one to advance the following conjecture: IOE and eME children were able to acquire the high ranking of WEAKC early in the acquisition process, through robust cues from SHOCC and stress assignment; as schwa instability set in, this ranking primed lengthened apocopated variants over unlengthened ones. Thus we have diachronic globality without teleology.

(4,21) MEOSL

	FTBIN	*[_σ μμμ]	*EMPTYV	DEP ^M	MAX ^{Seg}	WEAKC	PARSE ^{Seg}	MAX ^M
/na ^M ne ^M /								
[_σ [_σ na ^M] _σ ne ^M]]			*					
[_σ [_σ na ^M] _σ]m]	*				*		*	*
[_σ [_σ na ^M m ^M]]		*			*	*		
[_σ [_σ na ^M m ^M]]					*	!*		
[_σ [_σ na ^M]m]					*		*	

N.B. The mora associated with [m] in [_σ[_σna^Mm^M]] is a positional μ-licenser; hence, it fulfils DEP^M.

4.4 GLOBALITY IN LEXICAL CHANGE AND STEM-LEVEL DEFAULTS

Luick (1898) regarded his *Stufe 3* as the *Normalmaß* of ME stressed monosyllables. As we saw in §4.3.1.1, this claim ran into trouble over the retention of -VC monosyllables throughout the period: e.g. *gōd*, *vāt*, etc. In our optimality-theoretic analysis, the source of Luick's problem becomes clear: high-ranking DEP^h prohibited mora insertion. Nonetheless, close inspection of tableaux (4,16) and (4,17) reveals that an asymmetry does exist between /CV^hC/ and /CV^h^uC/ stems. In relation to the stem-level constraint hierarchy given in (4,19), the optimal parse of a /CV^hC/ input form, viz. [⌊_oCV^hC^h⌋], proves *less harmonic* than the optimal parse of input /CV^h^uC/, viz. [⌊_oCV^h^u⌋C]. In the former, FTBIN and DEP^h block final consonant extrasyllabicity; the latter, in contrast, manages to fulfil not only FTBIN and DEP^h, but also WEAKC.

(4,22) Asymmetry between /CV^hC/ and /CV^h^uC/ stems at the stem level

input	output	DEP ^h	WEAKC	PARSE ²⁻³
/CV ^h C/ ⊗	[⌊ _o CV ^h C ^h ⌋]		*	
/CV ^h ^u C/ ⊗	[⌊ _o CV ^h ^u ⌋C]			*

I interpret this result as indicating that, in ME, two morae is the *default underlying specification* of vowel length in the environment /C_o__C/. Let me clarify the meaning of this assertion. The principle of Lexicon Optimization holds that the child acquires a lexicon by choosing those underlying representations which map onto the appropriate surface forms with the least serious violations of the constraint hierarchy. Typically, this is interpreted as implying that, in neutralization environments, the child will —whenever possible— project onto UR whatever predictable specification occurs on the surface. I am suggesting, however, that this interpretation is unduly narrow and should be extended. I assert that, for each environment where a high-ranking faithfulness constraint maintains contrast, Lexicon Optimization also defines a default specification: i.e. given two contrastive feature values in some environment, the one which causes the output of the stem level to incur the least serious violations of the stem-level hierarchy is identified by the child as the default.

This proposal recaptures one of the key insights of rule-based LPM: the phonology of the stem level defines context-specific default underlying specifications. Kiparsky (1995)

expresses this insight in terms of underspecification. In Radical Underspecification Theory, for each environment *E* where a feature [±F] is lexically contrastive, there is a rule *R* that specifies the 'default' value of [F] in *E*. As a result, default values may be left underlyingly unspecified; they will be filled in by *R* applying in structure-building mode. Underlying representations need only be specified for the 'marked' value of [F]; this prevents *R* from assigning the default value. In (3,23), for example, [+F] is the default value of [F] in the environment X__Y.

(4,23)

UR X[]Y X[-F]Y
R (structure-building) [] → [+F] / X __ Y X[+F]Y *blocked*

The advantage of a constraint-based approach is that it removes the inherent stipulativeness of rules. If a default rule does not do some work in structure-changing fashion, creating alternations, then there will be no evidence for it other than the very facts which identify a particular specification as default (e.g. lexical diffusion; see below). In an optimality-theoretic framework, in contrast, defaults *emerge* from the hierarchy.

Crucially, Kiparsky (1995) suggests that context-specific default rules determine the direction of lexical diffusion processes. He observes that the modification of underlying specifications during lexical diffusion is not random:

If we assume that lexical diffusion is nothing more than the substitution of one phoneme for another in the lexical representations of words, we have no explanation either for the direction of the change, nor for the envelope of phonological conditions that continues to control it.

Kiparsky (1995: 651)

He suggests, accordingly, that lexical diffusion consists in the removal of marked specifications from underlying representations: as the default rule ceases to be blocked, the default feature value propagates from item to item. In this sense, lexical diffusion simplifies the lexicon. I think that Kiparsky's proposal is correct in essence, but I will suggest that defaults are defined relative to the stem-level constraint hierarchy, as suggested above. In this sense, diffusion is Lexicon Optimization in output-changing, rather than output-preserving, fashion.

Goldsmith (1995) criticizes Kiparsky on the grounds that, if all defaults tend to propagate, his account of diffusion would appear to predict the levelling of robust contrasts, whilst it is typically marginal or unstable contrasts that suffer diffusion. I do not think that Goldsmith's observation raises a serious obstacle. The question to be addressed here is under what circumstances the child is forced to turn from the primary linguistic data to the grammatical default for the purposes of determining the content of underlying representations. In other words, what needs refining is not the structural characterization of the notion of default, but its practical application. Goldsmith himself provides a pointer when he notes that marginal contrasts (which will not be robustly expressed in the child's trigger experience) are the most likely targets.

I suggest that, in ME, the default pattern identified in tableau (4,22) was responsible for the irregular lengthening of *whale*-type words. Here are some more examples:²²

(4,24)	OE	PDE
	<i>blæd</i>	<i>blade</i>
	<i>cōl</i>	<i>coal</i>
	<i>crān</i>	<i>crane</i>
	<i>geōc</i>	<i>yoke</i>
	<i>hōp</i>	<i>hope</i>
	<i>mōt</i>	<i>mote</i>
	<i>slæd</i>	<i>slade</i>
	<i>þōl</i>	<i>thole</i>

Bermúdez-Otero (1998b, c) gives an account of *weapon*-type words following a similar line of argument; in essence, it is argued that constraints on foot structure cause a form of Trochaic Shortening in underlying representations. Interestingly, the default pattern for monosyllables is also at work in assigning length specifications to loans from Old French.

The assignment of length specifications to stressed vowels in Old French loans followed an extremely simple pattern, although the somewhat convoluted statements of the handbooks do not make it immediately apparent (cf. Minkova & Stockwell, 1996: 18). Exceptions exist, but they are not numerous:

²² Ritt (1997) gives a somewhat inflated estimation of the size of the *whale* class.

(i) Stressed vowels were specified as long in word-final syllables closed by a single consonant. This generalization applies both to monosyllables and to oxytonic polysyllables: e.g. *bas* > *base*, *bec* > *beak*, *fol* > *fool*, *pris* > *price*; *desfāt* > *defeat*, *despīt* > *despite*, *devōt* > *devout*.²³

(ii) Stressed vowels were generally specified as short in non-final syllables, including penultimate open syllables: e.g. *tréble* > *treble*, *sóuple* > *supple*; *jolif* > *jolly*, *paláis* > *palace*, *trelliz* > *trellis* (see Appendix B, §B.2 and §B.3); *qualité* > *quality*, *vanité* > *vanity*. However, long specifications were allowed in the environment of MEOSL. Thus, the stressed vowel is long in words subject to apocope of stem-final schwa (e.g. *dāme* > *dame*, *fāce* > *face*), although these items can also be handled by the generalization covering word-final syllables. More significantly, the length of the stressed vowel is unpredictable in disyllabic stems whose post-tonic rhyme contains a sonorant consonant: e.g. *tāble* > *table*, *póudre* > *powder* (cf. *treble* and *supple* above); *bacin* > *basin*, *odúr* > *odour*, vs *forāin* > *foreign*, *galón* > *gallon* (see Appendix B, §B.1).

These generalizations make it clear that the assignment of length specifications to stressed vowels in monosyllabic loans complied with the default pattern identified in (4,22). Additionally, there is evidence that vowel length values could not have been fully determined by the phonology of Old French. This confirms that the stem-level ranking WEAKC » PARSE⁹⁸ played a significant rôle in the adaptation of loans:

(i) Old French lacked phonemic vowel length distinctions (*pace* Bliss 1955; see Pope 1934: §1170).²⁴ Accordingly, vowel length specifications in ME loans do not reflect phonological contrasts.

(ii) On the basis of evidence from qualitative changes, Pope claims that vowel length in was determined by the following allophonic rule: vowels were allophonically long in tonic open syllables and in tonic word-final syllables closed by a single consonant; elsewhere, they were allophonically short (see Pope 1934: §574, §575, §580).²⁵ If Pope is right, it is plausible that this allophonic distribution may have influenced the assignment of vowel length specifications to ME loans. According to her rule, for example, stress-

²³ I use the acute accent to indicate the original position of the tonic syllable in Old French.

²⁴ In Old French there were two vowel phonemes in the vicinity of [e], one derived from Late Latin tonic unchecked /a/, the other from Late Latin checked /e/, but their failure to merge need not imply a length contrast (Bliss 1952/3: §3; Pope 1934: §231—§233).

²⁵ In essence, Pope postulates the continuity of a well-known early Gallo-Roman rule (Pope 1934: §198).

shifted forms became accented on an originally short vowel: e.g. *jō'tif* > ME *'jōlyf* 'jolly'. Nevertheless, the assumption that allophony determined ME length contrasts proves inadequate: ME loans display unpredictable variation both in originally tonic open syllables (e.g. *table*, *powder*, vs *treble*, *supple*) and in originally counter-tonic syllables (e.g. *basin*, *odour*, vs *foreign*, *gallon*).

(iii) Finally, the length of stressed vowels in ME loans cannot have been determined by the quality of the vowel, since identical vowels have both short and long reflexes: e.g. counter-tonic *a* turns up as long in *basin*, but as short in *gallon*.

4.5 SUMMARY

The traditional approach to IOE and ME quantity is essentially teleological: it is fuelled by Luick's vision of a relentless pursuit of the *Normalmaß*. However, this approach takes for granted a system of empirical generalizations which has proved wholly inadequate, as I have been able to show in section 4.2 (see the summary in §4.2.6). Attempts at neutralizing Minkova's reformulation of MEOSL have fudged the issue: Lass (1987, 1992) resorts to a selective use of the empirical evidence (see §4.3.1.2); Ritt (1994) couches his analysis in probabilistic terms too vague to be useful (Bermúdez-Otero & McCully 1997).

Within OT, in contrast, it is easy to incorporate the compensatory formulation of MEOSL into a broader account of IOE and ME quantity changes. Moreover, the pattern of grammatical connections which emerges from an optimality-theoretic analysis is non-teleological: globality arises from the parallel interaction of old and new rankings within a level, as suggested in §4.1.3.

First, SHOCC was a form of closed syllable shortening triggered by *[_σμμμ]. Monosyllables with VVC rhymes escaped the change through final consonant extrasyllabicity, enforced by the ranking WEAKC » PARSE^{SEG}. In -VC monosyllables, however, extrasyllabicity was cancelled by high-ranking DEP^μ, thereby preserving foot binarity.

Secondly, MEOSL was triggered by the presence of a deletable schwa in the post-tonic rhyme. It had nothing to do with a preference for heavy stressed syllables; the constraint PKPROM was inactive in ME, its effects neutralized by dominant DEP^μ. The 'analytical problem' of the compensatory formulation of MEOSL (Riad 1992: 335-6) is neatly solved by constraint rankings independently established in the analysis of SHOCC:

in apocopated disyllabic stems (e.g. /namə/ → [na:m]), the floating mora was forced to dock onto the root-vowel because WEAKC required the extrasyllabification of the final consonant; since no mora-insertion took place, DEP^μ authorized the change.

Finally, the stem-level ranking WEAKC » PARSE^{SEG} creates an important asymmetry between length specifications in the environment /C₀__C¹/: two morae is the default. This default pattern correctly predicts the assignment of vowel length to monosyllables and oxytones borrowed from Old French. Similarly, it accounts for the direction of lexical change in the *whale* class.

These results are relevant beyond the limits of English historical phonology. First, they undermine the widely accepted notion of a Pan-Germanic quantity shift driven by Prokosch's Law.²⁶ Secondly, and more importantly, I have proposed an optimality-theoretic method for determining default input specifications, and I have suggested ways to apply it to the study of historical lexical change, including lexical borrowing and lexical diffusion.

²⁶ See Árnason (1980: ch. 3), Lahiri (1995), Lass (1987: 159-162, 169; and elsewhere), Prokosch (1938: 140), Riad (1992: ch. 7; 1995), Vennemann (1988: 30).

Prospects

It is true that grammars are formed in the child in accordance with the prescriptions of the linguistic genotype; consequently, they are economical and beautiful in ways imposed by UG. However, they are also shaped by the demands of haphazard experience, and this makes them historically contingent. There is no contradiction here [...]. But there is a tension.

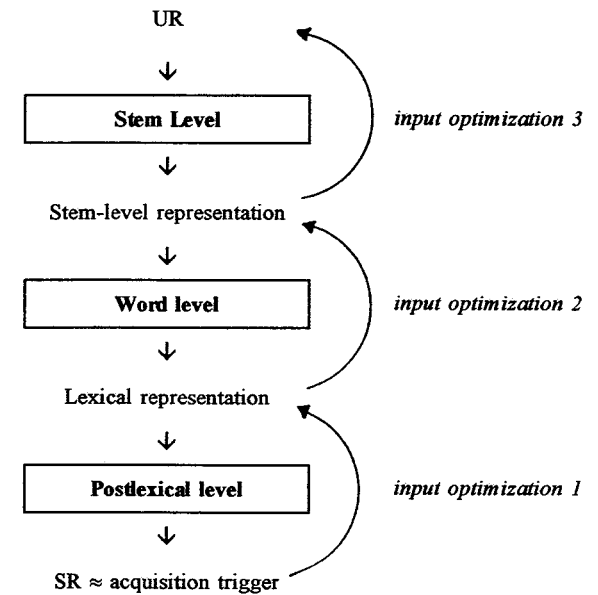
Lightfoot (1999: 111)

The preceding chapters have abundantly illustrated the fundamental tension at the heart of phonology between opacity and globality. In the recursive application of phonology over the hierarchical constituent structure of linguistic expressions there always lurks the danger of opacity. Phonological change causes this danger to become realized: as the child's grammar grows from the trigger experience, innovative constraint rankings introduced by hypocorrection at the postlexical level are free to disfigure the output of higher levels. No function is being served; this is just a Rube Goldberg feature of grammars. Yet within (non-cyclic) levels, at the same time, transparency reigns, and the overarching control of constraints on representations is felt in full force. I have suggested, following Kiparsky (1995: §2.1), that their power extends into the acquisition process: as constraint rankings are established in the nascent grammar, they exert a top-down selective pressure on the phonologization of automatic phonetic effects—hence diachronic globality, the intriguing obverse of opacity.

If a lesson emerges from the overall argument of this dissertation, it must be that grammatical theory and historical linguistics run astray when they run apart: without the theory of grammar, historical linguistics is dumb; without historical linguistics, grammar theory is blind. The late recrudescence of the abstractness debate has again intimated the fact. Phonologists are always tempted to impose draconian restrictions on the ability of grammars to encode opaque generalizations. The temptation should abate somewhat when one ponders how the upward percolation of constraint rankings in the grammar keeps the accretion of opaque interactions in check.

In this connection, the study of the life cycle of phonological processes holds exciting prospects. In §3.3.3.2, I advanced a new approach to this topic based on the idea that Lexicon Optimization applies recursively in the acquisition of stratified grammars:

(5,1)



When a phonological process develops sensitivity to some set of morphological structures, we know that automatic phonetic effects are not to blame. Some cue in the child's trigger experience must have disappeared or shifted; as a result, the child no longer has reason to posit at some level input representations which differ from the output, and Lexicon Optimization takes its course. Morphologization is, in this sense, of as much theoretical importance in phonology as obsolescence is in syntax (Lightfoot 1991); I hope to have shown that interleaved OT has equipped us with new and powerful theoretical tools with which to tackle this problem.

Appendix A

Morphological conditioning of SHOCC in the *Ormulum*

The data contained in this appendix were collected through a search of Holt's (1878) glossary. Words were selected according to two criteria: the stressed syllable was unequivocally closed, and the vowel was long in OE. Items such as <deofless> 'devil' pl./gen. (sg. <deofell> OE *dēofol*) were not included on the grounds that the medial cluster might be tautosyllabic (with resyllabification across the syncope site). OE forms are given in the right column for the purposes of comparison. Exceptions to the generalizations proposed are marked with an exclamation point. Only types are listed; the amount of tokens in the manuscript is not indicated, but is occasionally very large.

A.1 Underived stems

SHOCC applies regularly in underived stems:

<errnde> ¹	'errand, message'	<i>ǣrende</i>
<blosstme>	'blossom'	<i>blōstma</i>
<freollsenn>	'celebrate'	<i>frēols²</i>
<gittsunng>	'covetousness'	<i>gūsung³</i>
<lihht>	'light'	<i>lēohht</i>
<nohht>	'not'	<i>nōht</i>
<ohht>	'anything'	<i>ōht⁴</i>

Before *sC* clusters, shortening applies inconsistently: e.g.

<asskenn>	'ask'	<i>āscian</i>
<brest>	'breast'	<i>brēost</i>
<esstess>	'dainties'	<i>ēst</i>
<fosstrenn>	'foster' vb.	<i>fōstor</i>
<prest>	'priest'	<i>prēost</i>

¹ <errnde> occurs as nom. and acc. sg. There are no alternations to justify an unsyncopeated underlying representation /e:rændə/.

² *fr̥-hals > OE *frēols* 'free', whence *frēolsian* (Campbell, 1959: §238.2). Note that the SHOCC-triggering cluster occurs within the stem.

³ Related forms: *gūsian*, *gūserē*, *gūsiendnes*, etc. The -s- is to be analysed not as a derivational formative, but as part of the stem.

⁴ No longer synchronically related to <awihht> (OE *āwihht*) and <nawihht> (OE *nāwihht*).

A.2 Nominal inflection

Syncope attendant on the addition of nominal inflections fails to trigger SHOCC:

<dæcness>	'deacon' pl.	<i>dēacon</i>
<ehne> ~ <e3hne>	'eye' pl.	<i>ēage⁵</i>
<tacness>	'token, sign' pl.	<i>tācn</i>
! <hall3he> ⁶	'holy' infl.	<i>hālig</i>

A.3 Verbal inflection

SHOCC applies regularly to athematic past tense and past participle forms of weak verbs:

<brohhte>	'bring' pret. 3sg.	<i>bringan⁷</i>
<cwemnde>	'please' pret. 3sg.	<i>cwēman</i>
<demnde>	'deem' pret. 3sg.	<i>dēman</i>
<dredde>	'dread' pret. 3sg.	<i>adrēdan</i>
<fedde>	'feed' pret. 3sg.	<i>fēdan</i>
<flemnde>	'banish' pret. 3sg.	<i>flēmman</i>
<3emnde>	'heed' pret. 3sg.	<i>gēmman</i>
<hidd>	'hide' p. ptc.	<i>hȳdan</i>
<keppte>	'keep' pret. 3sg.	<i>cēpan</i>
<kidde>	'make known' pret. 3sg.	<i>cȳban</i>
<ledde>	'lead' pret. 3sg.	<i>lēdan</i>
<sleppte>	'sleep' pret. 3sg.	<i>slēpan</i>
<sohhte>	'seek' pret. 3sg.	<i>sēcan</i>
<spredd>	'spread' p. ptc.	<i>sprēdan</i>
<tahhte>	'teach' pret. 3sg.	<i>tāccan</i>
<pothhte>	'think' pret. 3sg.	<i>þencan</i>
<puhhte>	'appear' pret. 3sg.	<i>þyncan</i>
<wepptenn>	'weep' pret. 3pl.	<i>wēpan</i>

⁵ Orm's nom. sg. form is <e3he>; his plural form appears to involve double marking, with -e being added to the normal weak ending (cf. OE *ēagan*). The spelling <ehne>, suggesting a shortened root-vowel, appears just once in the manuscript, in contrast with well over a dozen occurrences of <ehne> ~ <e3hne>.

⁶ *Hali3* is exceptional in the *Ormulum* in that it is the only word where the suffix -i3 undergoes syncope under inflection (see Burchfield, 1956: 77, fn. 1, for a list of -i3 derivatives). Note also the OE weak noun *hālgā* 'saint'. Stem suppletion is therefore likely.

⁷ OE pret. 3sg. *brōhte*, from an athematic Germanic form *brapx-t-; cf. also OE *pōhte* (*þencan*) and *þūhte* (*þyncan*) (Campbell, 1959: §753.9.b.5).

A.4 Derivation

A.4.1 Zero-derived weak verbs

Syncope attendant on the addition of verbal inflections to zero-derived weak verbs fails to trigger SHOCC:

<ahnenn> ⁸	'obtain'	<i>āgen</i> (whence <i>āgnian</i>)
<becnenn>	'beckon'	<i>bēacn</i>
<huslenn>	'to admit to Communion'	<i>hūsl</i>
<tacnenn>	'betoken'	<i>tācn</i>
<wæpnedd>	'arm' p. ptc.	<i>wāpn</i>
! <crisstenn>	'christen'	<i>Crīsten</i> ⁹
! <hallghenn>	'consecrate'	<i>hālig</i> (see note 6)

A.4.2 -p/-t

The suffix *-p/-t* triggers SHOCC consistently:

<ahhte>	'property'	<i>āgan</i> 'own, possess'
<ma33be>	'kin'	<i>māg</i>
<lappe>	'hostility'	<i>lāp</i>
<seollpe>	'happiness'	<i>sāel</i>
<wrappe>	'wrath'	<i>wrāp</i>

A.4.3 Other suffixes

All derivational suffixes other than *-p/-t* fail to trigger SHOCC:

<æddmodle33c>	'humility'	<i>mōd</i>
<ædmodli3>	'humbly'	<i>mōd</i>
<ædmodnesse>	'humility'	<i>mōd</i> ¹⁰
<anfald>	'onefold'	<i>ān</i>
<buhsumm>	'obedient, pliable'	<i>būgan</i> 'bend'
<clænle33c>	'cleanliness'	<i>clāne</i>

⁸ Note that in Orm's dialect *hn-* no longer is a permissible onset cluster: e.g. <nesshe> 'tender' OE *hnesce*.

⁹ 'Christian', but Orm has <Cristene> ~ <Cristene>, whence also <Cristennom>.

¹⁰ Here the relevant feature of the forms <æddmodle33c>, <ædmodli3> (also <æddmodli3>) and <ædmodnesse> (also <æddmodnesse>) is the absence of vowel shortening in the second root *mōd* after the addition of a consonant-initial derivational suffix. For the behaviour of the first root, see A.5 below.

<clænnesse> ¹¹	'cleanness'	<i>clāne</i>
<cupli3>	'familiarly'	<i>cūp</i>
<deoplikerr>	'deeply' comp.	<i>dēop</i>
<fiffald>	'fivefold'	<i>fīf</i>
<flæshlic>	'carnal'	<i>flāesc</i>
<3æple33c>	'caution'	<i>gēap</i>
<galnesse>	'wantonness'	<i>gāl</i>
<gastlic>	'spiritual'	<i>gāst</i>
<godle33c>	'goodness'	<i>gōd</i>
<godnesse>	'goodness'	<i>gōd</i>
<halsumm>	'wholesome'	<i>hāl</i>
<hehlike>	'high'	<i>hēah</i>
<leflī3>	'affectionately'	<i>lēof</i>
<onnlicnesse>	'likeness'	<i>gelīc</i>
<meocle33c>	'meekness'	Orm <meoc>
<meocnesse>	'meekness'	Orm <meoc>
<sellcuplike>	'unusually'	<i>cūp</i>
<secnedd> ¹²	'sicken' p. ptc.	<i>sēoc</i>
<soþlike>	'truly'	<i>sōþ</i>
<swētlike>	'sweetly'	<i>swēte</i>
<wislike>	'wisely'	<i>wīs</i>
! <clennlike>	'cleanly'	<i>clāne</i>
! <clennsenn>	'cleanse'	<i>clāne</i>
! <dunnwarrd>	'downward'	<i>adūne</i>
! <herrsumm>	'obedient'	<i>hīeran</i>
! <liccness> ¹³	'likeness'	<i>gelīc</i>
! <mannsenn>	'excommunicate' vb.	<i>mān</i>
! <wissenn> ¹⁴	'instruct'	<i>wīs</i>

¹¹ Not *<clennesse>. Orm's orthography occasionally allows triple consonant graphs: e.g. <drunkennesse> 'drunkenness' ll. 14741, 15377, 15389 (Holm, 1922: 94); <skillles> 'ignorant' l. 3715. Moreover, the length of the root-vowel is redundantly specified by the graph <æ>, which uniquely designates the long vowel /ɛ:/ (OE *ā, ēa*), its short counterpart being <e>: cf. <clennlike> below (see Anderson & Britton, 1997: §49-51). Note also the antonym <unncleannesse>.

¹² Derived by means of an *-n-* weak verb formative (OE *-n-i-an*; see Lass, 1994: 203).

¹³ But cf. <onlicnesse> above.

¹⁴ Whence <wissinn> 'instruction'. In OE, *wīsan* alternated with *wissian*, which presumably already had a short vowel (Campbell, 1959: §287). Forms such as <clennsenn>, <mannsenn> and <wissenn> may suggest that in Orm's dialect the presence of the *-s-* formative in thematic weak verbs (OE *-s-i-an*; see Lass, 1994: 203) triggers SHOCC consistently. If so, the derivational verbal suffix *-senn* would pattern with the derivational nominal suffix *-p/-t*. The evidence is, unfortunately, insufficient: note, in particular, that it is doubtful whether Orm perceived <wissenn> as containing the *-s-* formative, given that he does not write <wissenn>.

A.4.4 Suffixoids¹⁵

Suffixoids fail to trigger SHOCC:

<læffull> ¹⁶	'believing'	<i>gelēafa</i>
<nīpfull>	'envious'	<i>nīð</i>
<soþfasst>	'true, faithful'	<i>sōþ</i>
<unnþæwfull>	'immoral'	<i>þēaw</i>
<whilwendlic>	'temporary'	<i>hwīl</i>
! <wissdom> ¹⁷	'wisdom'	<i>wīs</i>

A.5 Compounding

Transparent compounding fails to trigger SHOCC:

<æstdale>	'eastward'	<i>ēast</i>
<bocstaff>	'letter, character'	<i>bōc</i>
<breostlin>	'breast-plate of linen'	<i>brēost</i>
<bridgume>	'bridegroom'	<i>brȳd</i>
<dædbote>	'penitence'	<i>dǣd</i>
<dæþshildiȝ>	'guilty, condemned'	<i>dēaþ</i>
<driȝcrafftess>	'magical arts'	<i>drȳ</i>
<driȝmenn>	'magicians'	<i>drȳ</i>
<dunstiȝhinng>	'descent'	<i>adūne</i>
<fictre>	'fig-tree'	<i>fīc</i>
<hehfaderr>	'God the Father'	<i>hēah</i>
<larfaderr>	'teacher'	<i>lār</i>
<larspell>	'doctrine'	<i>lār</i>
<licwurrþiȝ>	'agreeable'	<i>gelīc</i>
<shephirde>	'shepherd'	<i>scēap</i>
<supdale>	'southward'	<i>sūþ</i>
<ūtþresstenn>	'burst out'	<i>ūt</i>
<ūtledenn>	'lead out'	<i>ūt</i>
<ūtnummenn>	'outstanding'	<i>ūt</i>
<ūtwiþþ>	'beyond, out of'	<i>ūt</i>
<wifmann> ¹⁸	'woman'	<i>wīf</i>

¹⁵ In the classification of morphemes as suffixes, suffixoids, or compound formatives, I follow Sauer (1992).

¹⁶ Not *<leffull>; see note 11.

¹⁷ It is significant that a good many morphologically unexpected applications of SHOCC in the *Ormmulum* are concentrated in a handful of roots: <clennlike>, <clennsenn>, <unncleensedd>, <Crisstene>, <crisstnenn>; <hallȝe>, <hallȝhenn> (plural noun), <hallȝhenn> (vb.); <wissenn>, <wissinnȝ>, <wissdom>. These appear to be frozen relics; their synchronic status is probably suppletive.

¹⁸ Alongside the grammaticalized form <wimmann>.

! <ædmod> ~ <æddmod> ¹⁹	'humble'	<i>ēade</i>
! <chappmenn>	'chapmen, merchants'	<i>cēap</i>
! <goddsPELL>	'Gospel'	<i>gōd</i> ²⁰
! <laffdiȝ>	'lady'	<i>hlāf</i>
! <weppmann>	'man'	<i>wāpn</i> ²¹

¹⁹ Both Fulk (1996: 496-7) and Anderson & Britton (1997: 49-50) observe that the form <æddmod>, as well as its derivatives <æddmodleȝȝc>, <æddmodliȝ> and <æddmodnesse>, displays conflicting orthographic clues: on the one hand, the graph <æ> implies a long vowel (see footnote 11 above); on the other, the double consonant implies shortness. The change -ðm- > -dm- (Campbell, 1959: §424) may be regarded as obscuring the relationship of the compound with its base: OE *ēade* 'easy, friendly', Orm <æþ> (cf. Fulk, 1996: 504-5).

²⁰ In this form shortening may already have occurred in OE (Luick, 1964: §204.1; Campbell, 1959: §285), although this is hard to ascertain. Orm is aware of the correct etymology of the word (Dedication, ll. 157ff.).

²¹ The truncation of the first element, already encountered in OE (Campbell, 1959: §477.5), may be a sign of grammaticalization.

Appendix B

Unapocopated disyllables and MEOSL

This appendix lists those MEOSL-candidates included in Minkova's (1982) corpus which remain disyllabic in PDE, with the sole exclusion of *process* (see footnote 14 in chapter 4). Only PDE reflexes are given: for the OE and ME forms of native items, see Minkova (1982: 33-40); for Anglo-Norman items, see Bliss (1952/3: §56). The list is broken down into three groups according to the structure of the post-tonic rhyme; within each group, reflexes are classified as lengthened or unlengthened. The double bar || separates reflexes of items attested in OE from OF loans. The total number of entries (with etymological doublets receiving two entries) is 231.

B.1 The post-tonic rhyme contains a sonorant consonant: 166 items

B.1.1 Lengthened: 39 items

acorn, acre, beaver, besom (also short), *chafer* (cf. *chaffer*), *cradle* (cf. *credde*), *even, gable, haven, hazel, ladle, maple, navel, open, over* (cf. *uvver*), *raven, staple, taper, treadle, weasel* || *bacon, basin, blazon, capon, favour, flavour, label, labour, mason, paper, patient, savour, razor, tabor, vacant, vapour, azure* (also short), *moment, odour*.

B.1.2 Unlengthened: 127 items

addle, aspen, bastard, batten, besom (also long), *better, blather, bottom, bracken, brothel, cackle, camel, canon, castle, chaffer* (cf. *chafer*), *clatter, cocker, cackle, copper, credde* (cf. *cradle*), *edder, father, feather, fennel, fester, fetter, fettle, gammon, gather, gavel, gravel, hammer, hatchel, heaven, hovel, hover, kettle, Latin, latter, leather, maslin* (obs.), *nether, nettle, otter, oven, uvver* (dial.; cf. *over*), *pebble, pepper, pottle, rather, reckon, repple* (obs.), *saddle, seven, shackle, shovel, smother, sollar, swaddle, talent, tetter, throttle, throstle, water, wattle, weather, weither, whether* || *alum, azure* (also long), *baron, barren, barrel, cattle, channel, chattel, clamour, dragon, flatter, gallon, hazard, latten, mallard, manor, panel, sachel, satin, tabard, talon, tassel, travel, valour, vassal, warrant, beryl, bezant,*

cellar, deavour, desert, felon, herald, kennel, lecher, lesson, metal, pennon, peril, present, record, revel, second, tenant, tenor, treasure, tremor, venom, coffin, collar, common, coral, florin, foreign, honour, moral, volume; colour, covin.

B.2 The post-tonic rhyme contains an unchecked vowel: 24 items

B.2.1 Lengthened: 0 items

B.2.2 Unlengthened: 24 items

barrow, belly, berry, body, callow, fallow, ferry, harry, harrow, heavy, holly, many, mellow, narrow, nephew, penny, poppy, ready, sallow, shadow, steady, tarry, yellow || *jolly*.

B.3 The post-tonic rhyme consists of vowel plus obstruent: 41 items

B.3.1 Lengthened: 1 item

naked.

B.3.2 Unlengthened: 40 items

basket, chalice, collop, eddish, gannet, haddock, jaspis, planet, provost, radish, relic, trivet || *anet, anise, barrat, claret, damask, faggot, habit, latchet, marish, palace, palate, statute, brevet, jealous, legate, prelate, senate, trellice, bonnet, closet, crotchets, forest, profit, rocket, rochet, socket, solace; cherish*.

References

- Alderete, J. (1995). Faithfulness to prosodic heads. Ms., University of Massachusetts, Amherst. (ROA-94, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.) Forthcoming in Hermans, B. & M. van Oostendorp (eds). *The derivational residue in phonology*. Amsterdam: John Benjamins.
- Allen, W. S. (1973). *Accent and rhythm. Prosodic features of Latin and Greek: a study in theory and reconstruction*. Cambridge: Cambridge University Press.
- Al-Mohanna, F. (1997). The superheavy syllable in Arabic: an OT account. Paper presented at the Autumn Meeting of the Linguistic Association of Great Britain, University of Hertfordshire, 6 September 1997.
- Anderson, J. & D. Britton (1997). Double trouble: geminate versus simplex graphs in the *Ormulum*. In Fisiak, J. (ed.) *Studies in Middle English linguistics*. Berlin: Mouton de Gruyter. 23-58.
- Anderson, S. R. (1981). Why phonology isn't natural. *Linguistic Inquiry* 12: 493-539.
- Anderson, S. R. (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.
- Anttila, A. & Y.-m. Y. Cho (1997). Variation and change in Optimality Theory. Ms., Stanford University, <http://www-linguistics.stanford.edu/Linguistics/papers>. (To appear in *Lingua*).
- Archangeli, D. & D. Pulleyblank (1994). *Grounded phonology*. Cambridge, MA: The MIT Press.
- Árason, K. (1980). *Quantity in historical phonology: Icelandic and related cases*. Cambridge: Cambridge University Press.
- Aronoff, M. (1994). *Morphology by itself*. Cambridge, MA: MIT Press.
- Bagemihl, B. (1991). Syllable structure in Bella Coola. *Linguistic Inquiry* 22: 589-646.
- Baković, E. J. (1995). Strong onsets and Spanish fortition. Ms., Rutgers University. (ROA, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Beckman, J. N. (1998). *Positional faithfulness*. PhD dissertation, University of Massachusetts, Amherst. (ROA-234, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Benua, L. (1995). Identity effects in morphological truncation. In J. Beckman, S. Urbanczyk & L. Walsh (eds). *Papers in Optimality Theory*. (University of Massachusetts Occasional Papers in Linguistics 18). Amherst, MA: Graduate Linguistic Student Association.
- Benua, L. (1997). *Transderivational identity: phonological relations between words*. PhD dissertation, University of Massachusetts, Amherst.
- Bermúdez-Otero, R. (1998a). Prosodic optimization: the Middle English length adjustment. *English Language and Linguistics* 2: 169-197.
- Bermúdez-Otero, R. (1998b). Harmonic evaluation of lexical items: evidence from Middle English. Paper presented at the 6th Manchester Phonology Meeting, Manchester, 23 May 1998.
- Bermúdez-Otero, R. (1998c). Lexicon Optimization: irregular vowel length changes in Middle English. Paper presented at the X International Conference on English Historical Linguistics, Manchester, 22 August 1998.
- Bermúdez-Otero, R. & C. B. McCully (1997). Review of Ritt (1994). *Journal of Linguistics* 33: 620-5.
- Berwick, R. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.
- Bickmore, L. S. (1995). Accounting for compensatory lengthening in the CV and moraic frameworks. In Durand, J. & F. Katamba (eds). *Frontiers of phonology: atoms, structures, derivations*. London: Longman. 119-148.
- Black, H. A. (1994). *Constraint-ranked derivation: a serial approach to optimization*. PhD dissertation, University of California, Santa Cruz.
- Bliss, A. J. (1952/3). Vowel quantity in Middle English borrowings from Anglo-Norman. *Archivum Linguisticum* 4: 121-147, 5: 22-47. Reprinted in Lass, R. (ed.) (1969). *Approaches to English historical linguistics*. New York: Holt, Rinehart and Winston. 164-207.
- Bliss, A. J. (1955). Quantity in Old French and Middle English. *Archivum Linguisticum* 7: 71-86.
- Boberg, V. (1896). Undersøgelser om de danske vokalers kvantitet. *Arkiv för nordisk filologi* 12: 315-366.
- Boersma, P. (1997). Sound change in Functional Phonology. Ms., University of Amsterdam. (ROA-237-1297, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.)
- Booij, G. (1996). Lexical phonology and the derivational residue. In J. Durand & B. Laks (eds). *Current trends in phonology: models and methods*. Salford: European Studies Research Institute. Vol. 1: 69-96.

- Booij, G. (1997). Non-derivational phonology meets Lexical Phonology. In Roca, I. (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press. 261-288.
- Booij, G. & J. Rubach (1987). Postcyclic versus postlexical rules in Lexical Phonology. *Linguistic Inquiry* 18: 1-44.
- Borowsky, T. (1993). On the word level. In Hargus, S. & E. M. Kaisse (eds). *Studies in Lexical Phonology*. (Phonetics and Phonology 4) San Diego: Academic Press. 199-234.
- Borowsky, T., J. Itô & R.-A. Mester (1984). The formal representation of ambisyllabicity: Evidence from Danish. *Papers from the Annual Meeting of the North East Linguistic Society* 14: 34-48.
- Bräuer, H. (1961). *Slavische Sprachwissenschaft I: Einleitung, Lautlehre*. Sammlung Götschen Band 1191/1191a. Berlin: de Gruyter.
- Braune, W. (1880). *Gotische Grammatik mit Lesestücken und Wörterverzeichnis*. 19th edn (1981) by E. A. Ebbinghaus. Tübingen: Max Niemeyer Verlag.
- Braune, W. (1886). *Althochdeutsche Grammatik*. 13th edn (1975) by H. Eggers. Tübingen: Max Niemeyer Verlag.
- Broselow, E. (1995). Skeletal positions and moras. In J. A. Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell. 175-205.
- Broselow, E., S.-I. Chen & M. Huffman (1997). Syllable weight: convergence of phonology and phonetics. *Phonology* 14: 47-82.
- Browman, C. P. & L. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* 6: 201-251.
- Browman, C. P. & L. Goldstein (1990). Tiers in Articulatory Phonology, with some implications for casual speech. In Kingston, J. & M. E. Beckman (eds). *Papers in Laboratory Phonology I. Between the grammar and physics of speech*. Cambridge: Cambridge University Press. 341-397.
- Brunner, K. (1948). *Abriss der mitttelenglischen Grammatik*. (2nd edn.) Halle: Max Niemeyer Verlag. Translated from the 5th edn by Johnston, G. (1970). *An outline of Middle English grammar*. Oxford: Basil Blackwell.
- Burchfield, R. W. (1956). The language and orthography of the Ormulum MS. *Transactions of the Philological Society* 56-87.
- Calabrese, A. (1994). Sievers' Law in Gothic: a synchronic analysis with some notes on its diachronic development. *The Linguistic Review* 11. 149-194.

- Calabrese, A. & C. Romani (1991). Syllable structure in aphasia: a case study. In Bertinetto, P. M. & M. Loporcaro (eds). *Certamen phonologicum II*. Torino: Rosenberg and Sellier.
- Campbell, A. (1959). *Old English Grammar*. Oxford: Clarendon Press.
- Carlton, T. R. (1990). *Introduction to the phonological history of the Slavic languages*. Columbus: Slavica.
- Chen, M. (1972). The time dimension: contribution toward a theory of sound change. *Foundations of Language* 8: 457-498. Reprinted in Wang, W. S.-Y. (ed.) (1977). *The lexicon in phonological change*. Mouton: The Hague. 197-265.
- Chomsky, N. & M. Halle (1968). *The sound pattern of English*. New York: Harper and Row.
- Clayton, M. L. (1976). The redundancy of underlying morpheme-structure conditions. *Language* 52: 295-313.
- Clements, G. N. (1985a). The geometry of phonological features. *Phonology Yearbook* 2: 225-252.
- Clements, G. N. (1985b). Compensatory lengthening and consonant gemination in LuGanda. In L. Wetzels & E. Sezer (eds). *Studies in compensatory lengthening*. Dordrecht: Foris. 37-77.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (eds). *Papers in laboratory phonology, I: Between the grammar and physics of speech*. Cambridge: Cambridge University Press. 283-333.
- Clements, G. N. (1992). The sonority cycle and syllable organization. In Dressler *et al.* (eds). 63-76.
- Clements, G. N. & E. Hume (1995). The internal organization of speech sounds. In J. A. Goldsmith (ed.) *The handbook of phonological theory*. Oxford: Blackwell. 245-306.
- Cole, J. (1995). The cycle in phonology. In J. A. Goldsmith (ed.) *The handbook of phonological theory*. Oxford: Blackwell. 70-113.
- Collinge, N. E. (1985). *The laws of Indo-European*. Philadelphia: John Benjamins.
- Croft, W. (1990). *Typology and universals*. Cambridge: Cambridge University Press.
- Crowhurst, M. & M. Hewitt (1997). Boolean operations and constraint interactions in Optimality Theory. Ms., University of North Carolina at Chapel Hill and Brandeis University. (ROA-229-1197, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).

- Davis, S. (1994). Geminate consonants in moraic phonology. *Proceedings of the West Coast Conference on Formal Linguistics* 13: 32-45.
- de Chene, E. B. & S. R. Anderson (1979). Compensatory lengthening. *Language* 55: 505-535.
- de Lacy, P. (1998). Sympathetic stress. Ms., University of Massachusetts, Amherst. (ROA-294-0199, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Dinnsen, D. (1980). Phonological rules and phonetic explanation. *Journal of Linguistics* 16: 171-191.
- Dobson, E. J. (1962). Middle English lengthening in open syllables. *Transactions of the Philological Society* 1962: 124-149.
- Dresher, B. E. (1983). Postlexical phonology in Tiberian Hebrew. *Proceedings of the West Coast Conference on Formal Linguistics* 2: 67-78.
- Dresher, B. E. (1992). A learning model for a parametric theory in phonology. In Levine, R. (ed.) *Formal grammar: theory and implementation*. New York: Oxford University Press. 290-317.
- Dresher, B. E. (1993). Cues and parameters in phonology. *Papers from the Annual Regional Meeting, Chicago Linguistic Society* 27. Part 1: 119-140.
- Dresher, B. E. (1994). Acquiring stress systems. In Ristad, E. S. (ed.) *Language computations. DIMACS Workshop on Human Language, March 20-22, 1992*. (DIMACS Series in Discrete Mathematics and Theoretical Computer Science 17). Providence, RI: American Mathematical Society. 71-92.
- Dresher, B. E. (1996). Learnability and phonological theory. In J. Durand & B. Laks (eds). *Current trends in phonology: models and methods*. Salford: European Studies Research Institute. Vol. 1: 245-266.
- Dresher, B. E. (1998). Paradigmatic levelling as a response to phonological incoherence. Plenary paper at the X International Conference on English Historical Linguistics, Manchester, 21-26 August 1998.
- Dresher, B. E. & J. D. Kaye (1990). A computational learning model for metrical phonology. *Cognition* 34. 137-195.
- Dresher, B. E. & A. Lahiri (1991). The Germanic foot: metrical coherence in Old English. *Linguistic Inquiry* 22: 251-286.
- Elugbe, B. O. (1982). The vowels of Proto-Edoid. *Journal of the Linguistic Association of Nigeria* 1: 107-115.

- Elugbe, B. O. (1989). Edoid. In Bendor-Samuel, J. (ed.) *The Niger-Congo languages: a classification and description of Africa's largest language family*. Lanham: University Press of America. 291-304.
- Everett, D. L. & I. Berent (1997). The comparative optimality of Hebrew roots: an experimental approach to violable identity constraints. Ms., ROA-235-1297, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.
- Fabb, N. (1988). English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory* 6: 527-539.
- Féry, C. (1997). Uni und Studis: die besten Wörter des Deutschen. *Linguistische Berichte* 172: 461-490.
- Féry, C. (1998). On the best optimality-theoretic account of German Final Devoicing. Ms., University of Tübingen. (ROA-274, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Féry, C. (1999). Truncation. Paper presented at the Workshop on Optimality Theory: Insights on Prosody and Morphology, Spring Meeting of the Linguistics Association of Great Britain, Manchester, 8 April 1999.
- Fidelholtz, J. (1967). English vowel reduction. Ms., MIT, Cambridge, MA.
- Fisiak, J. (1968). *A short grammar of Middle English*. Warszawa: Wydawnictwo Naukowe PWN.
- Fukuzawa, H. (1998). Multiple input-output faithfulness relations in Japanese. Ms., University of Maryland at College Park. (ROA-260-0598, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Fukuzawa, H., M. Kitahara & M. Ota (1998). Lexical stratification and ranking invariance in constraint-based grammars. *Chicago Linguistics Society* 34. (ROA, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Fulk, R. D. (1992). *A history of Old English meter*. Philadelphia: University of Pennsylvania Press.
- Fulk, R. D. (1996). Consonant doubling and open syllable lengthening in the *Ormulum*. *Anglia* 114: 481-513.
- García de Diego, V. (1959). *Manual de dialectología española*. Madrid: Ediciones Cultura Hispánica.
- Givón, T. (1971). Historical syntax and synchronic morphology: an archaeologist's field trip. *Chicago Linguistics Society* 7: 394-415.

- Gnanadesikan, A. (1995). Markedness and faithfulness constraints in child phonology. Ms., Rutgers University. (ROA, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Gnanadesikan, A. (1997). *Phonology with ternary scales*. PhD dissertation, University of Massachusetts, Amherst. (ROA, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Goldsmith, J. A. (1990). *Autosegmental and metrical phonology*. Oxford: Basil Blackwell.
- Goldsmith, J. A. (1993). Harmonic Phonology. In Goldsmith, J. A. (ed.). *The last phonological rule: Reflections on constraints and derivations*. Chicago: University of Chicago Press. 21-60.
- Goldsmith, J. A. (1995). Phonological theory. In J. A. Goldsmith (ed.) *The handbook of phonological theory*. Oxford: Blackwell. 1-23.
- Goldsmith, J. A. (1996). Tone in Mituku: how a floating tone nailed down an intermediate level. In J. Durand & B. Laks (eds). *Current trends in phonology: models and methods*. Salford: European Studies Research Institute. Vol. 1: 267-280.
- Gould, S. J. & R. C. Lewontin (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London, B*, 205: 581-598.
- Green, A. D. (1997). *The prosodic structure of Irish, Scots Gaelic, and Manx*. PhD dissertation, Cornell University.
- Greenberg, J. H. (1978). Diachrony, synchrony, and language universals. In Greenberg, J. H. (ed.) *Universals of human language*. (4 vols.) Stanford: Stanford University Press. Vol 1: 61-91.
- Greenberg, J. H. (1995). The diachronic typological approach to language. In Shibatani, M. & T. Bynon (eds.) *Approaches to language typology*. Oxford: Clarendon Press. 145-166.
- Gupta, A. (1987). Hindi word stress and the obligatory branching parameter. *Chicago Linguistics Society* 23: 134-148.
- Hale, M. & C. Reiss (1997a). Grammar Optimization: the simultaneous acquisition of constraint ranking and a lexicon. Ms., Concordia University. (ROA-231-1197, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Hale, M. & C. Reiss (1997b). Formal and empirical arguments concerning phonological acquisition. Ms., Concordia University. (ROA-233-1197, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>). To appear in *Linguistic Inquiry*.

- Halle, M. (1962). Phonology in generative grammar. *Word* 18: 54-72.
- Halle, M. & M. Kenstowicz (1991). The Free Element Condition and cyclic versus noncyclic stress. *Linguistic Inquiry* 22: 457-501.
- Halle, M. & K. P. Mohanan (1985). Segmental phonology of Modern English. *Linguistic Inquiry* 16: 57-116.
- Halle, M. & J.-R. Vergnaud (1987a). Stress and the cycle. *Linguistic Inquiry* 18: 45-84.
- Halle, M. & J.-R. Vergnaud (1987b). *An essay on stress*. (Current Studies in Linguistics 15). Cambridge MA: MIT Press. Re-edited 1990, with substantial typographical amendments.
- Hammond, M. (1997). Vowel quantity and syllabification in English. *Language* 73: 1-17.
- Hansen, A. (1962). *Den lydlige udvikling i dansk, fra ca. 1300 til nutiden. 1, vokalismen*. Copenhagen: CEC Gads forlag.
- Harms, R. (1968). *Introduction to phonological theory*. Englewood Cliffs, N.J.: Prentice-Hall.
- Harris, J. (1989). Towards a lexical analysis of sound change in progress. *Journal of Linguistics* 25: 35-56.
- Harris, J. (1992). Licensing Inheritance. *UCL Working Papers in Linguistics* 4. 359-406.
- Harris, J. (1994). *English sound structure*. Oxford: Blackwell.
- Harris, J. W. (1983). *Syllable structure and stress in Spanish: a nonlinear analysis*. Cambridge, MA: MIT Press.
- Harris, J. W. (1984). Autosegmental phonology, lexical phonology, and Spanish nasals. In Aronoff & Oehrlé (eds). *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students*. Cambridge, Mass.: MIT Press. 67-82.
- Harris, J. W. (1991a). The exponence of gender in Spanish. *Linguistic Inquiry* 22: 27-62.
- Harris, J. W. (1991b). The form classes of Spanish substantives. *Yearbook of Morphology* 2: 65-88.
- Harris, J. W. (1991c). With respect to accentual constituents in Spanish. In Campos, H. & F. Martínez-Gil (eds). *Current studies in Spanish linguistics*. Washington D.C.: Georgetown University Press. 447-473.
- Harris, J. W. (1993). Integrity of prosodic constituents and the domain of syllabification rules in Spanish and Catalan. In K. Hale & S. J. Keyser (eds). *The view from Building 20: essays in honor of Sylvain Bromberger*. Cambridge, MA: The MIT Press. 177-193.

- Harris, J. W. (1995). Projection and edge marking in the computation of stress in Spanish. In J. A. Goldsmith (ed.) *The handbook of phonological theory*. Oxford: Blackwell. 70-113.
- Hayes, B. (1982). Extrametricality and English stress. *Linguistic Inquiry* 13: 227-276.
- Hayes, B. (1986). Inalterability in CV phonology. *Language* 62: 321-351.
- Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20: 253-306.
- Hayes, B. (1995). *Metrical stress theory: principles and case studies*. Chicago/London: The University of Chicago Press.
- Hayes, B. (1996). Phonetically-driven phonology: the role of Optimality Theory and Inductive Grounding. Ms., UCLA. (ROA-158-1196, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Hechtenberg Collitz, K. (1906). Syllabication in Gothic. *Journal of English and Germanic Philology* 6: 72-91.
- Hermann, E. (1923). *Silbenbildung im Griechischen und in den anderen indogermanischen Sprachen*. Göttingen: Vandenhoeck & Ruprecht. 2nd edn 1978.
- Hoberman, R. D. (1995). Current issues in Semitic phonology. In Goldsmith, J. A. (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell Publishers. 839-847.
- Hock, H. H. (1986). Compensatory lengthening: in defense of the concept 'mora'. *Folia Linguistica Historica* 20: 431-460.
- Hogg, R. M. (1992a). *A grammar of Old English*. Vol. 1: *Phonology*. Oxford: Basil Blackwell.
- Hogg, R. M. (1992b). Phonology and morphology. In Hogg, R. (ed.) *The Cambridge history of the English language*. Vol. I: *The beginnings to 1066*. Cambridge: Cambridge University Press. 67-167.
- Holm, S. (1922). *Corrections and additions in the Ormulum manuscript*. Uppsala: Amqvist & Wiksell.
- Holt, R. (1878). *The Ormulum, with the notes and glossary of Dr. R. M. White*. 2 vols. Oxford: Clarendon Press.
- Hooper, J. B. (1976). *An introduction to Natural Generative Phonology*. New York: Academic Press.
- Hotzenköcherle, R. (1986). Aspekte und Probleme der Vokalquantität in Schweizerdeutschen. In Hotzenköcherle, R. *Dialektstrukturen im Wandel:*

- Gesammelte Aufsätze zur Dialektologie der deutschen Schweiz und der Walsertalgebiete Oberitaliens*. (ed. by R. Schläpfer & R. Trüb). Aarau: Verlag Sauerländer. 319-333.
- Hubbard, K. (1995a). 'Prenasalised consonants' and syllable timing: evidence from Runyambo and Luganda. *Phonology* 12: 235-256.
- Hubbard, K. (1995b). Morification and syllabification in Bantu languages. *Journal of African Languages and Linguistics* 16: 137-155.
- Hulst, H. G. van der (1994). On the relation between primary and secondary accent: Ten years later. Paper presented at the Phonology Workshop, Durham, September 16-17, 1994.
- Hume, E., J. Muller & A. van Engelenhoven (1997). Non-moraic geminates in Leti. *Phonology* 14: 371-402.
- Hutton, J. (1996). The Old English accent, *giet ongean*. Ms., University of Manchester.
- Hyman, L. M. (1985). *A theory of phonological weight*. Dordrecht: Foris.
- Hyman, L. M. (1994). Cyclic phonology and morphology in Cibemba. In J. Cole & C. Kisseberth (eds). *Perspectives in phonology*. Stanford: CSLI Publications. 81-112.
- Idsardi, W. J. (1997). Phonological derivations and historical changes in Hebrew spirantization. In Roca, I. (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press. 261-28
- Inkelas, S. (1993). Deriving cyclicity. In Hargus, S. & E. M. Kaisse (eds). *Studies in Lexical Phonology*. (Phonetics and Phonology 4) San Diego: Academic Press. 75-110.
- Inkelas, S. (1994). The consequences of optimization for underspecification. Ms., University of California, Berkeley. (ROA-40-1294, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Inkelas, S. & C. O. Orgun (1995). Level ordering and economy in the lexical phonology of Turkish. *Language* 71: 763-793.
- Itô, J. (1986). *Syllable theory in prosodic phonology*. PhD dissertation, University of Massachusetts, Amherst. Published (1988) New York: Garland Press.
- Itô, J. & R. A. Mester (1992). Weak Layering and Word Binariness [Preliminary version]. LRC-92-09, Linguistics Research Center, Cowell College, University of California, Santa Cruz.
- Itô, J. & R. A. Mester (1995). The core-periphery structure of the lexicon and constraints on reranking. In J. Beckman, S. Urbanczyk & L. Walsh (eds). *Papers in Optimality*

- Theory*. (University of Massachusetts Occasional Papers in Linguistics 18). Amherst, MA: Graduate Linguistic Student Association. 181-209.
- Itô, J. & R. A. Mester (1997a). Correspondence and compositionality: the Ga-gyō variation in Japanese phonology. In Roca, I. (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press. 419-462.
- Itô, J. & R. A. Mester (1997b). Sympathy Theory and German truncations. To appear in V. Miglio & B. Morén (eds). *Proceedings of the Hopkins Optimality Workshop/Maryland Mayfest 1997*. University of Maryland Working Papers in Linguistics 5. (ROA-211-0897, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Itô, J., R. A. Mester & J. Padgett (1995). Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26: 571-613. (ROA-38, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Jakobson, R. (1929). Remarques sur l'évolution phonologique du russe. *Travaux du Cercle Linguistique de Prague* 2. Reprinted (1962) in *Selected writings*. Vol. I: *Phonological studies*. The Hague: Mouton & Co. 7-116.
- Jakobson, R. (1936). Sur la théorie des affinités phonologiques. Report to the Fourth International Congress of Linguists, Copenhagen, August 1936. Reprinted (1962) in *Selected writings*. Vol. I: *Phonological studies*. The Hague: Mouton & Co. 234-246.
- Jordan, R. (1934). *Handbuch der mittellenglischen Grammatik. I. Teil: Lautlehre*. 2nd edn rev. Heidelberg. Translated and revised by Crook, J. (1974). *Handbook of Middle English Grammar: phonology*. The Hague/Paris: Mouton.
- Kager, R. (1995). Surface opacity of metrical structure in Optimality Theory. Ms., University of Utrecht. (ROA, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>). Forthcoming in Hermans, B. & M. van Oostendorp (eds). *The derivational residue in phonology*. Amsterdam: John Benjamins.
- Kager, R. (1997). Rhythmic vowel deletion in Optimality Theory. In Roca, I. (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press. 463-499.
- Kaisse, E. M. (1985). *Connected speech: The interaction of syntax and phonology*. Orlando: Academic Press.
- Kaisse, E. M. (1987). Rhythm and the cycle. *Papers from the Regional Meeting of the Chicago Linguistics Society* 23:2. 199-209.

- Kaisse, E. M. & P. A. Shaw (1985). On the theory of Lexical Phonology. *Phonology Yearbook* 2: 1-30.
- Kaisse, E. M. & S. Hargus (1993). Introduction. In S. Hargus & E. M. Kaisse (eds). *Studies in Lexical Phonology*. (Phonetics and Phonology 4) San Diego: Academic Press. 1-19.
- Kawasaki, H. (1986). Phonetic explanation for phonological universals: the case of distinctive vowel nasalization. In J. J. Ohala & J. J. Jaeger (eds). *Experimental phonology*. Orlando: Academic Press. 81-103.
- Kaye, J. (1974). Opacity and recoverability in phonology. *Canadian Journal of Linguistics* 19: 134-149.
- Kaye, J. (1975). A functional explanation of rule ordering in phonology. *Papers from the 11th Regional Meeting of the Chicago Linguistic Society: Papers from the parasession on functionalism*. 244-252.
- Kaye, J. (1990). 'Coda' licensing. *Phonology* 7: 301-330.
- Kaye, J. (1995). Derivations and interfaces. In Durand, J. & F. Katamba (eds). *Frontiers of phonology: atoms, structures, derivations*. London: Longman. 289-332.
- Kean, M.-L. (1974). The strict cycle in phonology. *Linguistic Inquiry* 5: 179-203.
- Keller, R. E. (1961). *German dialects: phonology and morphology, with selected texts*. Manchester: Manchester University Press.
- Kenstowicz, M. (1994). *Phonology in generative grammar*. Oxford: Blackwell.
- Kenstowicz, M. (1995). Cyclic vs. non-cyclic constraint evaluation. *Phonology* 12: 397-436.
- Kenstowicz, M. (1996). Base-identity and uniform exponence: alternatives to cyclicity. In J. Durand & B. Laks (eds). *Current trends in phonology: models and methods*. Salford: European Studies Research Institute. Vol. 1: 363-393.
- Kenstowicz, M. & C. Kisseberth (1977). *Topics in phonological theory*. New York: Academic Press.
- Kenstowicz, M. & C. Kisseberth (1979). *Generative phonology*. New York: Academic Press.
- Kim, M. (1993). On lengthening in the open syllables of Middle English. *Lingua* 91: 261-277.
- Kiparsky, P. (1968). How abstract is phonology? Bloomington: Indiana University Linguistics Club [Mimeograph]. Published (1973) in Fujimura, O. (ed.) *Three dimensions of linguistic theory*. Tokyo: TEC Company. 5-56. Reprinted in Kiparsky, P. (1982). *Explanation in phonology*. Dordrecht: Foris. 119-63.

- Kiparsky, P. (1970). Historical linguistics. In J. Lyons (ed.) *New horizons in linguistics*. New York: Penguin. 302-315.
- Kiparsky, P. (1971). Historical linguistics. In Dingwall, W. O. (ed.) *A survey of linguistic science*. College Park: Linguistics Program, University of Maryland. 576-649. Reprinted in Kiparsky, P. (1982b: 57-80).
- Kiparsky, P. (1973). Abstractness, opacity and global rules. Bloomington: Indiana University Linguistics Club. Also in Fujimura, O. (ed.) (1973). *Three dimensions of linguistic theory*. Tokyo: TEC Company. 57-86.
- Kiparsky, P. (1982a). From Cyclic Phonology to Lexical Phonology. In van der Hulst, H. & N. Smith (eds). *The structure of phonological representations*. Dordrecht: Foris. Part I: 131-75.
- Kiparsky, P. (1982b). *Explanation in phonology*. Dordrecht: Foris.
- Kiparsky, P. (1984). On the lexical phonology of Icelandic. In Elert, C.-C., I. Johansson & E. Strangert (eds.) *Nordic prosody III: papers from a symposium*. Stockholm: Almqvist & Wiksell. 135-64.
- Kiparsky, P. (1985). Some consequences of Lexical Phonology. *Phonology Yearbook 2*: 85-138.
- Kiparsky, P. (1988). Phonological change. In Newmeyer, F. J. (ed.) *Linguistics: the Cambridge survey*. Cambridge: Cambridge University Press. 363-415.
- Kiparsky, P. (1993). Blocking in nonderived environments. In Hargus, S. & E. M. Kaisse (eds). *Studies in Lexical Phonology*. (Phonetics and Phonology 4) San Diego: Academic Press. 277-313.
- Kiparsky, P. (1995). The phonological basis of sound change. In Goldsmith, J. A. (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell Publishers. 640-670.
- Kiparsky, P. (1998a). *Paradigm effects and opacity*. Ms., Stanford University.
- Kiparsky, P. (1998b). Sievers' Law as prosodic optimization. In J. Jasanoff, H. C. Melchert & L. Oliver (eds). *Mir Curad. Studies in honor of Calvert Watkins*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Kiparsky, P. (1998c). Covert generalization. In A. Rally (ed.) (forthcoming). *Proceedings of the First Mediterranean Morphology Conference*.
- Kirchner, R. (1995). Going the distance: synchronic chain shifts in Optimality Theory. Ms., UCLA. (ROA, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.)
- Kirchner, R. (1996). Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry 27*: 341-350.

- Kisseberth, C. (1970). On the functional unity of phonological rules. *Linguistic Inquiry 1*: 291-306.
- Kisseberth, C. (1973). Is rule ordering necessary in phonology? In Kachru, B. B., R. B. Lees, Y. Malkiel, A. Pietrangeli & S. Saporta (eds). *Issues in linguistics: papers in honor of Henry and Renée Kahane*. Urbana: University of Illinois Press. 418-441.
- Kraehenmann, A. (1998). Geminate are not all created equal. Paper presented at the 6th Manchester Phonology Meeting, Manchester, 21-23 May 1998.
- Kraska-Szlenk, I. (1995). *The phonology of stress in Polish*. PhD dissertation, University of Illinois, Urbana-Champaign.
- Labov, W. (1989). The exact description of the speech community: short a in Philadelphia. In R. Fasold & D. Schiffrin (eds). *Language change and variation*. Washington D. C.: Georgetown University Press. 1-57.
- Ladefoged, P. & I. Maddieson (1996). *The sounds of the world's languages*. Oxford: Blackwell.
- Lahiri, A. (1982). *Theoretical implications of analogical change: Evidence from Germanic languages*. PhD dissertation, Brown University.
- Lahiri, A. (1991). Stress and quantity in Middle Dutch and Middle English. Paper presented at Eurotyp, Salzburg, October 1991.
- Lahiri, A. (1995). Pervasion, simplification and optimization in language change. Paper presented at the XII International Conference on Historical Linguistics, Manchester, 15 August 1995.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In Lakatos, I. & A. Musgrave (eds). *Criticism and the growth of knowledge. Proceedings of the International Colloquium in the Philosophy of Science, London, 1965, volume 4*. Cambridge: Cambridge University Press. 91-196.
- Lakoff, G. (1993). Cognitive Phonology. In Goldsmith, J. A. (ed.) *The last phonological rule: Reflections on constraints and derivations*. Chicago: University of Chicago Press. 117-145.
- Lakoff, R. (1972). Another look at drift. In Stockwell, R. P. & R. K. S. Macaulay (eds.) *Linguistic change and generative theory: essays from the UCLA Conference on Historical Linguistics in the Perspective of Transformational Theory*. Bloomington: Indiana University Press. 172-198.
- Lass, R. (1974). Linguistic orthogenesis? Scots vowel quantity and the English length conspiracy. In Anderson, J. M. & C. Jones (eds.) *Historical linguistics: proceedings*

- of the First International Conference on Historical Linguistics, Edinburgh, 2-7 September 1973. Amsterdam: North-Holland. Vol II: 311-352.
- Lass, R. (1980). *On explaining language change*. Cambridge: Cambridge University Press.
- Lass, R. (1985). Minkova noch einmal: MEOSL and the resolved foot. *Folia Linguistica Historica* 6: 245-266.
- Lass, R. (1987). Language, speakers, history and drift. In Koopman, W. F., R. Lass, O. Fischer & R. Eaton (eds.) *Explanation and linguistic change*. Amsterdam: Benjamins. 151-176.
- Lass, R. (1992). Phonology and morphology. In Blake, N. (ed.) *The Cambridge history of the English language*. Vol. II: 1066-1476. Cambridge: Cambridge University Press. 23-155.
- Lass, R. (1994). *Old English: a historical linguistic companion*. Cambridge: Cambridge University Press.
- Lass, R. (1997). *Historical linguistics and language change*. Cambridge: Cambridge University Press.
- Lee, B.-G. (1996). The emergence of the faithful. Ms. Seoul University/University of Massachusetts, Amherst. (ROA-167-1296, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Levin, J. (1985). *A metrical theory of syllabicity*. PhD dissertation, MIT.
- Lightfoot, D. (1989). The child's trigger experience: degree-0 learnability. *Behavioural and Brain Sciences* 12: 321-375.
- Lightfoot, D. (1991). *How to set parameters: arguments from language change*. Cambridge, MA: MIT Press.
- Lightfoot, D. (1999). *The development of language: acquisition, change, and evolution*. Oxford: Blackwell.
- Lin, Y.-H. (1997). Syllabic and moraic structures in Piro. *Phonology* 14: 403-436.
- Lindau, M. (1975). A phonetic explanation to reduced vowel harmony systems. *Working Papers (Lund University)* 11: 43-54.
- Lindau, M. (1978). Vowel features. *Language* 54: 541-563.
- Lowenstamm, J. & J. Kaye (1985). Compensatory lengthening in Tiberian Hebrew. In L. Wetzels & E. Sezer (eds.) *Studies in compensatory lengthening*. Dordrecht: Foris. 97-132.

- Lubowicz, A. (1998). Derived environment effects in OT. Ms., University of Massachusetts, Amherst. (ROA-239-0198, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Lutz, A. (1986). The syllabic basis of word division in Old English manuscripts. *English Studies* 3: 193-210.
- Luick, K. (1898). Beiträge zur englischen Grammatik, III: Die Quantitätsveränderungen im Laufe der englischen Sprachentwicklung. *Anglia* 20: 335-362.
- Luick, K. (1964 [1914-1940]). *Historische Grammatik der englischen Sprache*. Stuttgart: Bernhard Tauchnitz / Oxford: Basil Blackwell.
- Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- Maddieson, I. (1993). Splitting the mora. *UCLA Working Papers in Phonetics* 83: 9-18.
- Maddieson, I. & P. Ladefoged (1993). Phonetics of partially nasal consonants. In Huffman, M. & R. Krakow (eds.) *Nasals, nasalization, and the velum*. (Phonetics and Phonology 5). San Diego: Academic Press. 251-301.
- Malone, J. (1993). *Tiberian Hebrew phonology*. Winona Lake, Indiana: Eisenbrauns.
- Marchand, H. (1960). *The categories and types of Present-day English word-formation: a synchronic-diachronic approach*. Wiesbaden: Otto Harrassowitz.
- Marchand, J. W. (1973). *The sounds and phonemes of Wulfila's Gothic*. The Hague: Mouton.
- Marcus, M. P. (1980). *A theory of syntactic recognition for natural language*. Cambridge, MA: MIT Press.
- Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition* 46: 53-85.
- Martin, J. (1992). In support of the Domino Condition. Ms., Program in Linguistics, University of Michigan, Ann Arbor.
- Mascaró, J. (1976). *Catalan phonology and the phonological cycle*. PhD dissertation, MIT. Distributed by Indiana University Linguistics Club.
- McCarthy, J. J. (1998). Sympathy and phonological opacity. Ms., University of Massachusetts, Amherst. (ROA-252-0398, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- McCarthy, J. J. (to appear). Sympathy, cumulativity, and the Duke-of-York gambit. In C. Féry & R. v. d. Vijver. *The optimal syllable*. Cambridge: Cambridge University Press. (Available from <http://www-unix.oit.umass.edu/~jjmccart/cumul.pdf>.)
- McCarthy, J. J. & A. Prince (1986). *Prosodic morphology*. Ms., University of Massachusetts, Amherst, and Brandeis University.

- McCarthy, J. J. & A. Prince (1988). Quantitative transfer in reduplicative and templatic morphology. In Linguistic Society of Korea (ed.) *Linguistics in the morning calm* vol. 2. Seoul: Hanshin Publishing Co. 3-35.
- McCarthy, J. J. & A. Prince (1993a). *Prosodic Morphology I: constraint interaction and satisfaction*. Rutgers University Center for Cognitive Science Technical Report #2.
- McCarthy, J. J. & A. Prince (1993b). Generalized alignment. In G. Booij & J. van Marle (eds). *Yearbook of Morphology 1993*. Dordrecht: Kluwer. 79-153. Also Rutgers University Center for Cognitive Science Technical Report #7.
- McCarthy, J. J. & A. Prince (1994). The emergence of the unmarked: optimality in prosodic morphology. In M. González (ed.) *Proceedings of the North East Linguistic Society 24*. Amherst: Graduate Linguistic Student Association. 333-379. (ROA-13, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.)
- McCarthy, J. J. & A. Prince (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Walsh-Dickey & S. Urbanczyk (eds). *Papers in Optimality Theory*. (University of Massachusetts Occasional Papers in Linguistics 18) Amherst: Graduate Linguistic Student Association. 249-384. (ROA-60, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.)
- McMahon, A. M. S. (1991). Lexical phonology and sound change: the case of the Scottish Vowel Length Rule. *Journal of Linguistics* 27: 29-53.
- McMahon, A. M. S. (1994). *Understanding language change*. Cambridge: Cambridge University Press.
- Mester, R. A. (1994). The quantitative trochee in Latin. *Natural Language and Linguistic Theory* 12: 1-61.
- Milroy, J. (1992). *Linguistic variation and change: on the historical sociolinguistics of English*. Oxford: Blackwell.
- Milroy, J. & L. Milroy (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21: 339-384.
- Minkova, D. (1982). The environment for Open Syllable Lengthening in Middle English. *Folia Linguistica Historica* 3: 29-58.
- Minkova, D. (1985). Of rhyme and reason: some foot-governed quantity changes in English. In Eaton, R. et al. (eds). *Papers from the 4th International Conference on English Historical Linguistics*. Amsterdam: John Benjamins. 163-178.
- Minkova, D. (1997). Constraint ranking in Middle English stress-shifting. *English Language and Linguistics* 1: 135-175.

- Minkova, D. & R. P. Stockwell (1996). The origins of long-short allomorphy in English. Paper presented at the IX International Conference on English Historical Linguistics, Poznań, 26 August 1996.
- Mohanan, K. P. (1982). *Lexical Phonology*. PhD dissertation, MIT. Distributed by Indiana University Linguistics Club.
- Mohanan, K. P. (1986). *The theory of Lexical Phonology*. Dordrecht: D. Reidel.
- Morén, B. T. (1998). The puzzle of Kashmiri stress: implications for weight theory. Ms., University of Maryland. (ROA-273-0798, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Morin, Y.-C. (1992). Phonological interpretations of historical lengthening. Paper presented at the Seventh International Phonology Meeting, Krems.
- Morsbach, L. (1896). *Mittelenglische Grammatik*. Halle: Max Niemeyer.
- Mossé, F. (1949). *Manuel de l'anglais du moyen âge des origines au XIV^e siècle*. Vol. II: *Moyen-Anglais*. Paris: Éditions Montaigne. Translated by Walker, J. A. (1952). *A handbook of Middle English*. Baltimore/London: The John Hopkins University Press.
- Murray, Robert W. (1986). Urgermanische Silbenstruktur und die westgermanische Konsonantengemination. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 108: 333-356.
- Murray, R. W. & T. Vennemann (1983). Sound change and syllable structure in Germanic phonology. *Language* 59: 514-528.
- Myers, S. (1987). Vowel shortening in English. *Natural Language and Linguistic Theory* 5: 485-518.
- Myers, S. (1997). Expressing phonetic naturalness in phonology. In Roca, I. (1997). *Derivations and constraints in phonology*. Oxford: Clarendon Press. 125-152.
- Neef, M. (1996). *Wordesign. Eine deklarative Analyse der deutschen Verflexion*. Tübingen: Stauffenburg Verlag.
- Nesse, R. M. & G. C. Williams (1998). Evolution and the origins of disease. *Scientific American*. November, 279(5): 58-65.
- Nigoyi, P. & R. C. Berwick (1997). Evolutionary consequences of language learning. *Linguistics and Philosophy* 20: 697-719.
- Noske, R. (1992). Moraic vs. constituent syllables. In Eisenberg, P., K. H. Ramers & H. Vater (eds.) *Silbenphonologie des Deutschen*. Tübingen: Gunter Narr. 284-328.

- Odden, D. (1993). Interaction between modules in Lexical Phonology. In Hargus, S. & E. M. Kaisse (eds). *Studies in Lexical Phonology*. (Phonetics and Phonology 4) San Diego: Academic Press. 277-313.
- Ohala, J. J. (1989). Sound change is drawn from a pool of synchronic variation. In Breivik, L. E. & E. H. Jahr (eds). *Language change: contributions to the study of its causes*. Berlin: Mouton de Gruyter. 173-198.
- Ohala, J. J. (1992). What's cognitive, what's not, in sound change. In Kellermann, G. & M. D. Morrissey (eds). *Diachrony within synchrony: language history and cognition. Papers from the International Symposium at the University of Duisburg, 26-28 March 1990*. Frankfurt am Main: Peter Lang. 309-355.
- Ohala, J. J. (1993). The phonetics of sound change. In Jones, C. (ed.) *Historical linguistics: problems and perspectives*. London/New York: Longman. 237-278.
- Orgun, C. O. (1994). Monotonic cyclicity. Ms., University of California, Berkeley. (ROA, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Orgun, C. O. (1996a). *Sign-Based Morphology and Phonology, with special attention to Optimality Theory*. PhD dissertation, University of California, Berkeley.
- Orgun, C. O. (1996b). Sign-Based Morphology: a declarative theory of phonology-morphology interleaving. Ms., University of California, Berkeley. (ROA-122-0496, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Padgett, J. (1995). *Structure in Feature Geometry*. Stanford: CSLI Publications.
- Parkes, M. B. (1983). On the presumed date of the manuscript of the *Ormulum*: Oxford, Bodleian Library, MS Junius 1. In Stanley, E. G. & D. Gray (eds). *Five hundred years of words and sounds: a festschrift for Eric Dobson*. Cambridge: D. S. Brewer. 115-127.
- Pater, J. (1995). On the non-uniformity of weight-to-stress and stress preservation effects in English. Ms., McGill University. ROA-107, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.
- Pesetsky, D. (1979). Russian morphology and lexical theory. Ms., MIT.
- Pollard, C & I. Sag (1994). *Head-driven phrase structure grammar*. Chicago: CSLI Publications and University of Chicago Press.
- Pope, M. K. (1934). *From Latin to Modern French with special consideration of Anglo-Norman. Phonology and morphology*. (Revised edn 1952.) Manchester: Manchester University Press.

- Potter, B. (1994). Serial optimality in Mohawk prosody. *Papers from the Annual Regional Meeting of the Chicago Linguistics Society* 13: 347-361.
- Prince, A. (1990). Quantitative consequences of rhythmic organization. *Chicago Linguistics Society* 26.2: 355-395.
- Prince, A. & P. Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms., Rutgers University Center for Cognitive Science Technical Report #2.
- Prokosch, E. (1938). *A comparative Germanic grammar*. Baltimore: Linguistic Society of America.
- Pulleyblank, D. (1994). Underlying mora structure. *Linguistic Inquiry* 25: 344-353.
- Pulleyblank, D. (1997). Optimality Theory and features. In Archangeli, D. & D. T. Langendoen (eds). *Optimality Theory: an overview*. Oxford: Blackwell. 59-101.
- Pulleyblank, D. & W. J. Turkel (1996). Optimality Theory and learning algorithms: the representation of recurrent featural asymmetries. In Durand, J & B. Laks (eds). *Current trends in phonology: models and methods*. Salford: European Studies Research Institute. 655-686.
- Pullum, G. (1976). The Duke of York gambit. *Journal of Linguistics* 12: 83-102.
- Raffelsiefen, R. (1996). A constraint-based explanation of stability effects in historical phonology. Paper presented at the Fourth Phonology Meeting, Manchester, 17 May 1996.
- Riad, T. (1992). *Structures in Germanic prosody: a diachronic study with special reference to the Nordic languages*. Stockholm: Department of Scandinavian languages, Stockholm University.
- Riad, T. (1995). The quantity shift in Germanic: a typology. *Amsterdamer Beiträge zur älteren Germanistik* 42: 159-184.
- Rialland, A. (1991). L'allongement compensatoire: Nature et modèles. In Laks, B. & A. Rialland (eds). *Architecture des représentations phonologiques*. Presse du CNRS.
- Ritt, N. (1994). *Quantity adjustment: vowel lengthening and shortening in early Middle English*. Cambridge: Cambridge University Press.
- Ritt, N. (1997). Now you see it, now you don't: Middle English lengthening in closed syllables. *Studia Anglica Posnaniensia* 32: 249-270.
- Ritzert, A. (1898). Die Dehnung der mhd. kurzen Stammsilbenvokale in den Volksmundarten des hochdeutschen Sprachgebietes auf Grund der vorhandenen

- Dialektliteratur. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 23: 131-222.
- Roca, I. (1986). Secondary stress and metrical rhythm. *Phonology Yearbook* 3: 341-370.
- Roca, I. (1991). Stress and syllables in Spanish. In Campos, H. & F. Martínez-Gil (eds). *Current studies in Spanish linguistics*. Washington D.C.: Georgetown University Press. 599-635.
- Roca, I. (1997). There are no "glides", at least in Spanish: an optimality account. *Probus* 9: 233-265.
- Rubach, J. (1996). Shortening and ambisyllabicity in English. *Phonology* 13: 197-237.
- Rubach, J. (1997). Extrasyllabic consonants in Polish: Derivational Optimality Theory. In Roca, I. (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press. 551-581.
- Rubach, J. & G. Booij (1990). Edge of constituent effects in Polish. *Natural Language and Linguistic Theory* 7: 121-158.
- Sapir, E. (1921). *Language*. New York: Harcourt, Brace & Co. Republished (1963) London: Rupert Hart-Davis Ltd.
- Sarrazin, G. (1898). Mittelenglische Vokaldehnung in offener Silbe und Streitbergs Dehnungsgesetz. *Archiv für das Studium der neueren Sprachen und Literaturen* 101: 65-86.
- Sauer, H. (1992). *Nominalkomposita im Frühmittelenglischen: mit Ausblicken auf die Geschichte der englischen Nominalkomposition*. Tübingen: Max Niemeyer Verlag.
- Schane, S. A. (1995). Diphthongization in Particle Phonology. In J. A. Goldsmith (ed.) *The handbook of phonological theory*. Oxford: Blackwell. 586-608.
- Schulze, W. (1908). Wortbrechung in den gotischen Handschriften. *Sitzungsberichte der Preußischen Akademie der Wissenschaften, Phil.-his. Klasse*, 610-24. Berlin. Reprinted (1966) in W. Schulze. *Kleine Schriften*. Göttingen: Vandenhoeck & Ruprecht.
- Scott, N. C. (1957). Notes on the pronunciation of Sea Dayak. *Bulletin of the School of Oriental and African Studies*. 20: 509-512.
- Selkirk, E. O. (1984a). On the major class features and syllable theory. In Aronoff & Oehle (eds). *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students*. Cambridge, Mass.: MIT Press. 107-136.
- Selkirk, E. O. (1984b). *Phonology and syntax: the relation between sound and structure*. Cambridge, MA: MIT Press.

- Sherer, T. D. (1994). *Prosodic phonotactics*. PhD dissertation, University of Massachusetts, Amherst. (ROA-54-03951, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Siegel, D. (1974). *Topics in English morphology*. PhD dissertation, MIT.
- Sievers, E. (1878). Zur Accent- und Lautlehre der germanischen Sprachen: II. Die Behandlung unbetonter Vokale. *Beiträge zur Geschichte der Deutschen Sprache und Literatur* 5: 63-163.
- Simmler, F. (1974). *Die westgermanische Konsonantengemination im Deutschen unter besonderer Berücksichtigung des Althochdeutschen*. München: Wilhelm Fink Verlag.
- Smolensky, P. (1993). Harmony, markedness, and phonological activity. Paper presented at Rutgers Optimality Workshop 1, Rutgers University, 23 October 1993. (ROA-87, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27. (ROA-118, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Spaelti, P. (1994). Weak edges and final geminates in Swiss German. *Papers from the Annual Meeting of the North East Linguistic Society* 24: 573-588. (ROA-18, Rutgers Optimality archive, <http://ruccs.rutgers.edu/roa.html>).
- Springer, O. (1975). Early Runic evidence of Sievers' Law. In O. Springer *Arbeiten zur germanischen Philologie und zur Literatur des Mittelalters*. München: Fink. 164-177.
- Sproat, R. (1985). *On deriving the lexicon*. PhD dissertation, MIT.
- Sprouse, R. (1996). Vowels that borrow moras: geminates and weight in OT. *Papers from the Annual Meeting of the North East Linguistic Society* 26: 393-408.
- Sprouse, R. (1997). A case for Enriched Inputs. Handout of paper presented at Trend VII, University of California Santa Cruz, 3 May 1997. (ROA-193-0597, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>).
- Stampe, D. (1972). *How I spent my summer vacation [A dissertation on Natural Phonology.]* PhD dissertation, University of Chicago. Published (1980) New York: Garland.
- Stanley, R. (1973). Boundaries in phonology. In Anderson, S. R. & P. Kiparsky (eds). *A Festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston. 185-206.
- Steriade, D. (1982). *Greek prosodies and the nature of syllabification*. PhD dissertation, MIT.

- Steriade, D. (1996). Paradigm uniformity and the phonetics-phonology boundary. Paper presented at the 5th Conference in Laboratory Phonology, Evanston, IL. Ms., UCLA.
- Stewart, I. (1989). *Does God play dice? The new mathematics of chaos*. 2nd edn 1997. London: Penguin.
- Stewart, J. M. (1971). Niger-Congo, Kwa. In Sebeok, T. A. (ed.) *Current trends in linguistics VII*. The Hague: Mouton. 179-212.
- Stockwell, R. P., J. D. Bowen & I. Silva-Fuenzalida (1956). Spanish juncture and intonation. *Language* 32. 641-665.
- Struijke, C. (1998). Reduplicant and output TETU in Kwakwala. In Fukuzawa, H., F. Morelli, C. Struijke & Y. Su (eds). *Papers in phonology*. (University of Maryland Working Papers 7). 150-178. (ROA-261-0598, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.)
- Sweet, H. (1888). *A history of English sounds from the earliest period, with full word-lists*. Oxford: Clarendon Press.
- Tranel, B. (1991). CVC light syllables, geminates and Moraic Theory. *Phonology* 8: 291-302.
- Vennemann, T. (1972). Rule inversion. *Lingua* 29: 209-242.
- Vennemann, T. (1983). Causality in language change: theories of linguistic preferences as a basis for linguistic explanations. *Folia Linguistica Historica* 4: 5-26.
- Vennemann, T. (1987) Muta cum Liquida: Worttrennung und Syllabierung im Gotischen: Mit einem Anhang zur Worttrennung in der Pariser Handschrift der althochdeutschen Isidor-Übersetzung. *Zeitschrift für deutsches Altertum und deutsche Literatur* 98: 165-204.
- Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change, with special reference to German, Germanic, Italian and Latin*. Berlin: Mouton de Gruyter.
- Vincent, N. (1978). Is sound change teleological? In Fisiak, J. (ed.) *Recent developments in historical phonology*. Mouton: The Hague. 409-429.
- Walker, R. (1994). Buriat syllable weight and head prominence. In Merchant, J., J. Padgett & R. Walker (eds). *Phonology at Santa Cruz* 3. University of California at Santa Cruz, Linguistics Research Center. 99-110. (ROA-142-0996, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.)
- Wang, W. S.-Y. (1969). Competing changes as a cause of residue. *Language* 45: 9-25.

- Wells, J. C. (1990). *Longman pronunciation dictionary*. Harlow: Longman.
- Wetzel, C.-D. (1981). *Die Worttrennung am Zeilenende in altenglischen Handschriften*. Frankfurt: Lang.
- Wetzels, L. (1985). Phonological timing in Ancient Greek. In L. Wetzels & E. Sezer (eds). *Studies in compensatory lengthening*. Dordrecht: Foris. 279-344.
- Withgott, M. M. (1983). *Segmental evidence for phonological constituents*. PhD dissertation, University of Texas, Austin.
- Wilbur, R. (1973). *The phonology of reduplication*. PhD dissertation, University of Illinois, Urbana-Champaign.
- Williamson, K. (1983). Vowel merger in harmony languages. *Journal of the Linguistic Association of Nigeria* 2: 61-82.
- Winitz, H., M. E. Scheib & J. A. Reeds (1972). Identification of stops and vowels for the burst portion of /p,t,k/ isolated from conversational speech. *Journal of the Acoustical Society of America* 51: 1309-1317.
- Winston, F. D. D. (1962). The nominal class system of Lokɔ. *African Language Studies* 3: 49-70.
- Wright, J. & E. M. Wright (1923). *An elementary Middle English grammar*. London: Oxford University Press.
- Zamora, J., & J. Guitart (1982). *Dialectología hispanoamericana: Teoría, descripción, historia*. Salamanca: Ediciones Almar.
- Zec, D. (1988). *Sonority constraints on prosodic structure*. PhD dissertation, Stanford University. Published (1994) New York: Garland.
- Zec, D. (1993). Rule domains and phonological change. In S. Hargus & E. M. Kaisse (eds). *Studies in Lexical Phonology*. (Phonetics and Phonology 4) San Diego, CA: Academic Press. 365-405.
- Zec, D. (1995a). The role of moraic structure in the distribution of segments within syllables. In Durand, J. & F. Katamba (eds). *Frontiers of phonology: atoms, structures, derivations*. London: Longman. 149-179.
- Zec, D. (1995b). Sonority constraints on syllable structure. *Phonology* 12: 85-129.
- Zoll, C. (1998). Positional asymmetries and licensing. Ms., MIT = Expanded handout of talk presented at the LSA meeting, New York, January 1998. (ROA-282-0998, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.)